5th International ARTIFICIAL INTELLIGENCE AND DATA SCIENCE Congress 2025

5. ULUSLARARASI YAPAY ZEKÂ VE VERİ BİLİMİ KONGRESİ

Proceedings Book:

ABSTRACT AND FULL PAPERS

Bildiriler Kitabı: Özetler ve Tam Metinler

Yayınevi: İZMİR KÂTİP ÇELEBİ ÜNİVERSİTESİ YAYINLARI
Publisher: IZMIR KATIP CELEBI UNIVERSITY PRESS
Yayın No. / Publication Number: 50

5TH INTERNATIONAL ARTIFICIAL INTELLIGENCE AND DATA SCIENCE CONGRESS (ICADA 2025) FULL PAPERS AND ABSTRACTS
5. ULUSLARARASI YAPAY ZEKÂ VE VERİ BİLİMİ KONGRESİ TAM METİNLER VE ÖZETLER

Editörler / Editors
Prof. Dr. Ayşegül ALAYBEYOĞLU
Dr. Öğr. Üyesi Semih ÇAKIR

Yardımcı Editör / Associate Editor
Öğr. Gör. Dr. Ümit SARP

Tasarım / Design
Büşra UYAR

E-ISBN
978-625-95496-4-4

Basım / Printing
Necmettin Erbakan University Press (NEU PRESS)
Yaka Mah. Yeni Meram Cad. Kasım Halife Sok. B Blok
No: 11 Meram / KONYA / TÜRKİYE
0332 221 0 575 - www.neupress.org

Sertifika No / Certificate No
48888

İKÇÜ YAYINEVİ 1. Baskı, İzmir/TÜRKİYE
Haziran 2025, ii + 839 sf., 21x29.7 cm

Tüm Hakları Saklıdır. Bu kitabın basım, yayın, satış hakları İKÇÜ Yayınevinin izni alınmadan kitabın tümü ya da bölümleri herhangi bir bilgi depolama ve erişim sistemiyle veya mekanik, elektronik, fotokopi, manyetik ya da başka yöntemlerle çoğaltılamaz.

All Rights Reserved. All or parts of this book cannot be reproduced by any information storage and access system or by mechanical, electronic, photocopying, magnetic or other methods without the permission of IKCU Publishing House.

PREFACE

As İzmir **Kâtip Çelebi University Artificial Intelligence and Data Science Application and Research Center**, our goal is to contribute to the development of our country in the field of artificial intelligence by organizing congresses, journals, educational activities, scientific events, project initiatives, and collaborations in line with the **"Digital Türkiye" vision** and **"National Technology Initiative"** development goals outlined in the "National Artificial Intelligence Strategy" prepared by the Presidency of the Republic of Türkiye Digital Transformation Office and the Ministry of Industry and Technology.

In line with these goals, we are pleased to have organized the fifth edition of the International Congress on Artificial Intelligence and Data Science, which we host annually as İzmir Katip Çelebi University Artificial Intelligence and Data Science Application and Research Center, in collaboration with Zonguldak Bülent Ecevit University Karaelmas Cyber Security Application and Research Center this year.

We believe that this collaboration between two different university research centers which focused on **"Artificial Intelligence"** and "**Cyber Security"** areas that have gained significant importance both in our country and globally due to rapid technological advancements—is highly significant as it marks a first of its kind in our country. We also believe that this cooperation will lead to important academic and sectoral outcomes.

In line with the roadmap outlined by the President of the Council of Higher Education (YÖK), Prof. Dr. Erol Özvar, who stated that "**Turkish universities have a very important responsibility regarding the production of artificial intelligence,**" we, as İzmir Katip Çelebi University Artificial Intelligence and Data Science Application and Research Center, would like to emphasize our continued commitment to contributing to our country's progress in the field of artificial intelligence through our work.

We would like to express our deepest gratitude to everyone who supported our congress, which featured scientific studies from numerous countries and universities spanning from the United States to the Philippines. Special thanks go to İzmir Katip Çelebi University Rector Prof. Dr. Saffet Köse and Zonguldak Bülent Ecevit University Rector Prof. Dr. İsmail Hakkı Özölçer, as well as to our invited speakers: YÖK Member Prof. Dr. Derviş Karaboğa, Prof. Dr. Tülay Yildirim, Prof. Dr. Sabina Barakovic, Prof. Dr. Bijan Najafi, Prof. Dr. Kunal Mankodiya, Assoc. Prof. Dr. Ali Turab, and our panelist speakers.

We would also like to sincerely thank our Organizing Committee Team, who worked with great effort and enthusiasm throughout the organization process, and our Sponsors for their contributions. We would be delighted to invite you to our next congress.

Congress Chair
Prof. Dr. Ayşegül Alaybeyoğlu

PREFACE

In today's world, where digitalization is increasingly integrated into every aspect of life, **cybersecurity** has become a strategic priority for the protection of both individuals and institutions. The growing variety and complexity of threats targeting information systems necessitate a multidisciplinary approach to research and application in this field. Particularly, the effective use of **artificial intelligence** technologies in cybersecurity brings about both new opportunities and emerging risks.

Focusing on this critical intersection, the "5th International Artificial Intelligence and Data Science Congress" conference was organized under the auspices of **Zonguldak Bülent Ecevit University**, in collaboration with the **Karaelmas Cybersecurity Application and Research Center** and the İzmir Kâ**tip Çelebi University Artificial Intelligence and Data Science Application and Research Center**. The aim of the conference is to bring together academic research and sectoral experience to assess the current landscape of cybersecurity and to discuss innovative solutions powered by artificial intelligence.

This **proceedings book** presents a selection of papers carefully reviewed and presented during the conference. Covering a wide range of crucial topics—such as threat detection, intrusion prevention, network security, cryptography, AI-based defense mechanisms, and ethical considerations—this volume offers valuable contributions to both theoretical knowledge and practical implementation.

We sincerely thank all the researchers who contributed their work, the reviewers who meticulously evaluated the papers, and everyone who played a role in the organization of this event. We hope that this conference and the proceedings book will inspire future research and collaborations in the fields of cybersecurity and artificial intelligence.

Sincerely,

Congress Organizing Committee Chair
Assist. Prof. Dr. Semih Çakır

ÖNSÖZ

İzmir Kâ**tip Çelebi Üniversitesi Yapay Zeka ve Veri Bilimi Uygulama ve Araştırma Merkezi** olarak hedefimiz Cumhurbaşkanlığı Dijital Dönüşüm Ofisi Başkanlığı ve Sanayi ve Teknoloji Bakanlığı tarafından hazırlanan **"Ulusal Yapay Zeka Stratejisi"** nde belirtilen **"Dijital Türkiye"** vizyonu ve **"Milli Teknoloji Hamlesi"** kalkınma hedefleri doğrultusunda kongre, dergi, eğitim, bilimsel etkinlikler, proje faaliyetleri ve işbirlikleri gerçekleştirerek ülkemizin yapay zekâ alanındaki gelişim sürecine katkı sağlamaktır.

Bu hedefler doğrultusunda; İzmir Katip Çelebi Üniversitesi Yapay Zeka ve Veri Bilimi Uygulama ve Araştırma Merkezi olarak her sene düzenlemiş olduğumuz Uluslararası Yapay Zeka ve Veri Bilimi Kongresinin beşincisini Zonguldak Bülent Ecevit Üniversitesi Karaelmas Siber Güvenlik Uygulama ve Araştırma Merkezi iş birliği ile gerçekleştirmiş olmanın mutluluğunu yaşamaktayız.

Özellikle de son dönemlerde teknolojinin hızla gelişmesi ile birlikte tüm Dünyada ve Ülkemizde büyük önem kazanan **"Yapay Zeka"** ve **"Siber Güvenlik"** alanlarında araştırma yapan iki farklı üniversite merkezinin işbirliği içinde olmasının ülkemizde bir ilk olması yönüyle de büyük önem taşıdığı ve bu iş birliğinin önemli akademik ve sektörel çıktılarının olacağı inancındayız.

YÖK Başkanımız Sayın Prof. Dr. Erol Özvar'ın **"yapay zekâyı üretmekle alakalı Türkiye'de üniversitelerin önünde çok önemli bir ödev bulunduğuna"** dair sözleri ile çizmiş olduğu yol haritası doğrultusunda, biz de İzmir Kâtip Çelebi Üniversitesi Yapay Zeka ve Veri Bilimi Uygulama ve Araştırma Merkezi olarak ülkemizin yapay zeka alanındaki gelişim sürecine çalışmalarımızla katkı sağlamaya devam edeceğimizi belirtmek isteriz.

Amerika'dan Filipinlere kadar uzanan birçok farklı ülke ve üniversitelerden bilimsel çalışmaların sunulduğu kongremize desteklerinden dolayı başta İzmir Katip Çelebi Üniversitesi Rektörü Sayın Prof. Dr. Saffet Köse ve Zonguldak Bülent Ecevit Üniversitesi Rektörü Sayın Prof. Dr. İsmail Hakkı Özölçer olmak üzere Davetli Konuşmacılarımız YÖK Üyesi Sayın Prof. Dr. Derviş Karaboğa'ya, Sayın Prof. Dr. Tülay Yıldırım'a, Sayın Prof. Dr. Sabina Barakovıc'e, Sayın Prof. Dr. Bijan Najafi'ye, Sayın Prof. Dr. Kunal Mankodıya'ya, Sayın Doç. Dr. Ali Turab'a ve Panelist Konuşmacılarımıza en derin şükranlarımızı sunarız.

Kongrenin düzenlenme sürecinde çok emek sarf eden ve büyük bir heyecanla çalışan Düzenleme Kurulu Ekibimize ve katkılarından dolayı Sponsorlarımıza çok teşekkür eder, bir sonraki kongremize sizleri davet etmekten büyük bir memnuniyet duyarız.

Kongre Başkanı
Prof. Dr. Ayşegül Alaybeyoğlu

ÖNSÖZ

Dijitalleşmenin her geçen gün daha fazla hayatımıza entegre olduğu günümüzde, siber güvenlik alanı hem bireylerin hem de kurumların güvenliği açısından stratejik bir öncelik haline gelmiştir. Bilgi sistemlerine yönelik tehditlerin çeşitlenmesi ve karmaşıklığının artması, bu alandaki araştırma ve uygulamaların disiplinler arası bir bakış açısıyla ele alınmasını zorunlu kılmaktadır. Özellikle yapay zekâ teknolojilerinin siber güvenlik uygulamalarında etkin bir biçimde kullanılmaya başlanması, hem fırsatları hem de yeni riskleri beraberinde getirmektedir.

Bu önemli kesişim noktasını merkezine alan "5th International Artificial Intelligence and Data Science Congress" kongresi, Zonguldak Bülent Ecevit Üniversitesi ev sahipliğinde; **Karaelmas Siber Güvenlik Uygulama ve Araştırma Merkezi** ile İzmir **Kâtip Çelebi Üniversitesi Yapay Zeka ve Veri Bilimi Uygulama ve Araştırma Merkezi** iş birliğinde düzenlenmiştir. Kongre, akademik araştırmaları ve sektörel deneyimleri bir araya getirerek, siber güvenliğin mevcut durumunu değerlendirmek ve yapay zekâ destekli yeni çözüm önerilerini tartışmak amacı taşımaktadır.

Elinizde bulunan bu bildiri kitabı, kongre kapsamında sunulan ve titizlikle seçilmiş bildirileri içermektedir. Siber güvenlikte tehdit algılama, saldırı önleme, ağ güvenliği, kriptografi, yapay zekâ temelli siber savunma mekanizmaları ve etik sorunlar gibi pek çok önemli konuyu kapsayan bu eser; hem teorik altyapıya hem de uygulamaya yönelik katkılar sunmaktadır.

Bu değerli çalışmalara katkı sunan tüm araştırmacılara, bildirileri bilimsel titizlikle değerlendiren hakemlere ve organizasyon sürecinde emeği geçen herkese içten teşekkürlerimizi sunarız. Kongremizin ve bu bildiri kitabının, siber güvenlik ve yapay zekâ alanlarında yapılacak yeni çalışmalara ilham vermesini diliyoruz.

Saygılarımla,

Kongre Düzenleme Kurulu Başkanı
Dr. Öğr. Üyesi Semih Çakır

BOARDS & COMMITEES

KURULLAR VE KOMİTELER

HONORARY PRESIDENT (Onursal Başkanlar)

Prof. Dr. Saffet Köse
Rector of İzmir Kâtip Çelebi University

Prof. Dr. İsmail Hakkı Özölçer
Rector of Zonguldak Bülent Ecevit University

CONGRESS PRESIDENT (Kongre Başkanı)

Prof. Dr. Ayşegül Alaybeyoğlu
İzmir Kâtip Çelebi University

PRESIDENT OF ORGANIZING COMMITTEE (Düzenleme Kurulu Başkanı)

Assist. Prof. Dr. Semih Çakır
Zonguldak Bülent Ecevit University

SCIENTIFIC SECRETARIAT (Bilimsel Sekreter)

Instructor Dr. Sümeyye Sınır
İzmir Kâtip Çelebi University

TECHNICAL SUPPORT (Teknik Destek)

Ali Gaffar Toksoy
Zonguldak Bülent Ecevit University

ORGANIZING COMMITTEE (Düzenleme Kurulu)

Prof. Dr. Şenol Hakan Kutoğlu
Zonguldak Bülent Ecevit University Vice Rector

Prof. Dr. Ayşegül Alaybeyoğlu
İzmir Kâtip Çelebi University

Prof. Dr. Yılmaz Yıldırım
Zonguldak Bülent Ecevit University

Prof. Dr. Merih Palandöken
İzmir Kâtip Çelebi University

Prof. Dr. Aytuğ Onan
İzmir Kâtip Çelebi University

Prof. Dr. Mustafa Alkan
President of Information Security Society

Dr. Nilgün Özcan
GSB General Directorate of Education Research and Coordination

Assoc. Prof. Dr. Ali Narin
Zonguldak Bülent Ecevit University

Assoc. Prof. Dr. Levent Aydın
İzmir Kâtip Çelebi University

Assoc. Prof. Dr. Mehmet Ali Belen
İskenderun Technical University

Asst. Prof. Dr. Ceren Kaya
Zonguldak Bülent Ecevit University

Asst. Prof. Dr. Semih Çakır
Zonguldak Bülent Ecevit University

Asst. Prof. Dr. Tuğba Palabaş
Zonguldak Bülent Ecevit University

Asst. Prof. Dr. Dinçer Biçer
Zonguldak Bülent Ecevit University

Asst. Prof. Dr. Nesibe Yalçın
Erciyes University

Asst. Prof. Dr. Sibel Ünaldı
Bilecik Şeyh Edebali University

Asst. Prof. Dr. Fatma Günseli Çıklaçandır
İzmir Kâtip Çelebi University

Asst. Prof. Dr. Ümit Aydoğan
İzmir Kâtip Çelebi University

Asst. Prof. Dr. Gözde Çay
İzmir Kâtip Çelebi University

Asst. Prof. Dr. Kadriye Filiz Balbal
Dokuz Eylül University

Inst. Dr. Ümit Sarp
İzmir Kâtip Çelebi University

Inst. Dr. Sümeyye Sınır
İzmir Kâtip Çelebi University

Inst. Dr. Mutluhan Çamur
Zonguldak Bülent Ecevit University

Inst. Bağdagül Gümüş Akyol
Zonguldak Bülent Ecevit University

Res. Asst. Murat Varol
Zonguldak Bülent Ecevit University

SCIENTIFIC COMMITTEE (Bilim Kurulu)

Prof. Dr. Şenol Hakan Kutoğlu
Zonguldak Bülent Ecevit University Vice Rector

Prof. Dr. Servet Karasu
Zonguldak Bülent Ecevit University Vice Rector

Prof. Dr. Bekir Hakan Bakkal
Zonguldak Bülent Ecevit University Vice Rector

Prof. Dr. Yasin Bulduklu
İzmir Kâtip Çelebi University Vice Rector

Prof. Dr. Muhsin Akbaş
İzmir Kâtip Çelebi University Vice Rector

Prof. Dr. Süleyman Akbulut
İzmir Kâtip Çelebi University Vice Rector

Prof. Dr. Hamit Aydın
Zonguldak Bülent Ecevit University, Dean of Faculty of Engineering

Prof. Dr. Kadir Demir
Zonguldak Bülent Ecevit University, Dean of Faculty of Science

Prof. Dr. Gokcen Bombar
İzmir Kâtip Celebi University, Dean of Faculty of Engineering and Architecture

Prof. Dr. Murat Kayacan
İzmir Kâtip Çelebi University

Prof. Dr. Ayşegül Alaybeyoğlu
İzmir Kâtip Çelebi University

Prof. Dr. Mustafa Alkan
Gazi University

Prof. Dr. Trushit Upadhyaya
C.S.Patel Institute of Technology, Charotar University of Sci. & Tech.

Prof. Dr. Dante Augusto Couto Barone
Federal Do Rio Grande Do Sul Uni. - Zonguldak Bülent Ecevit Uni.

Prof. Dr. Chemseddine Zebiri
Ferhat Abbas University of Setif, Algeria

Prof. Dr. Badrul Hisham bin Ahmad
Universiti Teknikal Malaysia Melaka, Malaysia

Prof. Dr. Zahriladha Zakaria
Universiti Teknikal Malaysia Melaka Durian Tunggal, Malaysia

Prof. Dr. Slawomir Koziel
Reykjavik University

Prof. Dr. Yılmaz Yıldırım
Zonguldak Bülent Ecevit University

Prof. Dr. Ozan Artun
Zonguldak Bülent Ecevit University

Prof. Dr. Cihan Karakuzu
Bilecik Şeyh Edebali University

Prof. Dr. Gülay Tezel
Konya Technical University

Prof. Dr. Bahadır Karasulu
Çanakkale Onsekiz Mart University

Prof. Dr. Ahmet Özmen
Sakarya University

Prof. Dr. Fikret Gölgeleyen
Zonguldak Bülent Ecevit University

Prof. Dr. Merih Palandöken
İzmir Kâtip Çelebi University

Prof. Dr. Aytuğ Onan
İzmir Kâtip Çelebi University

Prof. Dr. Kemal Polat
Bolu Abant İzzet Baysal University

Prof. Dr. Nurhan Karaboğa
Erciyes University

Prof. Dr. Can Kızılateş
Zonguldak Bülent Ecevit University

Assoc. Prof. Dr. Arbi Haza Nasution
Universitas Islam Riau, Indonesia

Assoc. Prof. Dr. Anna Pietrenko-Dabrowska
Gdansk University of Technology, Zweryfikowany

Assoc. Prof. Dr. Tologon Karataev
Nile University of Nigeria

Assoc. Prof. Dr. Md Baharul Islam
Florida Gulf Coast University, FL, USA

Assoc. Prof. Dr. Sabina Barakovic
University of Sarajevo, Bosnia and Herzegovina

Assoc. Prof. Dr. Hatice Yıldız Durak
Necmettin Erbakan University

Assoc. Prof. Dr. Evrim Güler
Bartın University

Assoc. Prof. Dr. Fatih Erdoğdu
Zonguldak Bülent Ecevit University

Assoc. Prof. Dr. Seyfullah Gökoğlu
Bartın University

Assoc. Prof. Dr. Süleyman Uzun
Sakarya University of Applied Sciences

Assoc. Prof. Dr. Yasemin Önal
Bilecik Şeyh Edebali University

Assoc. Prof. Dr. Cem Çırak
İzmir Kâtip Çelebi University

Assoc. Prof. Dr. Hilal Kahraman
İzmir Kâtip Çelebi University

Assoc. Prof. Dr. Mustafa Ergun
Ondokuz Mayıs University

Assoc. Prof. Dr. Sehran Dilmaç
İzmir Kâtip Çelebi University

Assoc. Prof. Dr. Seda Oğuz Ünal
Sivas Cumhuriyet University

Assoc. Prof. Dr. Şule Ayar Özbal
Yaşar University

Assoc. Prof. Dr. Ali Narin
Zonguldak Bülent Ecevit University

Assoc. Prof. Dr. Levent Aydın
İzmir Kâtip Çelebi University

Assoc. Prof. Dr. Mehmet Ali Belen
İskenderun Technical University

Assoc. Prof. Dr. Bekir Emiroğlu
İzmir Kâtip Çelebi University

Assoc. Prof. Dr. Mücahit Yalçın Öztüfekçi
İzmir Kâtip Çelebi University

Assoc. Prof. Dr. Murat İskefiyeli
Sakarya University

Assoc. Prof. Dr. İbrahim Etem Zinderen
Atatürk University

Assoc. Prof. Dr. Gizay Daver
Zonguldak Bülent Ecevit University

Assoc. Prof. Dr. Ersin Kanat
Zonguldak Bülent Ecevit University

Assoc. Prof. Dr. Huriye Çolaklar
Bartın University

Asst. Prof. Dr. Arpan Desai
Pandit Deendayal Energy University, Gandhinagar, Gujarat, India

Asst. Prof. Dr. Hesham Alhumyani
Taif University, Taif, Saudi Arabia

Asst. Prof. Dr. Gaye Ediboglu Bartos
Alba Regia Technical Faculty, Obuda University, Hungary

Asst. Prof. Dr. Vinayakumar Ravi
Prince Mohammad Bin Fahd University

Asst. Prof. Dr. Kadriye Filiz Balbal
Dokuz Eylul University

Asst. Prof. Dr. Aslı Eyecioğlu Özmutlu
Bartın University

Asst. Prof. Dr. Fatih Sarıkoç
Erciyes University

Asst. Prof. Dr. Fatma Güler Eroğlu
Atılım University

Asst. Prof. Dr. Hakan Üçgün
Bilecik Şeyh Edebali University

Asst. Prof. Dr. Vedat Marttin
Bilecik Şeyh Edebali University

Asst. Prof. Dr. Rıdvan Yayla
Bilecik Şeyh Edebali University

Asst. Prof. Dr. Bengi Yıldız
Bilecik Şeyh Edebali University

Asst. Prof. Dr. Zennure Uçar
İzmir Kâtip Çelebi University

Asst. Prof. Dr. Onan Güren
İzmir Kâtip Çelebi University

Asst. Prof. Dr. Ahmet Emin
Karabük University

Asst. Prof. Dr. Ezgi Aydemir
Iğdır University

Asst. Prof. Dr. Ünzile Yaman
İzmir Kâtip Çelebi University

Asst. Prof. Dr. Semih Çakır
Zonguldak Bülent Ecevit University

Asst. Prof. Dr. Ceren Kaya
Zonguldak Bülent Ecevit University

Asst. Prof. Dr. Tuğba Palabaş
Zonguldak Bülent Ecevit University

Asst. Prof. Dr. Dinçer Biçer
Zonguldak Bülent Ecevit University

Asst. Prof. Dr. Nesibe Yalçın
Erciyes University

Asst. Prof. Dr. Sibel Ünaldı
Bilecik Şeyh Edebali University

Asst. Prof. Dr. Fatma Günseli Çıklaçandır
İzmir Kâtip Çelebi University

Asst. Prof. Dr. Ümit Aydoğan
İzmir Kâtip Çelebi University

Asst. Prof. Dr. Durmuş Özkan Şahin
Ondokuz Mayıs University

Asst. Prof. Dr. Gözde Çay
İzmir Kâtip Çelebi University

Asst. Prof. Dr. Aykut Karakaya
Zonguldak Bülent Ecevit University

Dr. Fnu Neha
Kent State University

Dr. Hakan Can Altunay
Ondokuz Mayıs University

Dr. Gülsade Kale
Kilis 7 Aralık University

Dr. Jagriti Saini, Founder
Eternal RESTEM, Himachal Pradesh, India

Dr. M. Mustafa Bahşı
Manisa Celal Bayar University

Dr. Ayşe Kurt Bahşı
Manisa Celal Bayar University

Dr. Ümit Sarp
İzmir Kâtip Çelebi University

Dr. Sümeyye Sınır
İzmir Kâtip Çelebi University

Dr. İsmail Terzi
Zonguldak Bülent Ecevit University

Dr. Gülhan Toğa
Erciyes University

Dr. Mutluhan Çamur
Zonguldak Bülent Ecevit University

Inst. İdil Özkan
Zonguldak Bülent Ecevit University

Inst. Bağdagül Gümüş Akyol
Zonguldak Bülent Ecevit University

Res. Asst. Murat Varol
Zonguldak Bülent Ecevit University

Anusha Narapureddy
Engineering at Apple

INVITED SPEAKERS

(Davetli Konuşmacılar)

Prof. Dr. Derviş Karaboğa

"Recent Criticisms About Metaphor-Based Optimization Algorithms, and Artificial Bee Colony-ABC"

Brief CV: Dr. Karaboğa is an academician, engineer, and researcher known for his significant contributions to fields such as artificial intelligence, optimization, and control systems. He has developed AI-based solution methods and algorithms, particularly focusing on optimization problems. Karaboğa completed his master's and doctoral studies in Turkey and has worked as a faculty member at various universities, contributing extensively to academic research. His work spans both theoretical and applied engineering problems, and his contributions to artificial intelligence and optimization have gained international recognition. Throughout his academic career, Prof. Dr. Karaboğa has mentored numerous students, helping shape the next generation of engineers. His achievements in research and education have earned him a solid reputation in the engineering community.

Prof. Dr. Tülay Yıldırım

"Data Dancing in the Cyber Shadow: Privacy and Security in the Age of Artificial Intelligence"

Brief CV: She received her BSc degree from Yıldız Technical University (YTÜ), Department of Electronics and Communication Engineering in 1990, MSc degree in 1992 and PhD degree in 1997 from the Department of Electrical and Electronics Engineering, University of Liverpool, UK. After working as an R&D engineer in a private company for a short time, she started his academic career as a Research Assistant at Yıldız Technical University, Department of Electronics and Communication Engineering in 1991. She became an Assistant Professor in 1997, Associate Professor in 2000, and Professor in 2006, and held various administrative positions such as Head of Department and Vice Dean at the same university. Her research interests include artificial intelligence, intelligent systems, cyber-physical systems, biomedical instrumentation, biometric personal identification and verification systems, electronic circuits and systems, and artificial neural networks.

Prof. Dr. Sabina Baraković

"Artificial Intelligence in Cyber Security"

Brief CV: She holds a PhD in electrical engineering. She is a university professor and has published over 8 books and 80 journal and conference papers. She has held over 60 invited talks and led and participated in over 15 international scientific projects. Also, she participates in several international working groups for cyber security and artifical intelligence within UN, NATO, OSCE, COE. She has received several international and domestic awards for her extraordinary scientific results. Her research interests are cyber security, smart concept, quality of life and experience.

Prof. Dr. Kunal Mankodiya

"Will Smart Fabrics Enable the AI Revolution in Human Health and Wellness?"

Brief CV: Kunal Mankodiya, PhD is a tenured Professor of Biomedical Engineering at the University of Rhode Island and Director of the Wearable Biosensing Lab (WBL). With significant support from federal (NSF, NIH), state (RI Commerce), and foundation grants, he has established a transformative research program focused on wearable and mobile medical technologies for monitoring the brain, body, and behavior. Dr. Mankodiya's team developed a smart glove system for Parkinson's disease assessment in home environments, funded by the prestigious NSF CAREER Award, and a digital health platform for cognitive assessment, supported by NIH. His interdisciplinary research integrates AI, IoT, and advanced sensing technologies to address real-world healthcare challenges. He is currently collaborating internationally with India to expand the research on Smart Gloves. Recognized with honors such as the NSF CAREER Award, Innovator of the Year, and Providence Business News' 40 Under 40, he has mentored over 18 graduate students and 3 postdocs and has published more than 115 research papers on cutting-edge health technologies.

Prof. Dr. Bijan Najafi

"Empowering Patients: Innovative Decentralized Solutions for Better Care"

Brief CV: He is a tenured Professor of Surgery at UCLA, specializing in digital health and biotechnologies. He serves as the Research Director of the UCLA Center for Advanced Surgical and Interventional Technology (CASIT) and Co-Director of the NSF IUCRC Center to Stream Healthcare in Place (C2SHIP). He earned his Ph.D. in bioengineering from the Swiss Federal Institute of Technology (EPFL) in Switzerland and completed a postdoctoral fellowship in neuroscience at Harvard University. Dr. Najafi has published over 250 articles, garnering more than 17k citations. In 2022, he was ranked in the top 1% by Expertscape in his field of research, which includes digital health, diabetic foot care, and wound care. In recognition of his significant contributions to digital health technologies aimed at preventing falls and limb gangrene, he was inducted into the AIMBE College of Fellows in 2023. Throughout his career, he has mentored over 500 individuals.

Assoc. Prof. Dr. Ali Turab

"Integrating Mathematical Modeling and Machine Learning for Animal Behavioral Analysis in T-Maze Experiments"

Brief CV: Dr. Ali Turab is an Associate Professor at Northwestern Polytechnical University (NPU), China. He holds a Ph.D. in Applied Mathematics from Thammasat University, Thailand. His postdoctoral research at TU and tenure as a Research Associate at Universiti Malaysia Pahang Al-Sultan Abdullah (UMPSA) have refined his expertise in Computational Modeling, Animal Cognition & Behavior, and Machine Learning Applications in Behavioral Science. His research spans Mathematical Modeling, Psychological Learning Theory, and AI-driven Behavioral Analysis, with a particular focus on Neural Networks, Fractional Calculus, and Nonlinear Analysis. Dr. Turab has contributed significantly to scholarly discourse through numerous publications and presentations at international conferences. He remains committed to advancing research and education through interdisciplinary collaboration.

TABLE OF CONTENTS

İÇİNDEKİLER

PREFACE **1**
ÖNSÖZ **3**
BOARDS & COMMITEES (Kurullar ve Komiteler) **5**
INVITED SPEAKERS (Davetli Konuşmacılar) **14**
ABSTRACTS (Özetler) **23**

A Bibliometric Review of Research on Artificial Intelligence in Health Sciences Education 24
Nermin Zor - Mehmet Kara

A Deep Learning Approach for Estimating Capture Dates of German Historical Images 25
Noah Sentürk - Patrick Philipp

A Machine Learning-Driven Approach for Pothole & Speed Bump Detection Using Sensor Fusion 27
Efecan Cebel - Huseyin Karacali - Nevzat Donum

AI for Kinesiology Research: Classification of Posture and Knee Pathology 28
İmran Göker - Ali Aksaç - Ümit Yılmaz - Fatma Önay Koçoğlu - Gülizar Pat - Erhan Kurmuş Hüseyin Yusuf Hocaoğlu

AI-Based Modeling of Structural Stress in Multi-Deck Passenger Ship 30
Melih Savran - Mustafa Kafalı

AI-Driven Aspect-Based Sentiment Analysis on Online Reviews for Detecting Dish Preferences and Trend Forecasting 31
Vicdan Yalçın - Ahmet Cumhur Öztürk - Mustafa Çetin

Airline Food Optimization, "Flight Max" 33
Sena Yılmaz - Adem Erdem Kurt - Emre Atlıer Olca

Alzheimer's Disease through Gamification in the Metaverse 35
Ceylin Alak - Emre Atlıer Olca

An overview using GenAI on Cyber Security Applications 37
Murat Varol - Murat İskefiyeli

Araştırma Yöntem ve Teknikleri Dersinde Yapay Zekâ Kullanımının Derse Yönelik Tutuma Etkisi 39
Turgay Erdemir - Hasan Çoruk

Artificial Intelligence in Education: Sentiment-Based Analysis of Pre-Service Teachers' Views 40
Fatih Erdoğdu - Seyfullah Gökoğlu

Artificial Intelligence Supported Energy Performance Certificate (EPC) Analysis and Recommendation System 41
Tolga Tanrıbuyruğu - Emre Atlıer Olca

Artificial intelligence based publication analysis for Zonguldak Bülent Ecevit University 43
Aysegul Tok

Blok Zincir Tabanlı Oylama Sistemi 45
Mehmet rasim Yıldız - Meryem Soysaldı Şahin

Credit Card Fraud Detection Using Machine Learning Methods 47
Abdullah Alawad - Vedat Marttin

Determining the Optimal Number of Charging Units at Electric Vehicle Charging Stations Using a Queueing Theory Approach 49
Bayram Dündar

Development of an AI-Based System for Automated Cooking Using Color Tracking 50
Fırat Koçyiğit - Ayşegül Alaybeyoğlu

Dijital Ekonomi ve Blok Zinciri Uygulamaları 52
Hakan Can Altunay

Finansal Kişilik Profilleri: Klontz Money Script Inventory ve Yapay Zekâ Destekli Sınıflandırma ile Finans Öğrencilerinin Paraya Dair İnançlarının Psikolojik Analizi 53
Özge Ece Çiftci - Ersin Kanat

From Zero-Divisor Graphs to Topological Indices: A Structural Study of Semigroup-Derived Graphs 54
Seda Oguz Unal

Genotoxicity of Fig Seed Oil and The Comparison of Different Artificial Intelligence Tools For Data Analysis 55
Gamzegül Açış - Begümhan Yılmaz Kardaş - Mehmet Emin Diken

Hastanelerde Hijyen ve Güvenlik İçin Yapay Zekâ Destekli Kamera Tabanlı Koruyucu Ekipman Kontrol Sistemi 57
Muhsin Kavan - Ayşegül Alaybeyoğlu

In Silico Target Profiling of a Novel Dithiocarbamate Derivative (GKF25) via AI-Driven Molecular Docking for Neuropsychiatric Applications 58
Unzile Yaman - Gulsah Karakaya - Ofcan oflaz

İşletmelerde Yapay Zekâ Çağı 60
İlknur Zeydan

Learning Engagement and Analysis for Real-time Navigation (LEARN) 61
Fatih Emre Akarsu - Emre Atlıer Olca

Makine Öğrenmesi ile Desteklenen Ömür Devri Yönetimi ve Önleyici Bakım Planlamasında Taşıma Verimliğinin Arttırılması 63
Sevdenur Uzer - Fırat Ateş

Makine Öğrenmesi Regresyon Modelleri Kullanılarak Altın Ons Fiyatlarının Tahmini 65
Halil Sezgin Can - Fırat Ateş

Makine Öğrenmesi Yöntemleri ile Kötü Amaçlı Yazılım Tespit Analizi 66
Erol Özçekiç - Ümit Yılmaz - Ümit Mamuk

MedMind: A Data-Guided Digital Medication Management Solution for Individuals with Cognitive and Sensory Impairments 67
Unzile Yaman - Muhammet Can Sekman

Operator-Theoretic and Quantum-Inspired Frameworks for Data-Driven Modeling of Nonlinear Dynamical Systems 68
Joanna Slawinska - Dimitris Glannakis

Oturum Konfor Tasarımının Objektif Olarak İncelenmesi 69
Süleyman Dündar - Esra Akgül

Ödev Destek Sistemi Olarak Yapay Zekâ Kullanımı 71
Hasan Çoruk - Turgay Erdemir

Öğretmen Adaylarının Büyük Verinin Eğitimde Kullanımına İlişkin Farkındalıklarının İncelenmesi 73
Ayşegül Ayyıldız Asil - Abdurrahman Asil

Perception of Olympic Opening Ceremonies Through Social Media Commentary: A Sentiment Analysis and Topic Modeling Approach 75
Hüseyin Koçak

Sanal Gerçeklik Teknolojisi ile Simülasyon Ortamında Deprem Hazırlık Eğitimi 76
Nur Kuban Torun

Savunma Sanayisinde Makine Öğrenmesi ile Desteklenen Önleyici Bakım Planlaması 77
Pınar Özel - Sebahattin İkikardeş

Siber Güvenlik Konusunda Öğretmen Adaylarının Farkındalıklarının İncelenmesi 78
Ayşegül Ayyıldız Asil - Abdurrahman Asil

Siber Savaşta Yapay Zekâ: Askeri Yapay Zekâ Uygulamalarında Evrensel Etik ve Hukukun Dijital Çatışmadaki Rolü 80
Erdoğan Kahraman

Smart Optimization in Logistics: A Conceptual Review 82
İrem Karakaya

Sualtı Görüntülerinde Derin Öğrenme ile Nesne Tespiti 83
Kadriye Filiz Balbal - Özge Çelik - Sebahattin İkikardeş

Sürdürülebilir Kalkınma Hedeflerinin Çocuklara Mobil Oyun Uygulaması ile Öğretimi 84
Nur kuban Torun

Sürdürülebilirlik Yönetiminde Yeşil Teknolojilerin Önemi 85
Yaprak Akçay Zileli

Trends in IoT Security: A Post-Quantum Perspective 86
Aykut Karakaya

Wheat Germ Oil as a Potential DNA-Protective Agent: Insights from SMART Assay and AI-Assisted Analysis 87
Melih Can Demirbek - Serap Doğan - Begümhan Yılmaz Kardaş

Yapay Zekâ ve Gazetecilik: Lisansüstü Tezlere İlişkin Bir İnceleme 89
Mutluhan Çamur

2020-2024 yılları arasında Web of Scinece Veri Tabanında Üretken Yapay Zeka ve Programlama Eğitimi Alanlarında Yayınlanmış Yayınların Bibliyometrik Analizi 90
Erdinç Karaman - Mesut Erdönmez

FULL PAPERS (Tam Metinler) 92

A Comparative Analysis of Deep Learning Models for Plant Leaf Disease Classification 93
Buse Calisir - Bihter Das

A Comparative Study of Machine Learning Models for Dyslexia Prediction with Explainable Artificial Intelligence Analysis 102
Şule Akça - Tamer Uçar

A Dynamic Trajectory Optimization Model Proposal in the Context of Intelligent Automation 113
Muhammet Raşit Cesur - Elif Cesur

Analyzing Curriculum Evolution in a Morphologically Rich Language Using Unsupervised NLP: A Case Study in Turkish Mathematics Education 118
Batın Örene - Ghislaine Le Grange

Artificial Intelligence in Multilingual Itinerary Planning: A Case Study of Grok in Diyarbakır 133
Vedat Kızıl

Artificial Intelligence Performance in Embedded Systems: Comparisons of Inference Time, Cost, and Energy Consumption 137
Taner Erol - Kemal Polat

Background Subtraction-Based Approach for Human Activity Recognition Using MmWave Radar 147
Doga Naici - Zehra Bilici Saliha Yesilyurt

Automated Retinal Disease Classification for Early Detection Using Machine Learning Models 159
Abdul Samad - Muhammed Kürsad Uçar

Benliğin Dijital Dönüşümü, Yapay Zekâ ve Yalnızlık 173
Nilgün Özcan

Blockchain-Enabled Smart Contracts for Sustainable Medical Waste Management: Enhancing Traceability, Compliance, and Efficiency 179
Arif Soyler - Esra Dundar Aravacık - Serhat Burmaoğlu

Calligraphy Drawing with a Simple Robotic Arm 202
Rıdvan Yayla - Hüseyin Hakan Çetinkaya - Semih Çakır

Chaotic Signal-Induced Firing Characteristic of Hodgkin-Huxley Neuron in Neural Network Topology Based on Newman-Watts Approach 211
Muhammed Elhatip - Tuğba Palabaş - Ziynet Pamuk

Collision Detection for the Safety of Autonomous Systems using Machine Learning Algorithms and AutoML Libraries 218
Elif Cesur - Handegül İpek - Hilal Gülmez - Büşra Çetiner

Comparative Analysis of AI Models for Predicting Customer Purchase Behavior in E-Commerce 230
Shefitjon Bregu - Yusuf Kartal

Comparative Analysis of Brain Tumor MRI Scans Segmentation Using SegFormer Transformer and Hybrid 236
Muhammad Ahmed Qureshi - Mahnoor Tauqeer - Yashfa Kamran

Comparatıve Analysıs of Large Language Models 254
Buse Cansu Yiğit - Ahmet Ali Süzen - Remzi Gürfidan

Kalp Yetmezliği Hastalığının Tahmininde Makine ve Derin Öğrenme Modellerinin Karşılaştırmalı Analizi 263
Taner Erol - Kemal Polat

Comparing the NFT minting process: Ethereum vs Solana 274
Gjergji Mulla - Kristi Gorea

Condillac ve Yapay Zekâ: Algıdan Zihinsel Yapıya Bir Yolculuk 283
Hilal Kahraman

Connectogram: A Novel Time-Aware Graph Representation for EEG-Based Epileptic Seizure Detection 287
Serkan Aksu - Bayram Akgül - Fatih Dinç

DDoS Saldırılarının Makine Öğrenmesi Yaklaşımları ile Tespit Edilmesi 299
Sedanur Şahinoğlu - Fatih Ertam

Deep Learning for Cardiomyopathy Diagnosis Using Cardiac MRI: A Comparative Analysis of Dropout and Optimization Techniques on Classification Performance 313
Yeliz Alper - Ceren Kaya

Deep Learning Models for Classification, Diagnosis, and Treatment of Rotator Cuff Tears: A Comprehensive Examination 325
Kemal Polat - Büşra Kaya Olgaç

Deepfake Video Detection: A Comparative Study of ResNeXt50 and LSTM Architectures 334
Leili Eslam Jamalgolzari - Aydin Cetin

Design of a Patient Monitoring Device for Effective Inhaler Medication Usage 347
Erhan Sesli - Hamdi Eren Cagilci - Mustafa Bagatur

Enhancing Ancillary Product Classifications in the Aviation Industry: A Comparative Analysis of Machine Learning Algorithms 359
Yavuz Selim Emir - Tamer Uçar

Dinamik Sınıflandırıcı Seçim Yöntemleri ile Metin Tabanlı Duygu Sınıflandırması: Mağaza Yorumları Üzerine Bir Çalışma 372
İsmail Terzi

Enkaz Arama Alanında Karınca Sistemi Kullanılarak Geliştirilen Bir Algoritma ile Yapılan İnceleme 382
Zeynep Karabulut - Mehmet Dinçer Erbaş

Etkin Kaynak Kullanımı Amacı ile Güvenli Bir API Yönetim Mimarisi Tasarımı 391
Muhammet Emin - Gökçen Gökyer Nalcı - Özlem Emin - Furkan Baran - Ali Doğan - Mustafa Kozan - Önder Demir

Forwarder Truck System 399
Eyüp Kerem Malkoç - Emre Atlıer Olca

Gestational Diabetes Prediction with ANFIS and Machine Learning Methods 407
Sevcan Emek - Mehmet Emre Yılmaz

GIS-based Approach to Assessing Land Use Change Impacts on Urban Heat Island Patterns 416
Rabia Bovkir

Improving E-Mail Security with Zero-Shot Learning: A BART-MNLI Model Perspective 424
Remzi Gürfidan - Ahmet Ali Süzen

Improving Information Access in the Healthcare Sector with RAG: Enhancing the Effectiveness of Patient Information Forms 431
Murat Genç

Investigation of NiTi-Based Shape Memory Alloys Design and Performance Using Artificial Intelligence and Machine Learning 451
Ümit Zeybek

Language Model Approach in Blockchain-Based Forensic Unit Management System 459
Güncel Sarıman - Tuğçe Güzle

Makine Öğrenmesi Yöntemleriyle İzmir İli Su Tüketiminin İncelenmesi ve Su Tüketimini Tahmin Etmede Etkin Faktörlerin Tespit Edilmesi 477
Fatih Tek - Fatma Günseli Yaşar Çıklaçandır

Manipulating Images with SL(2,Z): A Group-Theoretic Approach 483
Bilal Demir - Özden Koruoğlu

Meslek Yüksekokulu Öğrencilerinin Yapay Zekâya Yönelik Görüşlerinin Belirlenmesi: Ordu Üniversitesi Örneği 490
Servet Kılıç

Morphological and Statistical Analysis of Iris Flower Species: A Study on Classification and Differentiation 496
Kemal Polat - Büşra Kaya Olgaç

Numerical Solution of The Pantograph Type Differential Equation Using Jacobi Neural Network 514
M. Mustafa Nahşı - Ayşe Kurt Bahşı

Optimizing Urban Indoor Farming: UV Light and Automated Irrigation for Bok Choy Cultivation 523
Carlo R. Hadjirul - Jopel E. Enriquez - Ghaleeb Kangiluhan - Aradzna M. Kamman

Performance Analysis of a Graphene-Coated Fractal Microstrip Patch Antenna for 5G Applications 536
Beyza Nur Demirkoparan - Husnu Deniz Basdemir

Performance Analysis of the Cheetah Optimization Algorithm in Predictive Maintenance Forecasting 545
Nuriye Bektaşoğlu - Ali Narin - Uğur İleri

Platon ve Aristoteles Perspektifinden Yapay Zekâ 565
Burçak Dinçergök

Prediction of Water Levels in Dams and Reservoirs for Istanbul and Ankara with ANFIS Model 579
Sevcan Emek - Kübra Çelebi - Oğuz Aydın

Puma Optımızatıon in Use for Performance Revıew of an LNA Transıstor Throughout the Entıre Operatıon Band 593
Mehmed Sinan Kömek - Filiz Güneş

Real-Time Exercise Feedback Using Pose Estimation for Home-Based Physiotherapy 604
Ozan Ahmet Demir - Aysun Coşkun

Resizing 3D Brain MR Images: A Quantitative Assessment of Segmentation Differences 612
Murat Arıcan - Kemal Polat

RouteWise: Satellite Navigation Application for Time-Saving and Traffic-Aware Travel 620
Chylle Donovhan C. Garda - Rhezana Shira R. Bosante - Al-Asri S. Injal - Jaydee C. Ballaho

Sağlık Çalışanlarının Yapay Zekâya İlişkin Farkındalık Düzeylerinin Değerlendirilmesi 631
Semih Islıcık - Gamze Arıkan - Nazan Kartal - Osman Şahman

Serendipity 644
Arda Çildan - Ömer Çevrük - Emre Atlıer Olca

Siber Güvenlikte Anomali Tespiti: CIC-IDS2017 ve CIC DDoS2019 Veri Setleri ile Makine Öğrenmesi Tabanlı Yöntemlerin Karşılaştırmalı Performans Analizi 652
Alperen Tokgöz - Özal Yıldırım

Single Package IIoT Assisted Real Time Monitoring and Management Control System for Smart Factory 4.0 Industrial Applications 664
Yashfa Kamran - Habiba Tasneem - Noor Fatimah

Sosyal Ağlarda Geometrik Desen Tabanlı Yorumlanabilir Bağlantı Tahmini: Facebook Veri Seti Üzerinden GNN Uygulaması 679
Leyla Susamcıoğlu - Ümit Sarp - Ayşegül Alaybeyoğlu

Sosyal Medyadaki Sahte Hesapların Sınıflandırılması: Dengesiz Veri Üzerinde Model Performansı Analizi 692
Mehmet Varan - Aslı Yatkınoğlu - Fatih Soygazi - Bora Mocan

Time-to-Collision Prediction in Autonomous Vehicles Using Machine Learning Algorithms 709
Elif Cesur - Sena Cicek

Tomato Is Life: A Prototype for Sorting Tomatoes Based on Size, Colors, and Quality 714
Charl's Benedick Martell W. Inoferio - Rhea Jenn D. Sala - Mary Harolhette O. Rom-Jaydee C. Ballaho

Türkçe E-Ticaret Yorumları Üzerinde Duygu Analizi: BERT ile Klasik Yöntemlerin Karşılaştırılması 724
Melike Çiftçi - Nesibe Yalçın - Celal Öztürk

Uzaktan Algılama İçin Tarım Odaklı Çözünürlük Artırımının Yeşil Yoğun Bölgelerde İncelenmesi 732
Deniz Çelik - Ali Furkan Kamanlı - Buket Bingöl

Uzaktan Algılamada Segment Anything Model ile Tarım Alanlarının Tespiti ve Ekili Alan Analizi 738
Buket Bingöl - Ali Furkan Kamanlı - Deniz Çelik

Visualization and Edge Intersection Analysis of Polygonal Cycle Graphs 745
Ümit Sarp - Ahmet Emin - Bilal Demir

William ShakesBlake 2.0: Fine-Tuned Language Models for Hybrid Poetic Style Generation 757
Doga Nalci - Zehra Bilici - Saliha Yesilyurt

Yapay Zeka Şirketlerinin Söylemleri: Kurumsal Sloganlarda Anlam İnşası 769
İbrahim Etem Zinderen - Aslıhan Zinderen

Yapay Zekâ ve Habercilik Üzerine Yazılan Akademik Çalışmalar Üzerine Bir İnceleme 777
Alper Çolak

Yapay Zekânın Cezai Sorumluluğu: Sistematik Literatür Haritalaması Çalışması 793
Cansel Arslan - Aslıhan Tüfekci

Yapay Zekâ, Sanat, Müzik Alanındaki Gelişmelere Genel Bakış 803
Mehmet Şahin Akıncı

5G Çekirdek Ağında Yapay Sinir Ağı Modeli ile Saldırı Tespiti 817
Harun Şeker - Hüseyin Parmaksız

ABSTRACTS
(ÖZETLER)

A BIBLIOMETRIC REVIEW OF RESEARCH ON ARTIFICIAL INTELLIGENCE IN HEALTH SCIENCES EDUCATION

Nermin Zor*
Amasya University, Computer Education & Instructional Technology, Amasya, Türkiye
Mehmet Kara
Amasya University, Educational Sciences, Amasya, Türkiye
(Corresponding author) nermin67672@gmail.com

ABSTRACT

This study aims to investigate the current scholarly landscape of Artificial Intelligence in Health Sciences Education through bibliometric analysis. The data sources were the papers available on the Web of Science database published between 2000 and 2025, and are relevant to artificial intelligence and health sciences education, such as medical education and nursing education.

The database search was conducted in the title field. Within the search strategy, the document types of the article, proceeding, early access, review article, and book chapter were included in the analysis. Besides, the search results were refined through the filters of Web of Science categories and language. The papers published in English and indexed in the categories of Education Scientific disciplines, Education Educational research, and Education Special were included in the analysis. A total of 407 papers were included in the final bibliometric analysis. The data analysis was conducted by using the VosViewer tool.

The analysis revealed the findings regarding the co-authorship of organizations and countries, co-occurrence of keywords, citations and coupling of documents, sources, organizations, authors, and countries, and co-citations of the cited REFERENCES, authors, and sources. The findings were presented as tables and visualizations obtained from the VosViewer tool. The findings gathered from this study demonstrated the research trends in the use of artificial intelligence in health sciences education and offered recommendations for future research.

Keywords: Artificial intelligence, health sciences education, bibliometric analysis.

A DEEP LEARNING APPROACH FOR ESTIMATING CAPTURE DATES OF GERMAN HISTORICAL IMAGES

Noah Sentürk
Koblenz University, Faculty of Mathematics, Informatics & Technology, Remagen, Germany

Patrick Philipp*
Pforzheim University, Faculty of Business and Law, Pforzheim, Germany
(Corresponding author) patrick.philipp@hs-pforzheim.de

ABSTRACT

Introduction: The automatic estimation of capture dates for historical photographs is a challenging and largely unex- plored task in computer vision. Inspired by recent work in this area [1], this study introduces a deep learning approach that uses convolutional neural networks (CNNs) to classify historical German images into defined time periods without relying on manual annotations or metadata.

Methods: We constructed a dataset of 1333 historical images spanning 1890–2009, sourced from publicly available archives [2]. The images were divided into six key historical intervals: (1890–1917), (1918–1932), (1933–1945), (1946–1961), (1962–1989), and (1990–2009). We resized all images to 256×256 pixels and used data augmentation to improve generalization. The model is based on the Xception architecture, followed by global average pooling, dropout, and a dense classification layer. The dataset was split into 87% training and 13% test data.

Results: The CNN's best performance was observed for the class associated with the period 1890–1917 and 1990–2009, which were classified with relatively high accuracy and F1-scores. In contrast, intermediate periods such as 1918–1932 and 1933–1945 were more difficult to distinguish, often being confused with neighboring classes (cf. Figure 1). Corresponding sample images from each of the six historical periods used as classes are shown in Figure 2 for reference.

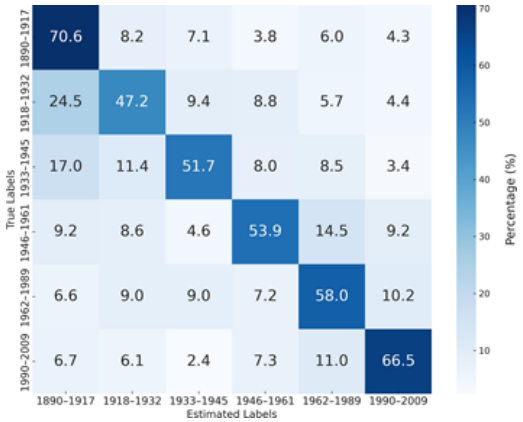


Figure 1: Confusion matrix of time periods.

Figure 2: Example images per time period.

Conclusion: This study shows that CNNs can successfully detect subtle visual patterns in historical photos to estimate their time period. Despite challenges with visually similar eras, the model surpasses human performance in several cases. Future work may include more data, finer time granularity, and transformer architec- tures.

Keywords: Historical images, deep learning, CNN, German history, image classification, estimation of capture date.

REFERENCES

[1] Y. Cimtay and H. ˙Ilk, "Capture Date Estimation of Historical Images by Using Deep Learning," in EKOTEK, 2024.1.

[2] Deutsches Historisches Institut, "Deutsche Geschichte in Dokumenten und Bildern," 2024. [Online]. Available at: https://germanhistorydocs.ghi-dc.org

A MACHINE LEARNING-DRIVEN APPROACH FOR POTHOLE & SPEED BUMP DETECTION USING SENSOR FUSION

Efecan Cebel*
TTTech Auto Turkey, İzmir, Türkiye

Huseyin Karacali
TTTech Auto Turkey, İzmir, Türkiye

Nevzat Donum
TTTech Auto Turkey, İzmir, Türkiye
**(Corresponding author) efecan.cebel@tttech-auto.com*

ABSTRACT

This study presents a machine learning-based approach for detecting potholes and speed bumps using only gyroscope and accelerometer data, eliminating the need for additional hardware such as GPS, cameras, or LiDAR. The proposed algorithm was trained on real-world road data and showed high success rates in evaluation tests through a lightweight neural network model in the Open Neural Network Exchange (ONNX) format. The model was trained on numerical data derived from gyroscope and accelerometer readings, specifically vibration frequencies and acceleration variation rates, collected across diverse road conditions and vehicle types. During preprocessing, advanced filtering techniques, including a Butterworth low-pass filter, were applied to reduce sensor noise while retaining essential signal features. Customized normalization and noise reduction procedures improved the clarity of vibration and acceleration patterns, supporting reliable road anomaly detection. Feature extraction emphasized high-frequency vibrations and sudden acceleration changes, which were then classified by the ONNX-based model. The use of the ONNX format allowed deployment on edge devices with limited computational capacity, enabling real-time processing with low latency and overhead. A key strength of this work is its high performance using only two sensor inputs and a computationally efficient model. Unlike traditional methods relying on image processing or LiDAR, this approach achieves comparable accuracy through sensor fusion alone. The ONNX format's portability and platform independence also ease integration into various hardware systems. Future work will focus on expanding the training dataset to cover more road conditions and vehicle types, and on exploring deep learning models like CNNs for finer classification of road irregularities.

Keywords: ONNX, Machine learning, Sensor fusion, Road anomalies, pothole and bump detection, CNN.

AI FOR KINESIOLOGY RESEARCH: CLASSIFICATION OF POSTURE AND KNEE PATHOLOGY

İmran Göker*
Üsküdar University, Bioengineering, İstanbul, Türkiye

Ali Alsaç
İstanbul University Cerrahpaşa, Industrial Engineering, İstanbul, Türkiye

Ümit Yılmaz
Muğla Sıtkı Koçman University, Software Engineering, Muğla, Türkiye

Fatma Önay Koçoğlu
İstanbul University, Artificial Intelligence & Data Engineering, İstanbul, Türkiye

Gülizar Pat
İstanbul University, Computer and Information Technologies, İstanbul, Türkiye

Erhan Kurmuş
Üsküdar University, Neuroscience, İstanbul, Türkiye

Hüseyin Yusuf Hocaoğlu
Muğla Sıtkı Koçman University, Software Engineering, Muğla, Türkiye
**(Corresponding author) imran.goker@uskudar.edu.tr*

ABSTRACT

Knee diseases are diseases that receive a lot of applications to polyclinics and may result in surgery. At this point, correct diagnosis and evaluation of the pathology for a fast and effective treatment process is very important in terms of patient health and correct management of resources. Artificial intelligence is widely used in the health field, as in many other fields, and supports doctors' decision-making processes with data-based predictions and definition outputs [1]. Recording the bioelectrical activity of skeletal muscles to aid in diagnosis is defined as Electromyography [2]. In this method, in addition to muscle activity, peripheral nerve conduction is also measured. This study aims to investigate the feasibility of classifying posture and knee pathology status using electromyography (EMG) signals from upper leg muscles that would be implemented into simulators being utilized in kinesiology research. A dataset comprising EMG signals, posture information (gait, standing, sitting), and knee pathology status (normal, pathological) was used. Time-domain (mean, standard deviation, RMS, zero-crossing rate) and frequency-domain (spectral power density via Fourier transform) features were extracted from the EMG signals. A Random Forest model was employed for both posture and knee pathology classification, achieving accuracies of 95% and 93%, respectively. The results demonstrate the potential of EMG signals for accurate classification in kinesiology applications, suggesting that these features can be

effectively integrated into simulators for detailed biomechanical analysis. Future research may focus on expanding the dataset, exploring alternative machine learning models, and optimizing hyperparameters to further enhance classification performance and robustness.

Table 1: Findings

Category	Posture Classification	Knee Pathology Classification
Accuracy	95%	93%
Precision	Standing: 0.95, Sitting: 0.94, Gait: 0.96	Normal: 0.92, Pathological: 0.94
Recall	Standing: 0.96, Sitting: 0.93, Gait: 0.97	Normal: 0.93, Pathological: 0.93
F1-Score	Standing: 0.95, Sitting: 0.94, Gait: 0.96	Normal: 0.92, Pathological: 0.94
Notes	EMG signals are effective for posture determination.	EMG signals are successful in determining knee pathology.

Keywords: EMG signals, artificial ıntelligence, random forest, kinesiology.

REFERENCES

[1] I. Goodfellow, Y. Bengio, A. Courville, "Deep learning", 2016, MIT Press.

[2] R. Merletti, P.A. Parker, "Electromyography: Physiology, engineering, and non-invasive applications", 2004, IEEE Press.

AI-BASED MODELING OF STRUCTURAL STRESS IN MULTI-DECK PASSENGER SHIP

Melih Savran*
İzmir Kâtip Çelebi University, Department of Mechanical Engineering, İzmir, Türkiye

Mustafa Kafalı
İzmir Kâtip Çelebi University, Department of Naval Architecture and Marine Engineering, İzmir, Türkiye
(Corresponding author) melih.savran@ikcu.edu.tr

ABSTRACT

In this study, the influence of critical design parameters on the stress distribution of a multi-deck passenger ship was mathematically modeled using an AI-based machine learning approach. The design parameters include the size of side openings; the number, height, and length of longitudinal and transverse bulkheads; and the geometry of the superstructure, deckhouse, and recesses. The relationship between these parameters and the resulting stress was established through various functional models, including polynomial, rational, trigonometric, and logarithmic forms, generated by the AI-based modeling method. The coefficient of determination (R^2) was employed as the primary metric to evaluate model performance. Among the models developed, the Absolute First Order Trigonometric Nonlinear (AFOTN) model demonstrated the highest prediction accuracy, achieving an R^2 value of 0.99 for both training and testing phases. To assess the model's robustness and reliability, a 10-run cross-validation was performed with randomly shuffled training and test datasets. The AFOTN model yielded average R^2 values of 0.98 for training and 0.95 for testing. The minimum stress predicted by the model was 4 MPa. Compared to the reference study, the proposed approach achieved a 33% improvement in stress reduction [1].

Keywords: Multi-deck passenger ship, AI-based modeling, machine learning, stress distribution.

REFERENCES

[1] J. Andric, P. Prebeg, M. Palaversa, and V. Zanic, "Influence of different topological variants on optimized structural scantlings of passenger ship," *Marine Structures*, vol. 78, pp. 102981, 2021.

AI-DRIVEN ASPECT-BASED SENTIMENT ANALYSIS ON ONLINE REVIEWS FOR DETECTING DISH PREFERENCES AND TREND FORECASTING

Vicdan Yalçın*
Aydın Adnan Menderes University, Department of Management Information Systems, Aydın, Türkiye

Ahmet Cumhur Öztürk
Aydın Adnan Menderes University, Department of Computer Technologies, Aydın, Türkiye

Mustafa Çetin
Aydın Adnan Menderes University, Department of Management Information Systems, Aydın, Türkiye
**(Corresponding author) yalcinvicdan@hotmail.com*

ABSTRACT

Online restaurant reviews contain valuable information about customer's dining experience. They also shape pREFERENCES and reputations. Sentiment analysis is a Natural Language Processing task for identifying emotions in text and categorizing them as negative, positive, or neutral. Aspect Based Sentiment Analysis (ASBA) in sentence level determines sentiments associated with specific aspects. Identifying food names and applying ASBA reveals opinions about specific dishes and helps determine which dishes are more liked or disliked. This helps restaurants refine menus and improve success. We scraped reviews from 2022–2024 on TripAdvisor.com for restaurants in Kuşadası, Turkey. We used GPT-4.0 Mini API to extract dish names and classify sentiments in related sentences. A structured prompt filtered out non-food aspects and ensured accurate classification into positive and negative. This AI-driven method enabled dish-specific sentiment identification and a labeled dataset for trend analysis. For forecasting, we used Facebook's Prophet model, a tool for seasonal trends and business predictions [1].

We refined the extracted dish names by eliminating non-specific terms such as "food", "breakfast", and "spicy", which appeared in GPT's output but did not refer to specific dishes. Additionally, we applied text normalization, lemmatization, and rule-based mapping to standardize dish names. All dish names were converted to lowercase, with special characters and extra spaces removed. We then used lemmatization and the Inflect library in Python to standardize singular and plural forms, ensuring that variations like steaks and steak were treated as the same entity. Furthermore, we created a predefined mapping dictionary to group similar dish names under a common label. For instance, t-bone steak, t bone steak, steak & prawns, and steak and mushroom pie were all mapped to the unified label steak. After dish names are refined they are systematically categorized based on their sentiment scores over three years (2022, 2023, and 2024). Next we applied Facebook's

Prophet model, a robust time-series forecasting method widely used for business analytics and trend prediction [1].

Our sentiment analysis revealed that dishes such as steak, pizza, sushi, calamari, pasta, prawns, fries, and sea bass received a significant number of positive sentiments over the three year period. The Forecasting results indicate that steak, pasta, and prawns will likely continue to receive predominantly positive sentiments in the coming years, and suggesting a strong customer preference and consistent satisfaction. Conversely, sushi and sea bass are projected to maintain positive sentiment but with a slower growth rate, indicating a potential area for quality improvement. On the negative sentiment side, steak, sushi, calamari, fried wonton, and lasagna emerged as dishes with the highest proportion of negative reviews. The forecasting model suggests that steak and sushi may continue receiving negative sentiments, suggesting areas for improvement in preparation and quality control. On the other hand lasagna and calamari show a declining trend in negative sentiments, which may indicate improved consistency in taste and quality. These findings align with recent studies emphasizing the role of consistency in food quality and customer satisfaction in online reviews [2].

Keywords: Large Language Models, Aspect Based Sentiment Analysis, Online Restaurant Reviews.

REFERENCES

[1] S.J. Taylor, and B. Letham, "Forecasting at scale", The American Statistician, 72, 1, 37-45, 2018. DOI:10.1080/00031305.2017.1380080.

[2] X. Zhang, Y. Li, J. Wang, and M. Chen, "Consumer sentiment and online food reviews: Examining the impact of consistency and quality perception", Journal of Business Research, 156, 113486, 2023. DOI:10.1016/j.jbusres.2023.113486.

[3] H. Chen, Y. Xu, L. Zhao, and P. Li, "AI-powered sentiment analysis for restaurant reviews: A deep learning approach to consumer behavior insights", Proceedings of AI and Business Analytics Conference, 84, 1, 56, 2023. DOI:10.3390/proceedings202308456.

AIRLINE FOOD OPTIMIZATION, "FLIGHT MAX"

Sena Yılmaz*
Maltepe University, Software Engineering, İstanbul, Türkiye

Adem Erdem Kurt
Maltepe University, Software Engineering, İstanbul, Türkiye

Emre Atlıer Olca
Maltepe University, Software Engineering, İstanbul, Türkiye
(Corresponding author) 200706021@st.maltepe.edu.tr

ABSTRACT

FlightMax is an advanced system designed to optimize in-flight catering processes within the aviation industry, focusing on reducing operational costs, improving resource utilization, and promoting environmental sustainability. The system leverages artificial intelligence (AI) to analyze critical variables such as flight routes, passenger profiles, flight duration, and historical data, thereby calculating the optimal food quantity required for each flight. This optimization reduces food waste and enhances operational efficiency by preventing overloading. Data is securely stored in a PostgreSQL database, ensuring reliable operations.

The system includes a user-friendly web-based interface with modules such as a dashboard for flight management, a catering planning tool for adjusting food quantities, and an analytics feature for generating reports based on historical data. These features simplify user interaction and improve catering efficiency.

Existing research in this field has focused on predictive models and inventory management to reduce food waste [2]. FlightMax builds on these concepts by integrating real-time AI-powered optimization tailored to the aviation industry. Furthermore, optimization methods in food processing workflows have been proven to minimize waste by analyzing variables like portion sizes and expiration dates [3], a strategy also adopted by FlightMax. Quality control measures for maintaining food freshness and improving logistics are incorporated, aligning with efforts to optimize food supply chains [4].

Airlines like Emirates and Etihad Airways have begun using AI to reduce food waste [5][6], although these approaches primarily focus on individual flights. FlightMax extends this by offering a multi-flight optimization strategy that improves overall resource usage across sequential flights.

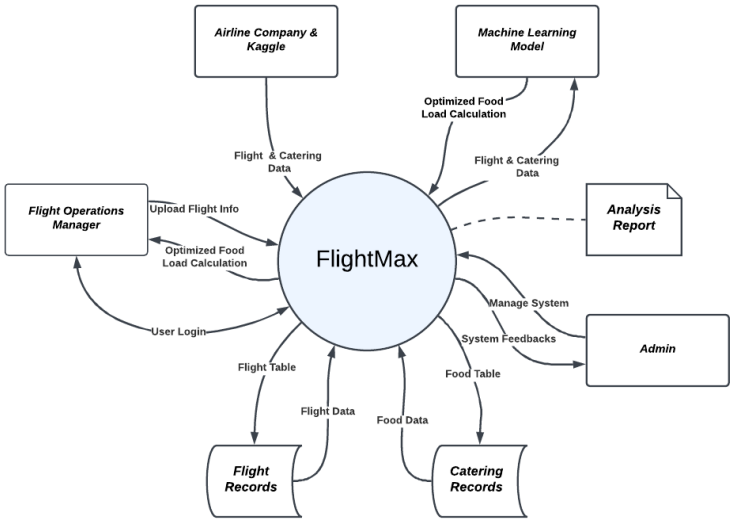


Figure 1: The FlightMax Context Diagram

The FlightMax Context Diagram shows the key interactions between the system and external entities. Users, such as the Flight Operations Manager, upload flight and catering data to the system, which processes it using the Machine Learning Model to calculate optimized food loads. The system generates analysis reports and provides feedback, while the Admin manages records like flight, food, and catering data. External data sources, like Airline Company and Kaggle, contribute additional flight and catering information to enhance system performance.

In conclusion, FlightMax modernizes aviation catering by integrating AI, data analytics, and user-friendly interfaces to reduce costs, enhance resource management, and minimize food waste, marking a significant step toward sustainable catering management in the aviation industry.

Keywords: FlightMax, artificial intelligence.

REFERENCES

[1] Global Air Travel Demand Continued Its Bounce Back in 2023, 31 January 2024, https://www.iata.org/en/pressroom/2024-releases/2024-01-31-02/?utm

[2] Van Rooijen, M. A., Gerdessen, J. C., Claassen, G. D. H., & de Leeuw, S. (2024). Optimizing household food waste: The impact of meal planning, package sizes, and performance indicators. Resources, Conservation and Recycling, 205, 107559. https://doi.org/10.1016/j.resconrec.2024.107559

[3] Julio R Banga ,Eva Balsa-Canto,Carmen G Moles ,Antonio A Alonso (2003). Improving food processing using modern optimization methods. https://doi.org/10.1016/j.ijpe.2009.11.026

[4] Aiying Rong, Renzo Akkerman ,Martin Grunow (2011) An optimization approach for managing fresh food quality throughout the supply chain. https://doi.org/10.1016/j.ijpe.2009.11.026

[5] https://ginjer.co.uk/consuming-data-how-ai-can-help-airlines-reduce-food-waste/

[6]https://aviationguideem.com/news/etihad-lumitics-to-use-technology-to-reduce-food-wastage-on-flights/

ALZHEIMER'S DISEASE THROUGH GAMIFICATION IN THE METAVERSE

Ceylin Alak*
Robert College, İstanbul, Türkiye

Emre Atlıer Olca
Maltepe University, Software Engineering, İstanbul, Türkiye
(Corresponding author) alacey.26@robcol.k12.tr

ABSTRACT

This paper explores the potential of gamification, virtual reality (VR), and artificial intelligence (AI) in the early diagnosis of Alzheimer's Disease. The proposed project aims to detect early signs of cognitive impairment in a non-invasive manner by transforming cognitive tests into an engaging VR game. Gamification has been proven to enhance user engagement and motivation in healthcare [1,2]. Additionally, VR technology has been increasingly used in healthcare due to its ability to enhance user experience, making both gamification and VR suitable for cognitive assessments [3, 4].

The game consists of tasks that assess memory, problem-solving, and spatial reasoning, which are among the key cognitive functions affected by Alzheimer's disease [5]. When the user enters the game, they are placed in a virtual environment where they engage in real-life tasks such as identifying shapes, organizing objects, and solving basic math problems. Performing these tasks in a gamified environment can lead to higher user engagement and more reliable data collection compared to traditional assessment methods [6].

For the data collection and analysis, machine learning models such as Random Forests (RF) and Long Short-Term Memory (LSTM) networks are used. RF is a robust decision-tree-based method that helps detect patterns in user responses, while LSTM networks enable tracking cognitive changes over time, allowing for a more accurate long-term analysis of cognitive decline [7,8].

Keywords: Gamification, virtual reality, Alzheimer's disease.

REFERENCES

[1] A. Bozkurt and E. Genç-Kumtepe, "Oyunlaştırma, Oyun Felsefesi ve Eğitim: Gamification," in Akademik Bilişim 2014, Mersin Üniversitesi, Mersin, pp. 147-156, 2014.

[2] E. Kutbay and N. Bozbuğa, "Health education: Gamification, health literacy, and the new era," in Metaverse Sağlık Eğitimi: Oyunlaştırma, Sağlık Okuryazarlığı ve Yeni Dönem, 2022. doi: 10.26650/B/ET07.2022.013.

[3] D. Kamińska et al., "Virtual reality and its applications in education: Survey," Information, vol. 10, no. 10, p. 318, 2019. doi: 10.3390/info10100318.

[4] VirtuAAL Research Group, "Goals," VIRTUAAL, Dec. 12, 2019. [Online]. Available: https://virtu-aal.eu/. Accessed: Sep. 8, 2024.

[5] A. R. Javed et al., "Artificial intelligence for cognitive health assessment: State-of-the-Art, open challenges and future directions," Cognitive Computation, vol. 15, no. 6, pp. 1767-1812, 2023. doi: 10.1007/s12559-023-10153-4.

[6] Alzheimer's Research UK, "Sea Hero Quest - Alzheimer's Research UK," Alzheimer's Research UK, Sep. 29, 2021. [Online]. Available: https://www.alzheimersresearchuk.org/research/for-researchers/resources-and-information/sea-hero-quest/. Accessed: Aug. 21, 2024.

[7] L. Breiman, "Random forests," Machine Learning, vol. 45, no. 1, pp. 5-32, 2001. doi: 10.1023/a:1010933404324.

[8] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735-1780, 1997. doi: 10.1162/neco.1997.9.8.1735.

AN OVERVIEW USING GENAI ON CYBER SECURITY APPLICATIONS

Murat Varol*
Zonguldak Bulent Ecevit University, Computer Engineering, Zonguldak, Turkiye
Murat İskefiyeli
Sakarya University, Computer Engineering Department, Zonguldak, Turkiye
(Corresponding author) muratv@beun.edu.tr

ABSTRACT

Artificial intelligence is a revolutionary technology that contributes to processes such as problem solving, data analysis and content generation by mimicking the functions of human intelligence. Classic AI analyzes existing data and makes classification or analysis. Generative AI, on the other hand, is frequently used in different fields such as health, education, art, informatics, security with its increasing creation power. Unlike Classic AI, GenAI can generate new content (text, code, images). In the field of cyber security, Classic AI is used to detect whether content is harmful or not. However, new malicious code, phishing, social engineering, misleading messages or fake content can be generated using Gen AI. In this paper, we review several studies that based Generative AI algorithms; anomaly Detection [1], intrusion detection [2], security concerns [3], industrial systems security [4,5] and Pentesting [6]. While Generative AI can be used as a powerful tool in a wide range of areas ranging from penetration testing to anomaly detection and Industrial system security, it also raises security concerns such as fake content generation and data manipulation. Against the security vulnerabilities of Generative AI, it is critical to develop fake content detection technologies, implement encryption algorithms, provide model training with ethical principles, develop user awareness practices and international policies. The main objective of this study is to explore the creative and analytical potential of GenAI and to recognize that applications in different sectors may have security risks. It also provides a scientific contribution to create public awareness and guide future impacts to make the use of AI more reliable.

Keywords: Generative AI, cyber security, artificial intelligence.

REFERENCES

[1] G. G. González, P. Casas and A. Fernández, "Fake it till you Detect it: Continual Anomaly Detection in Multivariate Time-Series using Generative AI," 2023 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW), Delft, Netherlands, 2023

[2] Rahman, Saifur, Et Al. Syn-Gan: A robust intrusion detection system using GAN-based synthetic data for IoT security. Internet of Things vol 26, 2024.

[3] S. Neupane et al., "Security Considerations in AI-Robotics: A Survey of Current Methods, Challenges, and Opportunities," in IEEE Access, vol. 12, pp. 22072-22097, 2024.

[4] M. S. Munir, S. Proddatoori, M. Muralidhara, W. Saad, Z. Han and S. Shetty, "A Zero Trust Framework for Realization and Defense Against Generative AI Attacks in Power Grid," *ICC 2024 - IEEE International Conference on Communications*, Denver, CO, USA, 2024.

[5] H. Lu, M. Du, K. Qian, X. He and K. Wang, "GAN-Based Data Augmentation Strategy for Sensor Anomaly Detection in Industrial Robots," in *IEEE Sensors Journal*, vol. 22, no. 18, pp. 17464-17474, 15 Sept.15, 2022.

[6] Hilario, E., Azam, S., Sundaram, J. et al. Generative AI for pentesting: the good, the bad, the ugly. Int. J. Inf. Secur. 23, 2075–2097, 2024.

ARAŞTIRMA YÖNTEM VE TEKNİKLERİ DERSİNDE YAPAY ZEKÂ KULLANIMININ DERSE YÖNELİK TUTUMA ETKİSİ

Turgay Erdemir*
Amasya Üniversitesi, Merzifon Meslek Yüksekokulu, Amasya, Türkiye

Hasan Çoruk
Zonguldak Bülent Ecevit Üniversitesi, Karadeniz Ereğli Meslek Yüksekokulu, Zonguldak, Türkiye
(İletişimden sorumlu yazar) erdemir05@gmail.com

ÖZET

Bu çalışma, araştırma yöntem ve teknikleri dersinde yapay zeka kullanımının öğrencilerin derse yönelik tutumlarına etkisini incelemeyi amaçlamaktadır. Günümüzde yapay zeka, eğitim ortamlarında giderek daha fazla kullanılmakta ve bilimsel araştırma süreçlerini destekleyerek öğrencilere kişiselleştirilmiş öğrenme deneyimleri sunmaktadır [1]. Araştırmada açıklayıcı sıralı karma yöntem kullanılmıştır. Çalışma grubu, 2024/2025 akademik yılında bir üniversitede araştırma yöntem ve teknikleri dersini alan 24 öğrenciden oluşmaktadır.

Veri toplama sürecinde nicel veriler için tutum ölçeği, nitel veriler için ise odak grup görüşmesi kullanılmıştır. Analiz sonuçlarına göre, yapay zeka kullanımının öğrencilerin derse yönelik genel tutumlarını olumlu yönde etkilediği bulunmuştur. Özellikle bilişsel özgüven ve ilgi boyutlarında anlamlı bir artış gözlemlenirken [3], dersin önemine ve mesleki yaşantıya uyarlanabilirliğine ilişkin algıda anlamlı bir değişim olmamıştır. Ayrıca, öğrenciler yapay zekanın araştırma süreçlerini hızlandırdığını, hata payını azalttığını ve akademik yazım sürecini desteklediğini belirtmişlerdir. Ancak, bazı öğrenciler yapay zeka kullanımının bilginin akılda kalıcılığına olumsuz etkisi olabileceğini ve derse katılımı azaltabileceğini ifade etmiştir [2].

Sonuç olarak, yapay zeka destekli eğitim ortamlarının derse yönelik ilgi ve bilişsel özgüven üzerinde olumlu etkileri olduğu görülmektedir. Ancak, bağımsız öğrenme becerilerini güçlendirmek, eleştirel düşünmeyi teşvik etmek ve dersin mesleki bağlamla daha fazla ilişkilendirilmesini sağlamak için ek stratejilere ihtiyaç duyulmaktadır. Gelecekte yapılacak çalışmaların, yapay zekanın uzun vadeli etkilerini incelemesi önerilmektedir.

Anahtar Kelimeler: Yapay zeka, bilimsel araştırma yöntemleri, öğrenci tutumu, karma yöntem, eğitim teknolojileri.

REFERANSLAR

[1] Kutlucan, E., & Seferoğlu, S. S. (2024). Eğitimde Yapay Zekâ Kullanımı: ChatGPT'nin KEFE ve PEST Analizi. Türk Eğitim Bilimleri Dergisi, 22(2), 1059-1083. https://doi.org/10.37217/tebd.1368821

[2] Selwyn, N. (2022). Should robots replace teachers? AI and the future of education. Polity Press.

[3] Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. International Journal of Educational Technology in Higher Education, 16(1), 1-27.

ARTIFICIAL INTELLIGENCE IN EDUCATION: SENTIMENT-BASED ANALYSIS OF PRE-SERVICE TEACHERS' VIEWS

Fatih Erdoğdu*
Zonguldak Bülent Ecevit University, Computer Technologies, Zonguldak, Türkiye
Seyfullah Gökoğlu
Bartın University, Computer Technology and Information Systems, Bartın, Türkiye
(Corresponding author) fatiherdogdu67@gmail.com

ABSTRACT

The aim of the study is to determine the distribution of positive, negative and neutral emotions by examining pre-service teachers' views on artificial intelligence through sentiment analysis methods. The study was conducted within the framework of descriptive research methods. The participants of the study were 48 pre-service teachers studying in different teaching departments. The participants wrote texts reflecting their thoughts on artificial intelligence and a data set consisting of 41,850 words in total was obtained. Python programming language was used to analyze the data and various sentiment analysis models were applied. In the analysis process, data preprocessing stages were carried out, and a manual list of emotionally charged words and a list of stop-words were used. According to the findings, 20.7% of the pre-service teachers' statements were classified as positive, 6.1% as negative and 73.2% as neutral. The overall emotion score (compound) was calculated as 1.0 and it was determined that the texts contained mostly positive emotions. The distribution of emotions was presented graphically. The results of the study show that pre-service teachers have a generally positive view of artificial intelligence. In the future, it is recommended to conduct such studies with larger samples and to use more advanced models in sentiment analysis.

Keywords: Artificial intelligence, sentiment analysis, pre-service teacher.

ARTIFICIAL INTELLIGENCE SUPPORTED ENERGY PERFORMANCE CERTIFICATE (EPC) ANALYSIS AND RECOMMENDATION SYSTEM

Tolga Tanrıbuyruğu*
Robert College, İstanbul, Türkiye

Emre Atlıer Olca
Maltepe University, Software Engineering Department, İstanbul, Türkiye
(Corresponding author) ttanribuyrugu@gmail.com

ABSTRACT

Energy efficiency is a crucial component in addressing global climate challenges and achieving Sustainable Development Goals (SDGs) [1]. Buildings contribute significantly to environmental impact, consuming approximately 30-35% of the world's energy and generating nearly 40% of global CO2 emissions [2]. The Energy Performance Certificate (EPC), or Energy Performance Certificate, serves as an essential tool in assessing and improving the energy efficiency of buildings. However, traditional EPC processes rely heavily on manual analysis, leading to inefficiencies in time and cost. This project introduces an innovative approach that integrates Artificial Intelligence (AI) to automate and enhance the EPC evaluation process. AI algorithms analyze data from the buildings to provide recommendations for energy performance improvements, allowing property owners to make well decisions. Furthermore, in the absence of an existing EPC, residents can manually input relevant parameters to generate energy optimization insights. The significance of this project is reinforced by its alignment with the United Nations Sustainable Development Goals (SDGs), particularly SDG 11 (Sustainable Cities and Communities) and SDG 13 (Climate Action). By enhancing the energy efficiency of buildings, this project contributes to reducing carbon footprints and improving sustainable urban development. Additionally, it provides communities and urban planners with precise, data-driven insights for large-scale energy optimization and environmental planning, supporting a transition toward more eco-friendly cities [3]. In conclusion, this project demonstrates how AI-driven automation can revolutionize the Energy Performance Certificate (EPC) evaluation process, making energy assessments more efficient, cost-effective, and accessible. By leveraging AI to analyze building data and provide actionable recommendations, it empowers property owners, residents, and urban planners to enhance energy efficiency and reduce environmental impact. Aligning with global sustainability efforts, this approach supports the transition toward greener cities and contributes to the broader goals of climate action.

Keywords: Energy efficiency, digital transformation, environmental sustainability, Energy Performance Certificate (EPC).

REFERANSLAR

[1] United Nations, "The 17 sustainable development goals," United Nations, 2024. https://sdgs.un.org/goals

[2] J. C. Lam, K. K.W. Wan, C.L. Tsang, and Liu Yang, "Building energy efficiency in different climates, Energy Conversion and Management", Volume 49, Issue 8, Pages 2354-2366, ISSN 0196-8904, 2008.

[3] C. Cong, J. Page, Y. Kwak, B. Deal, and Z. Kalantari, "AI Analytics for Carbon-Neutral City Planning: A Systematic Review of Applications," Urban Science, vol. 8, no. 3, p. 104, Aug. 2024.

ARTIFICIAL INTELLIGENCE BASED PUBLICATION ANALYSIS FOR ZONGULDAK BÜLENT ECEVIT UNIVERSITY

Aysegul Tok*

Zonguldak Bulent Ecevit University, Department of Library, Zonguldak, Turkiye

**(Corresponding author) aysegul.tok@beun.edu.tr*

ABSTRACT

This study is a quantitative research that aims to analyze 1317 academic publications of Zonguldak Bülent Ecevit University between 2022-2024 using artificial intelligence (AI) tools. The analysis of academic publications is important for evaluating research results and identifying scientific contributions. The study compares the data obtained with Polymer and Julius artificial intelligence tools and evaluates the effectiveness of these tools in academic content analysis. The main data of the study were collected from the Incites database, which provides comprehensive research reviews, and structured according to variables such as author, year, title, etc. The data was prepared in a numerical format and made suitable for analysis. In the analysis process, artificial intelligence tools such as Polymer and Julius were used for citation analysis and topic modeling. The results obtained were compared with the analysis performed with Excel, one of the traditional data analysis methods. The results of the study will show the potential and effectiveness of artificial intelligence tools in the process of academic publication analysis.

Table 1: Study Findings

Affiliation	WOS Documents	Times Cited	Q1 Documents	Q2 Documents	Q3 Documents	Q4 Documents	Citation Impact	Top %10 Documents	Higly Cites Paper
Zonguldak Bülent Ecevit University	1317	4708	234	223	189	141	3,574791192	110	12

```
term_freq = Counter(key_terms)

# Get top 15 terms
top_terms = dict(sorted(term_freq.items(), key=lambda x: x[1], reverse=True)[:15])

# Create bar plot
plt.figure(figsize=(12, 6))
plt.bar(top_terms.keys(), top_terms.values())
plt.xticks(rotation=45, ha='right')
plt.title('Most Common Research Topics (2022-2024)')
plt.xlabel('Terms')
plt.ylabel('Frequency')
plt.tight_layout()
plt.show()

# Print top terms with counts
print("\
Top research terms and their frequencies:")
for term, count in top_terms.items():
    print(f"{term}: {count}")
```

Code Explanation

This code snippet processes article titles to extract and visualize the most common research topics by counting the frequency of significant terms.

Text Click to view

Image Click to view

The analysis extracted key terms from the article titles to identify subject trends, and a bar chart was generated to visualize the most common research topics. Below is the chart and the top terms with their frequencies.

Figure 1: AI tool code snippet that processes article headlines (Julius)

Keywords: Artificial intelligence, academic research, publication analysis, data mining, Zonguldak Bulent Ecevit University.

REFERENCES

[1] S. Uğur, and İ. Usta, "Bibliographic Analysis of Master's Theses on the Use of Artificial Intelligence in Education in Turkey", MANAS Journal of Social Studies, 14, 1, 143-163, 2025.

[2] P.R. Srivastava and D. Sharma "Intellectual Structure and Publication Pattern in the Journal of Global Information Management: A Bibliometric Analysis During 2002-2020". Journal of Global Information Management, 29, 4, 1-31, 2021.

[3] M. Yılmaz, "A Critical View on Bibliometrics". Turkish Librarianship, 33, 1, 43-49, 2019. http://www.tk.org.tr/index.php/TK/article/view/2986

[4] E. Demircioğlu, C. Yazıcı, and B. Demir, "Artificial Intelligence Supported Mathematics Education: A Content Analysis", JSHSR, 11, 106, 771-785, 2024.

[5] B. Kaya Erdem, "The Impact of Artificial Intelligence on Media and Publishing". TRT Academy, 6, 13, 896-903, 2021. https://doi.org/10.37679/trta.1002525

BLOK ZİNCİR TABANLI OYLAMA SİSTEMİ

Mehmet Rasim Yıldız*
Ondokuz Mayıs Üniversitesi, Bilgisayar Mühendisliği, Samsun, Türkiye

Meryem Soysaldı Şahin
Ondokuz Mayıs Üniversitesi, Bilgisayar Mühendisliği, Samsun, Türkiye
**(İletişimden sorumlu yazar) rasim.yildiz@bil.omu.edu.tr*

ÖZET

Geleneksel oylama sistemleri; şeffaflık, güvenlik ve veri bütünlüğü gibi temel unsurlar açısından çeşitli zafiyetler barındırmaktadır. Bu nedenle, merkeziyetsiz, değiştirilemez ve şeffaf bir veri yapısı sunan blok zincir (blockchain) teknolojisi, elektronik oylama sistemlerinin güvenirliğini artırmak amacıyla önemli bir alternatif olarak öne çıkmaktadır. Özellikle şifreleme algoritmalarıyla desteklenen blok zincir tabanlı sistemler, oyların gizliliğini ve bütünlüğünü sağlamada kritik bir rol üstlenmektedir.

Bu çalışmada, blok zincir teknolojisi kullanılarak geliştirilen güvenli bir elektronik oylama sistemi geliştirilmiştir. Sistem, her bir oyun blok zinciri yapısı üzerinden doğrulanarak kayıt altına alınmasını sağlamakta ve bu sayede güvenli bir oy kullanma süreci hedeflenmektedir. Uygulamada, kullanıcılar sisteme kayıt olduktan sonra giriş yaparak oy verme ekranına yönlendirilmekte; aday olmak isteyen kullanıcılar ise sistem üzerinden adaylık başvurusu gerçekleştirebilmektedir. Oylama işlemi tamamlandıktan sonra, sonuç sayfasında en çok oy alan aday kullanıcıya sunulmaktadır.

Sistem, kullanıcı şifrelerinin ve kişisel bilgilerinin güvenliğini ön planda tutmakta; tüm veriler blok zincir altyapısı ile şifrelenerek korunmaktadır. Ayrıca, oylama sürecinin verimli ve sürdürülebilir şekilde yürütülmesini sağlamak amacıyla, sistem kaynaklarını zorlamayan bir zamanlayıcı mekanizması entegre edilmiştir. Geliştirilen sistem sayesinde, oy verileri bloklar hâlinde güvenli biçimde zincire eklenmiş; her oy, dijital imzalar ve hash algoritmalarıyla korunarak değiştirilemez hâle getirilmiştir. Oylama sonuçları, şeffaflık ilkesi doğrultusunda herkes tarafından denetlenebilir biçimde sunulmaktadır. Bu çalışma ile blok zincir teknolojisinin elektronik oylama sistemlerine entegrasyonu sağlanarak; güvenli, şeffaf ve müdahaleye kapalı bir elektronik oylama ortamı oluşturulmuştur.

Anahtar Kelimeler: Blok zincir, oylama sistemi, veri güvenliği.

REFERANSLAR

[1] C. Angsuchotmetee, P. Setthawong and S. Udomviriyalanon, "BlockVOTE: An Architecture of a Blockchain-based Electronic Voting System," in 23rd International Computer Science and Engineer-ing Conference (ICSEC), 2019, 110-116.

[2] D. Seftyanto, A. Amiruddin and A. R. Hakim, "Design of Blockchain-Based Electronic Election System Using Hyperledger: Case of Indonesia," in 4th International Conference on Information Technology, Information Systems and Electrical Engineering (ICITISEE), 2019, 228-233.

[4] C. Killer, B. Rodrigues, R. Matile, E. Scheid, & B. Stiller. "Design and implementation of cast-as-intended verifiability for a blockchain-based voting system", in Proceedings of the 35th annual ACM symposium on applied computing, 2020, 286-293.

[5] G. Rathee, R. Iqbal, O. Waqar and A. K. Bashir, "On the Design and Implementation of a Blockchain Enabled E-Voting Application Within IoT-Oriented Smart Cities," in IEEE Access, vol. 9, 34165-34176, 2021.

CREDIT CARD FRAUD DETECTION USING MACHINE LEARNING METHODS

Abdullah Alawad
Bilecik Seyh Edebali University, Computer Engineering Department, Bilecik, Türkiye
Vedat Marttin*
Bilecik Seyh Edebali University, Computer Engineering Department, Bilecik, Türkiye
(Corresponding author) vedat.marttin@bilecik.edu.tr

ABSTRACT

Machine learning (ML) has rapidly evolved, leading to transformative applications across numerous domains. Deep learning (DL) methods [1], in particular, have enabled the training of complex models on large-scale datasets, resulting in significant advances in areas such as computer vision, natural language processing, and speech recognition. This progress has increased the need for accurate and efficient modeling techniques, especially when working with vast amounts of data.

A review of the literature shows that various ML and DL methods have been employed for credit card fraud detection. Among the most commonly used datasets is Creditcard.csv [2], which contains 284,807 credit card transactions from Europe and 30 anonymized features transformed using Principal Component Analysis (PCA). The dataset's 'Class' variable labels transactions as legitimate (Class = 0) or fraudulent (Class = 1).

In fraud detection, ML and DL-based models are among the most widely adopted approaches. Notable algorithms include Logistic Regression [3], Naive Bayes, K-Nearest Neighbors (KNN), Random Forest, Support Vector Machines (SVM) [4], Decision Trees [5], and Autoencoders. In this study, the dataset underwent preprocessing, and the performance of 10 different machine learning algorithms was evaluated. The implementations were carried out in Python using relevant libraries. As shown in Table 1, XGBoost, Random Forest, and KNN outperformed the other methods in terms of classification performance.

Table 1: Study Findings

Methods	Class	Precision	Recall	F1-Score	Methods	Class	Precision	Recall	F1-Score
Logistic Regression	0	1.00	1.00	1.00	Naive Bayes	0	1.00	0.98	0.99
	1	0.83	0.44	0.57		1	0.06	0.82	0.11
KNN	0	**1.00**	**1.00**	**1.00**	Gradient Boosting	0	1.00	1.00	1.00
	1	**0.97**	**0.76**	**0.85**		1	0.74	0.60	0.66
Decision Tree	0	1.00	1.00	1.00	XGBoost	0	**1.00**	**1.00**	**1.00**
	1	0.71	0.82	0.76		1	**0.97**	**0.79**	**0.87**
Random Forest	0	**1.00**	**1.00**	**1.00**	LightGBM	0	1.00	1.00	1.00
	1	**0.97**	**0.78**	**0.86**		1	0.32	0.60	0.42
SVM	0	1.00	1.00	1.00	AutoEncoder	0	1.00	1.00	1.00
	1	0.80	0.80	0.80		1	0.46	0.39	0.42

Keywords: Machine learning (ML), Artificial intelligence (AI), Credit card fraud detection

REFERENCES

[1] Y.LeCun, Y.Bengio, and G.Hinton. "Deep learning". Nature, 521(7553):436–444, 2015.

[2] Anomaly Detection Credit Card Fraud Dataset, https://www.kaggle.com/code/annastasy/anomaly-detection-credit-card-fraud/input (Access Time: 10.04.2025)

[3] M.S.Kumar, V. Soundarya, S. Kavitha, ES Keerthika, and E Aswini. "Credit card fraud detection using random forest algorithm". In 2019 *3rd International Conference on Computing and Communications Technologies (ICCCT),* pages 149–153. IEEE, 2019.

[4] F. I. Meenakshi, and S. Singh. "Comparison and analysis of logistic regression, naive bayes and knn machine learning algorithms for credit card fraud detection". *International Journal of Information Technology,* 13(4):1503–1511, 2021.

[5] Z. Xie and X. Huang, "A Credit Card Fraud Detection Method Based on Mahalanobis Distance Hybrid Sampling and Random Forest Algorithm," in *IEEE Access*, vol. 12, pp. 162788-162798, 2024, Doi: 10.1109/ACCESS.2024.3421316.

DETERMINING THE OPTIMAL NUMBER OF CHARGING UNITS AT ELECTRIC VEHICLE CHARGING STATIONS USING A QUEUEING THEORY APPROACH

Bayram Dündar*
Bartın University, Bartın, Türkiye
**(Corresponding author) bdundar@bartin.edu.tr*

ABSTRACT

Decision-makers are seeking solutions in the transportation sector to mitigate the negative consequences of climate change and the growing energy crisis. Since electric vehicles are considered more efficient both environmentally and economically—and with advancements in technology—a rapid increase has been observed in the number of these vehicles on the market. To meet the charging needs of electric vehicles, reduce customer waiting times, and ensure cost efficiency, it has become essential for decision-makers to provide the optimum number of charging units. This study proposes a queuing theory-based approach to estimate the optimal number of charging units needed to maintain the average waiting time of users below a predefined service level threshold. To evaluate the performance of electric vehicle charging stations, this study addresses the problem using the M/M/s/C queuing model approach. In this model, the interarrival times of successive vehicles and the charging durations are assumed to follow a negative exponential distribution. The model is applied to several electric vehicle charging stations in Bartin. In Bartın province, the monthly predicted number of visits to charging stations by electric vehicles is approximately 2,300, according to TÜİK data. According to the current data, there are a total of 10 charging station points and a total of 11 charging units at these stations throughout the province. According to the queuing model analysis, waiting times and queuing are not currently observed at the stations. However, if the increase in the number of electric vehicles continues with the current trend, waiting time (W_q) values are expected to emerge within the framework of the queuing model starting next year. The findings are expected to support decision-makers in optimizing electric vehicle charging infrastructure under growing demand conditions.

Keywords: Optimization, queuing theory, electric vehicles.

DEVELOPMENT OF AN AI-BASED SYSTEM FOR AUTOMATED COOKING USING COLOR TRACKING

Fırat Koçyiğit*
Vestel Whıte Goods, Technology Development Department, Manisa, Türkiye

Ayşegül Alaybeyoğlu
İzmir Kâtip Çelebi University, İzmir, *Türkiye*
(Corresponding author) firat.kocyigit@vestel.com.tr

ABSTRACT

Introduction: This study develops a CNN-based model to determine food cooking status by analyzing color changes for more objective and reliable results. Data augmentation and various network architectures were tested to optimize prediction performance. Results show that CNN-based image processing enhances quality control in the food industry. The proposed AI solution has potential applications in kitchen automation and smart cooking systems.

Literature Research: Automated systems using computer vision and AI assess food color and cooking conditions efficiently. Studies highlight real-time image processing for tracking color changes. Using HSV transformation and segmentation, the system accurately detects cooking stages and readiness. This validates color-based food analysis with machine learning. [1-3].

Method: This study used CNN to predict food cooking status. Data was normalized, partitioned, and optimized to prevent overfitting. Accuracy was evaluated using activation functions and performance metrics. Key parameters were tuned for optimal results, ensuring reliable predictions.

Results: This study used CNN to predict oven-baked food status. Different configurations were tested, and key parameters optimized.(Figure-1) The system suits home and industrial kitchens and can be improved with larger datasets and sensors.

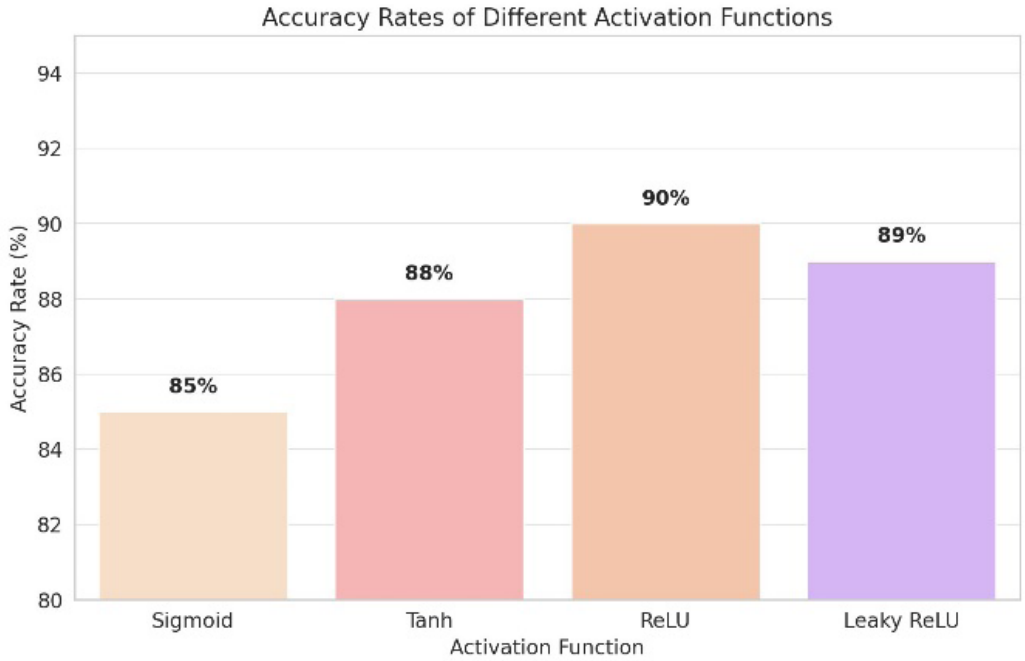


Figure 1. Example Results

Keywords: Convolutional neural networks, segmentation, activation functions.

REFERENCES

[1] Lin, C.-S., et al. (2021). "A Study of Automatic Judgment of Food Color and Cooking Conditions with Artificial Intelligence Technology." Processes, 9(1128), MDPI.

[2] Li, H.; Lee, W.S.; Wang, K. Identifying blueberry fruit of different growth stages using natural outdoor color images. Comput. Electron. Agric. 2014

[3] Annamalai, P.; Lee, W.S.; Burks, T.F. Color Vision Systems for Estimating Citrus Yield in Real-time. In Agricultural and Biological Engineering; University of Florida: Gainesville, FL, USA, 2004

DİJİTAL EKONOMİ VE BLOK ZİNCİRİ UYGULAMALARI

Hakan Can Altunay*
Ondokuz Mayıs Üniversitesi, Çarşamba Ticaret Borsası Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, Samsun, Türkiye
(İletişimden sorumlu yazar) hakancan.altunay@omu.edu.tr

ÖZET

Blok zinciri yönetimi ve ödeme sistemlerinin iyileştirilmesinde ortaya çıkan veri güvenliğini artırma probleminde, yapay zeka önemli bir potansiyele sahiptir. Yapay zeka ve blok zinciri teknolojisi birbirine entegre biçimde çalışır. Blok zinciri ile veri güvenliği merkeziyetsiz bir şekilde sağlanırken, verilerin analizi kısmında yapay zeka kullanılır. Bu çalışmada ekonomik işlemler, blok zinciri izlenebilirliği ve Nesnelerin İnterneti (IoT) içindeki güvenlik uygulamalarına odaklanılırken, yapay zeka ve blok zinciri arasındaki çalışma ilişkisi de incelenmiştir. Herhangi bir kurum ya da kuruluş, potansiyel risklerinin belirlenmesinde yapay zeka teknolojisini, siber saldırılardan korunup güvenliğini artırmak için ise blok zincir teknolojisini kullanabilir. Geleneksel sistemlerle entegrasyon ve ölçeklenebilme problemi halen önemli bir sorun olarak karşımıza çıkmaktadır. Fakat yapay zeka ve blok zinciri birleşimi, dijital ekonomi alanında gelecekte karşılaşılabilecek taleplere önemli bir çözümdür. Ayrıca, yapay zeka ve blok zinciri birleşiminin güçlü yönleri açıklanarak, günümüzde karşılaşılan güvenlik problemleri ve verimlilik sorununun üstesinden gelmek ve gelecekte yaşanabilecek problemlerin çözümü için nasıl bir temel oluşturulabileceği konusunda detaylı bir çerçeve sunulmaktadır.

Anahtar Kelimeler: Veri güvenliği, yapay zeka, blok zinciri.

REFERANSLAR

[1] C.K. Chuen, Alvin, A.K.M. Yusoff, and H. Webb, "The Convergence of Artificial Intelligence and Cryptocurrency: Transforming Digital Innovation and Security in Blockchain Ecosystems", The Palgrave Handbook of FinTech in Africa and Middle East: Connecting the Dots of a Rapidly Emerging Ecosystem, Singapore: Springer Nature Singapore, 1-18, 2025.

[2] L.U. Ashrapova, and N. M. Apsilyam, "Cybersecurity in the Digital Economy: New Challenges and Solutions." International Scientific-Electronic Journal Pioneering Studies and Theories, 1. 2, 9-13, 2025.

[3] Yang, Fan, et al. "Blockchain and Machine Learning in the Green Economy: Pioneering Carbon Neutrality Through Innovative Trading Technologies", IEEE Transactions on Engineering Management, 2025.

[4] H.C. Altunay, "Analysis of Cyber Attacks Using Honeypot", Black Sea Journal of Engineering and Science, 7(5), 954-959, 2024.

[5] Rizky, Agung, et al. "Optimizing Blockchain Digital Signature Security in Driving Innovation and Sustainable Infrastructure", Blockchain Frontier Technology, 4.2, 183-192, 2025.

[6] Zhang, Tianyu. "Analysis of the Realization for Trading Security of Cryptocurrency", Advances in Economics, Management and Political Sciences, 149, 38-47, 2025.

[7] Pineda, Magda, D.Jabba, and W.Nieto-Bernal, "Blockchain architectures for the digital economy: trends and opportunities", Sustainability, 16.1, 442, 2024.

[8] Jin, Wenjin, "Security and privacy of digital economic risk assessment system based on cloud computing and blockchain", Soft Computing, 28.3, 2753-2768, 2024.

FİNANSAL KİŞİLİK PROFİLLERİ: KLONTZ MONEY SCRİPT INVENTORY VE YAPAY ZEKÂ DESTEKLİ SINIFLANDIRMA İLE FİNANS ÖĞRENCİLERİNİN PARAYA DAİR İNANÇLARININ PSİKOLOJİK ANALİZİ

Özge Ece Çiftci*

Zonguldak Bülent Ecevit Üniversitesi, Psikolojik Danışma ve Rehberlik uygulama ve Araştırma Merkezi, Zonguldak, Türkiye

Ersin Kanat

Zonguldak Bülent Ecevit Üniversitesi, Uygulamalı Bilimler Yüksekokulu, Zonguldak, Türkiye

(İletişimden sorumlu yazar) ece.ciftci@beun.edu.tr

ÖZET

Bu çalışma, finans alanında eğitim gören üniversite öğrencilerinin paraya dair bilinç dışı inançlarını incelemeyi ve bu inançlara dayalı olarak psikolojik temelli finansal kişilik profilleri oluşturmayı amaçlamaktadır. Ölçüm aracı olarak Klontz Money Script Inventory (KMSI) kullanılmış, katılımcıların paraya yönelik inançları dört boyutta değerlendirilmiştir: Paradan Kaçınma (Money Avoidance), Paraya Tapma (Money Worship), Para = Statü (Money Status) ve Parasal Dikkat (Money Vigilance).

Veri analizi sürecinde, KMSI'dan elde edilen skorlar yapay zeka destekli algoritmalar aracılığıyla işlenmiş ve öğrenciler beş psikolojik-finansal profil altında sınıflandırılmıştır:

Güvenlik Odaklı, (parasal güvenliği önceleyenler),Statü Arayıcı, (parayı sosyal değerle ilişkilendirenler), Duygusal Harcayıcı, (anlık haz ve harcamayla rahatlayanlar),Sorgulayan Kaçınmacı, (paradan uzak durmayı tercih edenler) ,Çelişkili Kimlik. (birden fazla boyutta yüksek skor gösteren ve kararsız davranış sergileyenler)

Yapay zekâ destekli sınıflandırma sürecinde, katılımcıların KMSI alt boyutlarındaki puanları dikkate alınarak kümeleme algoritmaları ve karar ağacı modelleri kullanılmış; böylece veri temelli, nesnel ve tekrar edilebilir bir sınıflandırma elde edilmiştir. Bu sınıflandırma, öğrencilerin parayla kurduğu psikolojik ilişkiyi anlamaya yardımcı olurken, aynı zamanda finansal eğitim içeriklerinin kişiselleştirilmesine katkı sağlamaktadır.

Araştırmada veri toplama aracı olarak Klontz Money Script Inventory (KMSI) kullanılmıştır. Katılımcılar, Zonguldak Bülent Ecevit Üniversitesi'nde öğrenim gören Finans ve Bankacılık Bölümü öğrencilerinden oluşmaktadır. KMSI ölçeği Türkçeye uyarlanarak uygulanmıştır.

Anahtar Kelimeler: Finansal kişilik profilleri, güvenlik odaklı, statü arayıcı, duygusal harcayıcı, sorgulayan kaçınmacı, çelişkili kimlik.

FROM ZERO-DIVISOR GRAPHS TO TOPOLOGICAL INDICES: A STRUCTURAL STUDY OF SEMIGROUP-DERIVED GRAPHS

Seda Oguz Unal*
Sivas Cumhuriyet University, Department of Mathematics and Science Education, Sivas, Turkey
**(Corresponding author) sdaoguz@gmail.com*

ABSTRACT

Topological indices have played a key role in chemistry for over fifty years [4], but their mathematical foundations have gained deeper attention in recent research. These indices help uncover structural properties of molecules, providing valuable insights across industrial science, applied physics, biochemistry, and even environmental and toxicological studies [5]. Such topological indices, particularly those incorporating both degree and distance information, remain a subject of considerable interest in both theoretical and applied contexts. Motivated by the concept of zero-divisor graphs (see [1]), our work focuses on monogenic semigroup graphs, which play a significant role in the study of algebraic structures. Following the approach of [2], we adopt vertex adjacency rules while preserving the core framework. In this talk, we explore the interplay between algebraic structures and topological indices, particularly in graphs derived from semigroups. We establish exact formulas and fundamental structural results to advance theoretical developments.

Keywords: Monogenic semigroups, graphs, indices.

REFERENCES

[1] I. Beck, "Coloring of Commutating Rings", J. Algebra, 116, 208–226, 1988.

[2] K. C. Das, N. Akgüneş, A.S. Çevik, "On a graph of monogenic semigroup", J. Ineq. Appl., 44, 2013.

[3] I. Gutman, "Selected Properties of the Schultz Molecular Topological Index", J. Chem. Inf. Comput. Sci.,34, 1087-1089, 1994.

[4] R. Todeschini, V. Consonni, "Molecular Descriptors for Chemoinformatics", Wiley VCH: Weinheim, Germany, 2009.

[5] J. Devillers, A.T. Balaban, "Topological Indices and Related Descriptors in QSAR and QSPR", Gordon and Breach: Amsterdam, NY, USA , 1999.

GENOTOXICITY OF FIG SEED OIL AND THE COMPARISON OF DIFFERENT ARTIFICIAL INTELLIGENCE TOOLS FOR DATA ANALYSIS

Gamzegül Açış*
Balıkesir University, Molecular Biology and Genetics, Balıkesir, Türkiye

Begümhan Yılmaz Kardaş
Balıkesir University, Molecular Biology and Genetics, Balıkesir, Türkiye

Mehmet Emin Diken
Balıkesir University, Molecular Biology and Genetics, Balıkesir, Türkiye
(Corresponding author) ggamzeacis@gmail.com

ABSTRACT

The fig (Ficus spp. , Family: Moraceae) has been used medicinally to treat ailments such as sore throat, cough, colic, appetite loss, and diarrhea, due to their anti-inflammatory and antispasmodic properties. Fig seeds are rich in fats, proteins, tocopherols, and essential amino acids. Fig seed oil contains 14 fatty acids, including alpha-linolenic acid (ALA), an omega-3 derivative, and supports energy production, cellular regeneration, brain health, vascular flexibility, and blood pressure regulation. Its high potassium content is linked to lower insulin levels and improved lipid profiles, including increased HDL and decreased LDL cholesterol.

Antigenotoxicity tests evaluate whether a substance can protect DNA from damage caused by genotoxic agents. One such test is the somatic mutation and recombination test (SMART), a sensitive assay using Drosophila melanogaster as a model. This study aimed to assess the antigenotoxic potential of fig seed oil using the SMART assay. The frequencies and the statistical analyses were calculated using Excel. In testing the one-way alternative hypothesis, Student's t-test was used with the two-sample equal variance option. All tests were performed by performing 3 biological replicates. The results showed that fig seed oil is non-genotoxic and significantly reduces genetic damage induced by hydrogen peroxide (H_2O_2). Then, the Large Language Model (LLM) tools—ChatGPT, DeepSeek, and Microsoft Copilot—were used to calculate the results. All programs accurately computed frequency values, though there were minor variations in inhibition percentages. When ChatGPT was asked to administer the Student ttest, it correctly predicted statistical significance ($p<0.05$). However, DeepSeek and Microsoft Copilot could not conduct statistical analyses directly but offered methodological guidance.These findings highlight both the antigenotoxic potential of fig seed oil and the emerging role of LLM tools in genetic and molecular biology research.

Keywords: Fig seed oil, artificial intelligence, SMART antigenotoxicity. biology.

REFERENCES

[1] Nakilcioğlu-taş, E. (2019). Biochemical Characterization of Fig (Ficus carica L.) Seeds. Journal of Agricultural Sciences, 25(2), 232-237. https://doi.org/10.15832/ankutbd.398268

[2] Mert H, Mert N, Cibuk S, Yildirim S, Mert N. Antidiabetic Effect of Fig Seed Oil in Rats with Diabetes Induced by Streptozotocin. J Oleo Sci. 2024;73(5):717-727. doi: 10.5650/jos.ess23255. PMID: 38692894.

[3] 5.Yılmaz Çetinkaya, A., & Yurtsever, S. (2021). Bazı gıda katkı maddelerinin genotoksisitesini araştırmak için kullanılan Drosophila melanogaster'de somatik mutasyonlar ve rekombinasyon testi. Uluslararası Tarım Çevre ve Gıda Bilimleri Dergisi, 5(1), 65-73. https://doi.org/10.31015/jaefs.2021.1.9

HASTANELERDE HİJYEN VE GÜVENLİK İÇİN YAPAY ZEKÂ DESTEKLİ KAMERA TABANLI KORUYUCU EKİPMAN KONTROL SİSTEMİ

Muhsin Kavan
İzmir Kâtip Çelebi Üniversitesi, Yazılım Mühendisliği, İzmir, Türkiye

Ayşegül Alaybeyoğlu
İzmir Kâtip Çelebi Üniversitesi, Yazılım Mühendisliği, İzmir, Türkiye
**(İletişimden sorumlu yazar) muhsinkavan@gmail.com*

ÖZET

Hastanelerde cerrahlar, hemşireler ve diğer sağlık personeli, enfeksiyon riskini azaltmak için maske, eldiven, bone ve önlük gibi koruyucu ekipmanları eksiksiz şekilde kullanmakla yükümlüdür. Ancak zaman zaman bu kurallara uyulmaması veya eksik ekipman kullanımı nedeniyle enfeksiyon yayılma riski artmaktadır. Bu çalışmada, hastanelerde hijyen ve güvenlik standartlarını artırmak amacıyla yapay zeka destekli, kamera tabanlı bir koruyucu ekipman kontrol sistemi geliştirilmiştir.

Sistem, sağlık personelinin ekipman kullanımını kameralar aracılığıyla gerçek zamanlı izleyerek eksik veya hatalı kullanım durumlarında otomatik olarak uyarı verebilmekte ya da yönetim birimlerine bildirim gönderebilmektedir. Bu sayede, hastane ortamında anlık müdahale olanağı sağlanmakta ve enfeksiyon riski minimize edilmektedir.

Giriş: Bu çalışmanın amacı, yapay zeka ve görüntü işleme tekniklerini kullanarak sağlık personelinin koruyucu ekipman kullanımını denetleyen akıllı bir sistem tasarlamaktır [1]. Sistem, hem hastane çalışanlarının kurallara uygunluğunu sürekli denetleyecek hem de hijyen protokollerine katkı sağlayacaktır.

Metot: Çalışmada, sağlık personelinin koruyucu ekipmanlı ve ekipmansız görsellerinden oluşan bir veri seti oluşturulmuştur. Görüntü işleme aşamasında OpenCV ve TensorFlow/ PyTorch kütüphaneleri kullanılarak analiz gerçekleştirilmiştir. Maske, eldiven, bone ve önlük tespiti için YOLO ve Faster R-CNN gibi nesne tanıma algoritmaları uygulanmıştır [2]. Gerçek zamanlı analiz, hastane kameralarından gelen görüntüler üzerinde gerçekleştirilmiş ve eksik ekipman kullanımı durumunda sistem otomatik uyarı ve bildirim mekanizmalarını devreye sokmuştur.

Sonuç: Geliştirilen sistem, hastane ortamında yapılan testlerde yüksek doğruluk oranlarıyla çalışmış ve hijyen takibine büyük katkı sağlamıştır. Bu yapay zeka destekli yaklaşım, sağlık sektöründe enfeksiyon kontrolü için etkili bir dijital çözüm sunmakta olup, farklı sağlık kuruluşlarına entegre edilebilecek potansiyele sahiptir.

Anahtar Kelimeler: Görüntü işleme, koruyucu ekipman, derin öğrenme, nesne tanıma, gerçek zamanlı izleme, sağlık güvenliği, otomatik uyarı sistemi.

REFERANSLAR

[1] Richard Szeliski, *Computer Vision: Algorithms and Applications*, 2. Baskı, Springer, 2022.

[2] Joseph Redmon ve Ali Farhadi, "YOLOv3: An Incremental Improvement," *arXiv preprint arXiv:1804.02767*, 2018.

IN SILICO TARGET PROFILING OF A NOVEL DITHIOCARBAMATE DERIVATIVE (GKF25) VIA AI-DRIVEN MOLECULAR DOCKING FOR NEUROPSYCHIATRIC APPLICATIONS

Unzile Yaman*
İzmir Kâtip Çelebi University, Faculty of Pharmacy, Department of Pharmaceutical Toxicology, İzmir, Türkiye

Gulsah Karakaya
İzmir Kâtip Çelebi University, Faculty of Pharmacy, Department of Pharmaceutical Chemistry, İzmir, Türkiye

Ofcan Oflaz
Lokman Hekim University, School of Medicine, Department of Basic Medical Sciences, Department of Medical Biology, Ankara, Türkiye
**(Corresponding author) unzile.yaman@ikcu.edu.tr & unzileyaman@gmail.com*

ABSTRACT

Introduction: Early identification of bioactive compounds targeting central nervous system (CNS)-related proteins is critical in the development of novel therapeutic strategies for neurodegenerative and psychiatric disorders. In this study, we present the computational evaluation of GKF25, a novel fluorophenylpiperazine-containing dithiocarbamate derivative. Molecular docking, a data-driven *in silico* approach, was employed to investigate its potential interactions with key CNS-related protein targets. This study aims to reduce experimental burden in the drug discovery process by prioritizing high-affinity targets prior to *in vitro* validation, aligning with current trends in artificial intelligence and data-supported molecular modeling.

Methods: Protein target prediction was conducted using SwissTargetPrediction, identifying serotonin 2B receptor (5-HT2B), MAP kinase ERK1, monoamine oxidase B (MAO-B), and cannabinoid receptor 1 (CB1) as the most likely targets. The 3D structure of GKF25 was prepared using ChemDraw and molecular docking studies were performed comperatively to known reference ligands using CB-Dock, a cavity-detection–assisted AutoDock Vina-based platform.

Results and Conclusion: Docking simulations revealed that GKF25 showed better docking scores compared to Lisuride (5-HT2B), Temuterkib (ERK1), Rasagiline and Selegiline (MAO-B), and THC (CB1) (Figure 1).

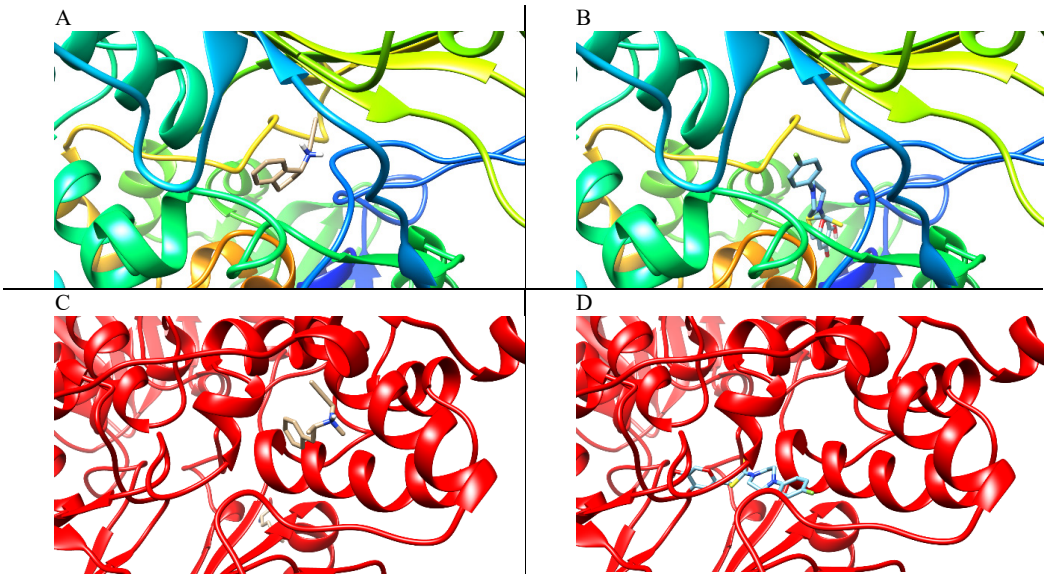


Figure 1: Monoamine Oxidase B protein model analysis.
(A) Rasagiline-2xfq (B) GKF25-2xfq (C) Selegiline-2byb (D) GKF25-2byb

These findings indicate that GKF25 may interact with neurotransmitter systems and signaling pathways linked to neuroprotection and neuroinflammation. AI-integrated docking supports its potential as a neuroactive candidate. Further in vitro and in vivo studies are needed to confirm its pharmacological and toxicological profile.

Keywords: Molecular docking, dithiocarbamate, neurodegeneration, artificial intelligence, *in silico* target prediction.

İŞLETMELERDE YAPAY ZEKÂ ÇAĞI

İlknur Zeydan*
Ağız ve Diş Sağlığı Uygulama ve Araştırma Merkezi, Zonguldak Bülent Ecevit Üniversitesi, Zonguldak, Türkiye
**(İletişimden sorumlu yazar) ilknur.oral@beun.edu.tr*

ÖZET

Yapay zekâ, karmaşık sorunlara ve görevlere yaklaşım tarzımızı dönüştürmeyi ustaca gerçekleştirmiş teknolojik bir gelişimdir. Bireyler kadar işletmeler de bu gelişimden nasibini almaktadır. İşletmelerin gelişimi, çağa ayak uydurabilmesi ve bilinçli karar alma düzeylerini geliştirebilmeleri için yapay zekânın kullanımı oldukça önemlidir. Bazı işletmeler yapay zekâyı iş süreçlerinde kullanmaya başlamışken bazıları henüz kullanmaya başlamamıştır. İşletmelerin çoğunluğu yapay zekânın ne olduğunu, nasıl ve nerelerde kullanacağını anlamaya çalışmakla meşguldür. İşletmeler yapay zekâyı bünyelerine dâhil etmek için çalışmalarda bulunmaktadır. 2024 yılında Avrupa Birliği'nde bulunan işletmelerin %13,48'i yapay zekâ teknolojilerini kullanmaktadır. Bu işletmelerin %41,17'si büyük ölçekli işletmelerdir. 2024 yılında yapay zekâyı en çok kullanan sektör, bilgi ve iletişim sektörü olmuştur. Türkiye'deki işletmeler de yapay zekânın sunduğu fırsatları benimsemeye başlamışlardır. 2024 yılında Türkiye'deki işletmelerin %17,5'i yapay zekâdan faydalanmıştır. Türkiye'de yapay zekânın en çok kullanıldığı alanlar %53 ile müşteri ilişkileri ve pazarlama olmuştur. Bu çalışmanın amacı işletmelere yapay zekânın önemini kavratmak, yapay zekânın hangi alanlarda ve nasıl kullanılabileceğini göstermektir. Bu çalışmada yapay zekâ nedir? İşletmelerde nasıl kullanılır? İşletmelerde uygulama alanları nelerdir? gibi sorulara yanıtlar verilecektir. Yapay zekânın işletmelerde kullanılmasının avantajlarından bahsedilecektir.

Anahtar Kelimeler: Yapay zekâ, işletmeler, yapay zekâ kullanım alanları, uygulamalar.

LEARNING ENGAGEMENT AND ANALYSIS FOR REAL-TIME NAVIGATION (LEARN)

Fatih Emre Akarsu*
Robert Collage, İstanbul, Türkiye
Emre Atlıer Olca
Software Engineering Department, Maltepe University, İstanbul, Türkiye
(Corresponding author) akafat.28@robcol.k12.tr

ABSTRACT

This project aims to automate student attendance in classroom environments and analyze students' attention and emotional states during lessons using artificial intelligence-powered image processing technology. Face recognition technology has been utilized to conduct attendance processes quickly and accurately, minimizing errors associated with traditional attendance methods such as paper-based or manual tracking [1]. By reducing time lost in attendance-taking, this implementation enhances teachers' time management and allows lesson durations to be used more efficiently.

Beyond attendance automation, this project also focuses on evaluating students' attention and motivation levels through emotional state analysis. Many studies have explored the use of image processing and emotion detection techniques in educational environments, demonstrating their potential to improve learning experiences. The collected data supports the identification of individual learning needs, enabling the development of personalized educational plans [2-3]. For instance, students with lower attention levels can benefit from more focused teaching strategies designed to enhance engagement.

Technically, the system collects data via cameras integrated into classrooms and processes it using artificial intelligence algorithms. The face recognition component achieves high accuracy and rapid analysis, even under varying lighting conditions and classroom settings [4]. Additionally, this project integrates emotion-based interaction analysis, offering a unique approach that combines attendance automation with adaptive learning insights. Compared to conventional attendance systems, which are often time-consuming and prone to errors, this method ensures more reliable student tracking while contributing to a more personalized educational approach.

All data processing procedures comply with the Personal Data Protection Law, with data security prioritized [5]. The findings indicate that integrating AI-driven image processing and emotion detection into classroom settings enhances both time efficiency and the overall learning experience. By merging automation with personalized career guidance, this project provides an innovative solution to optimize education through technology.

Keywords: Face recognition, artificial intelligence, emotional analysis, educational technology.

REFERENCES

[1] Parvinzamir, F. (2012). Finger-print based student attendance register

[2] Akinduyite, C., Adetunmbi, O., Olabode, O., & Ibidunmoye, E. (2013) Fingerprint-Based Attendance Management System

[3] Lim, T., Sim, S., & Mansor, M. (2009). RFID based attendance system

[4] Saparkhojayev, N., & Guvercin, S. (2012). Attendance Control System based on RFID-technology

[5] Türkiye Cumhuriyeti, Kişisel Verilerin Korunması Kanunu, Kanun No. 6698, Kabul Tarihi: 24/03/2016, Resmi Gazete Tarihi: 07/04/2016, Sayı: 29677.

MAKİNE ÖĞRENMESİ İLE DESTEKLENEN ÖMÜR DEVRİ YÖNETİMİ VE ÖNLEYİCİ BAKIM PLANLAMASINDA TAŞIMA VERİMLİĞİNİN ARTTIRILMASI

Sevdenur Uzer*
Balıkesir Üniversitesi, Fen Bilimleri Enstitüsü, Matematik Bölümü, Balıkesir, Türkiye

Fırat Ateş
Balıkesir Üniversitesi, Matematik Bölümü, Balıkesir, Türkiye
**(İletişimden sorumlu yazar) sevdenur.uzerr@outlook.com*

ÖZET

Lojistik oldukça eski bir tarihçeye sahip olmasına rağmen önemi 1. ve 2. Dünya Savaşları esnasında ortaya çıkmıştır. Bu savaşlar sırasında askeri mühimmatların uzun mesafelere taşınması, tedariği ve depolamanması oldukça büyük önem kazanmıştır. Bu açıdan yeni teknikler geliştirilmesi gerekmiştir. Geliştirilen tekniklerle beraber kullanılan araçların bakım ve arıza tamirini öngörememe sorunları ortaya çıkmıştır. Bu bakımdan lojistik alanında ana sistemler ve sistem parçaları için ömür devri yönetimi ve önleyici bakım planlarının geliştirilmesi son derece önemlidir. Bu amaçla lojistik alanında makine öğrenmesinin kullanımlarına baktığımızda, makine öğrenimi tabanlı optimizasyon algoritmalarının sürdürülebilir lojistiğe katkıda bulunduğu görülmekte olup bu katkının en başını "taşıma verimliliğini " çekmektedir. Dolayısıyla bu alanda taşıma verimliliğini arttırma problemi ortaya çıkmaktadır. Bu problemin çözümüne dair makine öğrenimi tabanlı optimizasyon algoritmalarının geliştirilmesi son derece önemlidir. Çünkü uzun mesafeli taşımacılıkla ilgili sektörlerde taşıma, tedarik zincirinde karbon emisyonlarının önemli bir kısmından sorumludur. Özellikle taşıma süreçlerinin daha etkin yönetilmesi ve sistem bileşenlerinin kullanım ömrü boyunca karşılaşılabilecek olası arızaların önceden tahmin edilerek önleyici bakım planlamalarının yapılması üzerine odaklanılmıştır. Bu bağlamda, yapay zekâ ve makine öğrenmesi temelli tekniklerin sağladığı avantajlardan yararlanılarak, taşıma süreçlerinin iyileştirilmesi ve sistemlerin sürdürülebilirliğinin artırılması amaçlanmaktadır. Yapılan çalışmada hazır veri setleri kullanılarak makine öğrenmesi ile önleyici bakım planları oluşturmak amaçlanmıştır. Bu amaç doğrultusunda regresyon modelleri için çeşitli hata metrikleri oluşturulmuş ve modeller karşılaştırılarak en iyi çalışan modele ulaşılmaya çalışılmıştır.

Anahtar Kelimeler: Makine öğrenmesi, lojistik, taşıma verimliliği, önleyici bakım planlaması.

Bu çalışma Cumhurbaşkanlığı Savunma Sanayi Başkanlığı, SAYZEK Akademik Tez Programı kapsamındadır.

REFERANSLAR

[1] Uçan, A. (2015). Türkiye'de Performansa Dayalı Lojistik Yönetimi: Kamu ve Özel Sektör Uygulamaları. Ankara: Türk Hava Kurumu Üniversitesi, Sosyal Bilimler Enstitüsü.

[2] Singh, R. (2024). Machine Learning and Sustainable Logistics: Analyzing the Role of Optimization Algorithms in Supply Chain Management.

[3] Özaslan, İ. H. (2019). Entegre Lojistik Destek Sisteminin Ömür Devri Yönetimine Etkisi: Hava Araçları Yedek Parça Kullanımı Üzerine Bir Uygulama.

[4] Avacharmal, R. (2022). Enhancing Algorithmic Efficacy: A Comprehensive Exploration of Machine Learning Model Lifecycle Management from Inception to Operationalization. Distributed Learning and Broad Applications in Scientific Research

MAKİNE ÖĞRENMESİ REGRESYON MODELLERİ KULLANILARAK ALTIN ONS FİYATLARININ TAHMİNİ

Halil Sezgin Can*
Milli Savunma Üniversitesi, Kara Astsubay Meslek Yüksek Okulu Müdürlüğü, Temel Bilimler Bölümü, Balıkesir, Türkiye

Fırat Ateş
Balıkesir Üniversitesi, Matematik Bölümü, Balıkesir, Türkiye
(İletişimden sorumlu yazar) halilsezgincan@gmail.com

ÖZET

Altın, yatırımcılar ve finans piyasaları açısından güvenilir bir yatırım aracı olarak tarih boyunca önemli olmuştur. Bu çalışma kapsamında, 2000-2025 yılları arasındaki günlük altın ons fiyatları temel alınarak, gelecekteki fiyat hareketlerinin tahmini üzerine makine öğrenmesi regresyon modelleriyle analiz gerçekleştirilmiştir. Çalışmada kullanılan veriler https://tr.investing.com/commodities/gold adresinden elde edilmiş ve Python programlama diliyle çeşitli regresyon modelleri uygulanarak analiz edilmiştir. Lineer Regresyon, Polinom Regresyon Ridge Regresyon, Lasso Regresyon, ElasticNet Regresyon Destek Vektör Regresyonu (SVR) gibi farklı regresyon modelleri kullanılarak model performansları karşılaştırılmış ve model doğruluğunu ölçmek amacıyla çeşitli hata metrikleri (MAE, RMSE ve R^2) hesaplanmıştır. Çalışmada ayrıca altının günlük en yüksek ve en düşük fiyatları tek bir veri seti olarak değerlendirilmiştir. Bu kapsamda, en yüksek ve en düşük değerlerin aritmetik ortalaması, geometrik ortalaması ve logaritmik ortalaması hesaplanarak, bu farklı ortalama yöntemlerinin tahmin performanslarına olan etkisi de karşılaştırmalı olarak incelenmiştir. Bu yöntemlerle elde edilen farklı ortalama değerlerinin regresyon modellerindeki tahmin performansına etkileri analiz edilerek, hangi ortalama türünün daha iyi sonuç verdiği belirlenmeye çalışılmıştır. Çalışmanın temel amacı, belirlenen zaman aralığındaki altın fiyatlarını doğru bir şekilde tahmin ederek yatırım stratejilerine ve finansal kararlara fikir sağlamaktır. Modellerin geleceğe yönelik tahmin performansları incelenmiş, sonuçlar görselleştirilmiş ve en başarılı tahmin performansına sahip model belirlenmiştir. Elde edilen sonuçlar, altın piyasasının trendlerinin anlaşılması ve gelecekteki yatırım kararlarına ışık tutabilecek niteliktedir. Bu bağlamda, makine öğrenmesi yöntemlerinin finans piyasalarında tahmin aracı olarak kullanılabileceği değerlendirilmiştir.

Anahtar Kelimeler: Makine öğrenmesi, makine öğrenmesi regresyon modelleri, altın ons fiyatları, tahminleme, finansal analiz.

REFERANSLAR

[1] Beyaztaş, M. & Sönmez, M. (2023). Finans Alanında Makine ve Derin Öğrenmenin Kullanılması: Lisansüstü Tezlerde Sistematik Literatür Taraması. *YBS Dergisi – Yönetim Bilişim Sistemleri Dergisi*

[2]Ermiş, B. (2021). Makine Öğrenmesi Teknikleri ile Hisse Senedi Fiyat Tahmini. *Uluslararası Yönetim Akademisi Dergisi*

[3] Kara, S., & Şahin, M. (2020). Makine Öğrenmesi Yaklaşımları ile Finansal Tahminleme: Borsa İstanbul Uygulaması. *İktisat ve Finans Araştırmaları Dergisi*

MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE KÖTÜ AMAÇLI YAZILIM TESPİT ANALİZİ

Erol Özçekiç*
Balıkesir Üniversitesi, Balıkesir Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, Balıkesir, Türkiye

Ümit Yılmaz
Balıkesir Üniversitesi, Bigadiç Meslek Yüksekokulu, Yönetim ve Organizasyon Bölümü, Balıkesir, Türkiye

Ümit Mamuk
Balıkesir Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, Bilgisayar Programcılığı Programı, Balıkesir, Türkiye
**(İletişimden sorumlu yazar) erolozcekic@balikesir.edu.tr*

ÖZET

Kötü amaçlı yazılımlar, dijital dünyada giderek daha karmaşık hale gelerek bireyler ve kurumlar için ciddi bir tehdit oluşturmaktadır. Bu durum, kötü amaçlı yazılımların tespiti için etkili ve yenilikçi yöntemlere duyulan ihtiyacı artırmaktadır. Çalışmada, CTU-13 Malware Dataset kullanılarak kötü amaçlı yazılımların tespiti ve analizi gerçekleştirilmiştir. İlk aşamada, veri ön işleme sürecinde eksik ve hatalı veriler düzeltilmiş, aykırı değerler temizlenmiş ve analiz için uygun hale getirilmiştir. Özellik mühendisliği aşamasında, veri setindeki trafik verileri incelenmiş ve akış süresi (Flow Duration) ile ortalama paket boyutu (Avg Packet Size) gibi önemli özellikler seçilmiştir. Bu özelliklerin seçiminde, kötü amaçlı aktivitelerin belirgin davranışlarını yansıtmaları ve analizlerde yüksek ayırt edici güce sahip olmaları etkili olmuştur.

Ardından, K-means algoritması kullanılarak yapılan kümeleme analizi, trafik kayıtlarını üç kümeye ayırmış ve ağ trafiğindeki kötü amaçlı davranışların daha iyi anlaşılmasına olanak tanımıştır. Bu işlem, sınıflandırma sürecine daha iyi temsil edilen veriler sağlamış ve kötü amaçlı aktivitelerin zararsız aktivitelerden ayrıştırılması için güçlü bir temel oluşturmuştur.

Sınıflandırma aşamasında, Naïve Bayes, Decision Tree, Random Forest, KNN, Gradient Boosting, AdaBoost, XGBoost ve LightGBM gibi makine öğrenmesi algoritmaları değerlendirilmiştir. Bu algoritmalar, kötü amaçlı aktivitelerin zararsız aktivitelerden ayrıştırılması için doğruluk, hassasiyet, geri çağırma ve F1 skoru gibi metriklerle test edilmiştir. Analiz sonuçları, Gradient Boosting ve XGBoost algoritmalarının %99.09 doğruluk oranıyla diğer yöntemlere kıyasla en yüksek performansı sergilediğini ortaya koymuştur. Performans değerlendirmesi, bu iki algoritmanın yalnızca yüksek doğruluk oranı sağlamakla kalmadığını, aynı zamanda hassasiyet ve geri çağırma metriklerinde de güçlü sonuçlar verdiğini göstermektedir.

Anahtar Kelimeler: Kötü amaçlı yazılım, makine öğrenmesi, özellik mühendisliği.

MEDMIND: A DATA-GUIDED DIGITAL MEDICATION MANAGEMENT SOLUTION FOR INDIVIDUALS WITH COGNITIVE AND SENSORY IMPAIRMENTS

Unzile Yaman*
İzmir Kâtip Çelebi University, Faculty of Pharmacy, Department of Pharmaceutical Toxicology, İzmir, Türkiye

Muhammet Can Sekman
İzmir Kâtip Çelebi University, Faculty of Pharmacy, İzmir, Türkiye
(Corresponding author) unzile.yaman@ikcu.edu.tr, unzileyaman@gmail.com

ABSTRACT

Medication safety and adherence are particularly challenging for individuals with cognitive, visual, or auditory impairments, who face significant barriers in managing complex pharmacotherapy. MedMind is a digital medication management solution designed to address these barriers through a multi-layered, accessibility-focused approach. The application delivers personalized reminders via audio, visual, and vibration notifications, ensuring compatibility with users' sensory capacities. In addition to dose and timing reminders, the system tracks prescription and medical report renewal dates, reducing disruptions in chronic treatment regimens. A built-in module analyzes potential drug–drug and drug–food interactions and provides alerts using a simplified, user-friendly format tailored to individual patient profiles. Caregivers and healthcare professionals are granted access, with permission, to monitor patient compliance and receive real-time updates in emergency situations. A poisoning response interface guides users or caregivers through first-response protocols in the event of accidental overdose or adverse reactions, with direct links to emergency services or poison information centers.

MedMind was awarded first place in a national health innovation competition and is currently under further evaluation for large-scale implementation. While specific technical components and algorithmic structures are not disclosed in this presentation due to ongoing development, we welcome collaborations and critical feedback from the AI and data science community. The project exemplifies how data-guided, patient-centered design can bridge accessibility gaps in digital health, enhance medication safety, and empower both patients and caregivers within vulnerable populations.

Keywords: Digital health, medication safety, accessibility, personalized alerts, cognitive impairment.

OPERATOR-THEORETIC AND QUANTUM-INSPIRED FRAMEWORKS FOR DATA-DRIVEN MODELING OF NONLINEAR DYNAMICAL SYSTEMS

Joanna Slawinska*
Department of Mathematics, Dartmouth College, Hanover, NH, USA
Dimitris Giannakis
Department of Mathematics, Dartmouth College, Hanover, NH, USA
(Corresponding author) joanna.m.slawinska@dartmouth.edu

ABSTRACT

We present a data-driven operator-theoretic framework for the analysis of nonlinear dynamical systems. Rooted in Koopman operator theory and enriched with kernel-based learning techniques, the method constructs finite-dimensional approximations of infinite-dimensional evolution operators, enabling nonlinear dynamics to be studied through linear evolution in a learned observable space.

The framework supports spectral decomposition for interpretable, compact representations of spatiotemporal dynamics. Variations based on vector-valued observables allow modeling of coupled subsystems and multivariate fields. By embedding nonlinear evolution into linear operator flows, the method captures dominant patterns without requiring intrusive solvers or governing equations.

Among several Koopman-based predictive schemes, a recent advance introduces a quantum-inspired formulation in which the evolving state is represented by a data-driven analog of a density operator, and observables evolve under a Koopman operator, mirroring Heisenberg dynamics. This approach enables uncertainty quantification and underpins a nonparametric data assimilation scheme. Operators are constructed empirically using time-ordered data in a basis learned via delay-coordinate embeddings and kernels.

The framework is well suited for partially observed or complex systems where first-principles models are unavailable. It offers a flexible, interpretable alternative to black-box learning, while maintaining ties to functional and spectral analysis. It also provides a pathway for developing subgrid-scale modeling strategies, including closure schemes and parameterizations, by capturing unresolved dynamics through coarse-grained evolution.

Current research focuses on implementing this framework on quantum-like platforms and quantum computers, aligning operator structure and spectral evolution with quantum computational primitives for scalable, data-driven modeling.

Keywords: Nonlinear dynamical systems.

OTURUM KONFOR TASARIMININ OBJEKTİF OLARAK İNCELENMESİ

Süleyman Dündar
Yataş Grup, Erciyes Üniversitesi, Fen Bilimleri Enstitüsü, Endüstriyel Tasarım Mühendisliği Anabilim Dalı, Kayseri, Türkiye

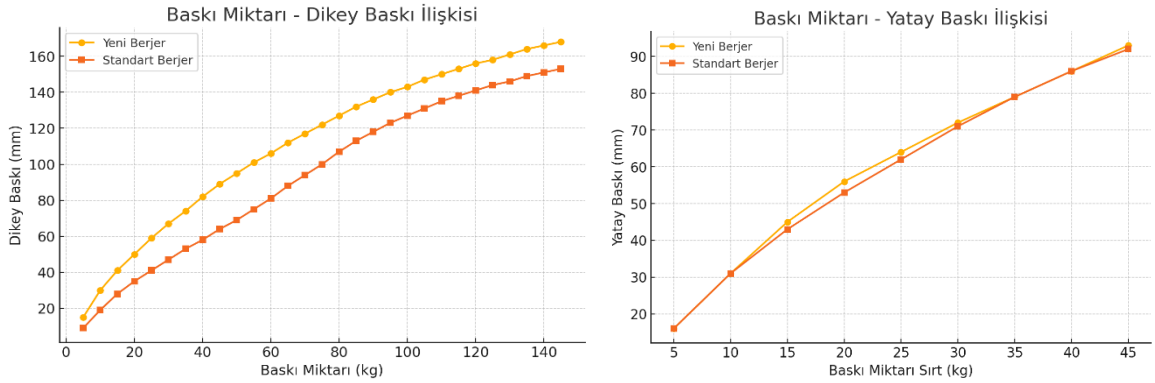
Esra Akgül*
Erciyes Üniversitesi, Mühendislik Fakültesi, Endüstriyel Tasarım Mühendisliği Bölümü, Kayseri Türkiye
(İletişimden sorumlu yazar) akgul@erciyes.edu.tr

ÖZET

Oturma grubu mobilyaları günlük hayatta en sık kullandığımız ürünler arasında yer almaktadır. Oturma grubu tasarımı yapılırken dikkate alınan başlıca parametreler; tarz, konfor, sağlamlık ve maliyettir. Mobilya sektöründe kullanılan ürünler konfor ve ergonomik açıdan her ne kadar kabul edilebilir olsa da, mevcut konfor yapılarının yetersiz kaldığı durumlar gözlemlenmektedir. Kullanıcıların kısa ve uzun süreli oturum pozisyonunda kalma durumlarında hissettikleri rahatlık algıları değişkenlik gösterebilmektedir. Bu durum, oturma gruplarında konfor algısının insan fizyolojisi ve psikolojisi üzerindeki önemini daha da artırmaktadır. Literatürde konfor mobilya ile ilişkili birçok araştırma yapılmıştır. Yan Liu ve arkadaşları, çalışmalarında, ergonomik mobilya tasarımında biyomekaniğin farklı mobilya tasarımlarında nasıl uygulandığı, biyomekanik analiz ve test yöntemleri üzerine araştırma yaparak, biyomekanik yöntemlerin mevcut durumunu analiz etmişlerdir [1]. Yong Zhang ve arkadaşları, ofis ortamında oturma konforunu değerlendirmek için çok seviyeli bir oturma pozisyonu değerlendirme sistemi oluşturmaktır. Bu sistem, eklem açıları, eklem torkları, eklem kuvvetleri ve kas kuvvetleri gibi parametreleri analiz ederek geliştirilmiştir [2]. Oturma ergonomisi kullanıcının antropometrisine en uygun tepkisel davranışı sergileyen malzeme veya malzemeler bütünü ile elde edilmiş yapı olarak tanımlanmaktadır. Kullanıcı ve konfor arasındaki etkileşimin kaynağı olarak, oturum konfor yapısı içerisinde kullanılan malzemelerin birbirleri ile etkileşimi de toplam konfor algısı bakımından bir o kadar önemlidir. Çalışmada, konfor katmanları arasında yeni nesil konfor destek malzemeleri ekleyerek, konfor algısını yükseltip daha yüksek ergonomik oturum konforunu sağlamak amaçlanmıştır.

Çalışma kapsamında, konfor düzeyinin artırılmasına yönelik tasarlanan oturma grubu ünitesi ve geleneksel olarak tasarlanan belirli konfor düzeyindeki oturma grubu ünitesi; basınç-yük- tepki temeline dayalı geliştirilen ve oturma gruplarında konfor testini objektif olarak değerlendirerek konforu nicel verilerle ifade eden objektif ölçüm yöntemi kullanılarak, ürünlerin konfor düzeyleri incelenmiştir [3].

Elde edilen sonuçlar, Şekil 1'de gösterilmiştir. Şekil 1'de verilen dikey baskı grafiği, oturum katmanına uygulanan (kg) yük karşısında, yeni ve standart oturum katmanının verdiği baskı ile itme ilişkisini göstermektedir.



Şekil 1: Konfor test cihazından elde edilen sonuçların karşılaştırılması

Standart oturum ile yeni oturum katmanı kıyaslanmasında çıkan sonuçların belirli baskı yükleri altında, standart oturum katmanına göre farklı baskı ve itme eğrisi çizdiği gözlemlenmektedir. Yeni konfor tasarımı yapılarak geliştirilen oturma grubu ünitesi, farklı yük/basınç altında objektif konfor test cihazına göre değerlendirildiğinde, geleneksel olarak tasarlanan oturma grubu ünitesinden daha yüksek konfor algısı oluşturabileceği düşünülmektedir.

Anahtar Kelimeler: Ergonomi, konfor, mobilya, ürün tasarımı.

REFERANSLAR

[1] Liu Y, Hu W, Kasal A, Erdil YZ., "Ergonomik Mobilya Tasarımında Uygulanan Biyomekaniğin Son Durumu", *Uygulamalı Bilimler*, 13, 22, 12120, 2023.

[2] Zhang, Y., "*Seating comfort in office environments: A study*", *Ergonomics Journal*, 45, 3, 123-135, 2023.

[3] Borlu, E., Atıcı, Ö., Demir, R., Durmuş, M., Kaya, Ü., & Ekinci, E. "Method and Device for Determining Comfort in Seating Groups and Comparison of Subjective and Objective Evaluations", The European Journal of Research and Development, *2*, 4, 57-66, 2022.

ÖDEV DESTEK SİSTEMİ OLARAK YAPAY ZEKÂ KULLANIMI

Hasan Çoruk*

Zonguldak Bulent Ecevit Üniversitesi, Karadeniz Ereğli Meslek Yüksekokulu, Zonguldak, Türkiye

Turgay Erdemir

Amasya Üniversitesi, Merzifon Meslek Yüksekokulu, Amasya, Türkiye

**(İletişimden sorumlu yazar) hasancoruk@beun.edu.tr*

ÖZET

Bu çalışma, yapay zekâ araçlarının öğrencilerin ev ödevlerini tamamlama süreçlerindeki rolünü incelemeyi amaçlamaktadır. Günümüzde yapay zekâ, eğitim ortamlarında giderek daha fazla kullanılmakta ve öğrenme süreçlerini destekleyerek öğrencilere kişiselleştirilmiş öğrenme deneyimleri sunmaktadır [3]. Araştırma, yapay zekânın kişiselleştirilmiş öğrenme desteği, adım adım rehberlik ve zaman yönetimi avantajlarını analiz ederken; akademik dürüstlük, özgünlük kaybı ve kültürel duyarlılık eksikliği gibi potansiyel riskleri de tartışmaktadır [1].

Araştırma nitel araştırma desenlerinden durum çalışması yöntemi ile yürütülmüştür [2]. Araştırmanın çalışma grubunu bir devlet üniversitesinde çeşitli bölümlerde okumakta olan öğrenciler oluşturmaktadır. Araştırma 2024-2025 bahar döneminde gerçekleştirilmiştir. Araştırma süreci dört haftadır. Her hafta derslerde anlatılan konulara yönelik ödevlendirmeler yapılmıştır. Öğrenenlerin ödevlerini yaparken ihtiyaç duymaları halinde yalnızca yapay zekâ araçlarından faydalanmaları istenmiştir. Tamamlanan ödevler bir çevrimiçi öğrenme yönetim sistemi aracılığıyla öğretim elemanına gönderilmiştir. Verilen ödev, sonraki hafta işlenmekte olan dersin ilk bölümünde öğretim elemanı ile tekrar çözülmüş, öğrencilerden kendi çalışmalarını değerlendirmeleri için zaman tanınmış ve bir sonraki konuya geçilmiştir. Takip eden 4 haftanın sonunda araştırma süreci tamamlanmıştır.

Veri toplama araçları olarak öğrencilere uygulanan süreç değerlendirme formu uygulaması kullanılmıştır. Bulgular, yapay zekâ uygulamalarının öğrencilere ödevlerinde yardımcı olması amacıyla kullanılabileceğini göstermektedir. Öğrencilere geri bildirim sağlayarak öğrenme verimliliğini artırdığı tespit edilmiştir [4]. Ancak aşırı derecede yapay zekâ uygulaması kullanımının eleştirel düşünme becerilerini zayıflatabileceğini göstermiştir. Sonuç olarak, yapay zekâ araçlarının eğitimde tamamlayıcı bir rehber olarak kullanılması önerilmektedir.

Anahtar Kelimeler: Yapay zekâ, eğitim teknolojileri, ChatGPT, DeepSeek, ödev desteği.

REFERANSLAR

[1] Selwyn, N. (2019). Should robots replace teachers? AI and the Future of Education. (1st ed.) Polity Press.

[2] Yıldırım, A., & Şimşek, H. (2023). Sosyal bilimlerde nitel araştırma yöntemleri (12. baskı). Seçkin Yayıncılık.

[3] Yörük, T. (2024). Eğitimde yapay zekâ ve kişiselleştirilmiş öğrenme. Eğitim Bilimleri Alanında Uluslararası Araştırmalar XXIII, 21.

[4] Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. International Journal of Educational Technology in Higher Education, 16(1), 1-27.

ÖĞRETMEN ADAYLARININ BÜYÜK VERİNİN EĞİTİMDE KULLANIMINA İLİŞKİN FARKINDALIKLARININ İNCELENMESİ

Ayşegül Ayyıldız Asil*
Hakkari Üniversitesi, Eğitim Fakültesi, Hakkari, Türkiye

Abdurrahman Asil
Selahaddin Eyyubi Ortaokulu, Hakkari, Türkiye
(İletişimden sorumlu yazar) aysegulkadi@hakkari.edu.tr

ÖZET

Araştırmanın amacı, öğretmen adaylarının eğitimde büyük verinin kullanımına ilişkin görüşlerini incelemektir.

Araştırmada nitel yöntem tercih edilmiştir. Bir problemin çözümüne ilişkin gözlem, görüşme ve doküman analizi gibi nitel veri toplama yöntemlerini kullanan nitel araştırma, daha önceden bilinen veya fark edilmemiş problemlerin algılanmasına, probleme ilişkin doğal olguların gerçekçi bir şekilde ele alınmasına yönelik öznel-yorumlayıcı bir süreci ifade etmektedir [1].

Araştırmada nitel araştırma yöntemlerinden olgu bilim yöntemine başvuruluştur. Olgu bilim araştırmalarında olgulara ilişkin yaşantıları ve bunlara yüklenen anlamları ortaya çıkarmak için araştırmanın odaklandığı olguyu yaşayan veya bu olguyu dışa vurabilecek bireylerle çalışılır [2].

Araştırmada kolay ulaşılabilir örnekleme yöntemi tercih edilmiştir. Yıldırım ve Şimşek'e göre (2008) "kolay ulaşılabilir durum örneklemesi" olarak da adlandırılır [3]. Bu örnekleme türü, araştırma yapılacak birey ya da grupların araştırma sürecine dâhil edilmesinin daha kolay ya da bunlara erişilmesinin daha kolay olmasıyla ilişkilidir. Çalışma Hakkâri Üniversitesi öğretmen adayları ile yürütülecektir. 20 öğretmen adayı araştırmaya dahil edilecektir.

Araştırmada "Büyük Veriye İlişkin Yarı Yapılandırılmış Görüşme Formu" kullanılmıştır.

Verilerin analizinde içerik analizi kullanılacaktır. İçerik analizi, sosyal gerçekliği araştıran nesnel, sistematik, tümdengelime dayalı okuma aracı olarak önceden belirlenen ölçütlere göre kavramlardan, metinlerden, sözlü veya yazılı materyallerden anlamlar çıkarmayı amaçlayan metodolojik araç ve teknikler bütünüdür [4].

Bu araştırmanın bulgularına göre, öğretmen adayları büyük veriyi en çok "bilgi havuzu" ve olarak tanımlamış, eğitimsel veri madenciliğini ise eğitim verilerini analiz etme süreci olarak görmüşlerdir. Eğitimde büyük veriye örnek olarak sınav notları verilmiştir. Büyük verinin eğitimdeki yararları arasında en çok öğrenci performansının analiz edilmesi şeklinde vurgulanmıştır. Buna karşın büyük veri kullanımında, veri güvenliği, gizlilik sorunları ve bilgi

kirliliği gibi zorluklara dikkat çekilmiştir. Büyük verinin eğitime etkileri arasında bilgiye hızlı erişim olumlu yön ön plana çıkarken; öneriler kısmında ise paydaşlara büyük veri eğitimi, veri analizi programları geliştirme ve düzenleyici kılavuzlar oluşturma gerektiği ifade edilmiştir.

Anahtar Kelimeler: Büyük veri, görüşme, öğretmen adayı.

REFERANSLAR

[1] Seale, C. (1999). Quality in qualitative research. Qualitative Inquiry, 5(4), 465-478.

[2] Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö.E., Karadeniz, Ş. ve Demirel, F. (2008). *Bilimsel araştırma yöntemleri.* Ankara: Pegem Akademi.

[3] Yıldırım, A., ve Şimşek, H. (2008). Sosyal Bilimlerde Nitel Araştırma Yöntemleri . Ankara: Seçkin.

[4] Tavşancıl, E. ve Aslan, E. A. (2001). Sözel, yazılı ve diğer materyaller için içerik analizi ve uygulama örnekleri. İstanbul: Epsilon.

PERCEPTION OF OLYMPIC OPENING CEREMONIES THROUGH SOCIAL MEDIA COMMENTARY: A SENTIMENT ANALYSIS AND TOPIC MODELING APPROACH

Hüseyin Koçak*

Ordu University, Finance, Banking, and Insurance, Ordu, Türkiye

**(Corresponding author) huseyinkocak@odu.edu.tr*

ABSTRACT

This study aims to examine how the opening ceremony of the 2024 Olympic Games is perceived by the public through social media comments. Accordingly, 6461 comments made on the opening ceremony video, which was published on the official Olympic channel on YouTube and reached 2.1 million views as of February 22, 2025, were extracted from English to make the data suitable for analysis. The obtained text data contains a total of 11,159 words. Pre-processing steps were applied to the data using the Python programming language, and repetitive words and unnecessary elements in the English language were cleaned. Then, sentiment analysis was performed using the VADER, NRC Sentiment Dictionary and Transformers models, and emotional tendencies in the text were determined. In addition, In the Topic Modeling phase, the coherence score metric was used to determine the optimal number of topics, and the 9 topics providing the highest coherence score were selected as the ideal number of topics. The topic modeling process was performed using the Non-Negative Matrix Factorization (NMF) method. The themes that emerged can be summarized as follows: Opening Ceremony and Paris Theme (ceremony, opening, paris, french, france), Tokyo 2020 and Performances (magnificent, tokyo, breaking, pop, masterpiece), Social Media and Comments (call, orientations, youtube, unique, use), Positive Emotions and Impressive Elements (incredible, positive, crowd, playing, honestly), Stadium and Organization (stadium, forever, masked, fact, incredible), Decision Process and Planning (uploading, beginning, edition, decision, unifying), Movements and Cultural Elements (movements, ancient, provides, wonderful), Youth and Sports Events (boy, speculation, sing, athletic, disruptive), Criticism and Discussions (insult, comment, troubles, despicable, amaze). The findings reveal that the Olympic opening ceremonies have a wide resonance in social media and that different emotional reactions and thematic focuses coexist in the public.

Keywords: Social media, sentiment analysis, topic modelling.

SANAL GERÇEKLİK TEKNOLOJİSİ İLE SİMÜLASYON ORTAMINDA DEPREM HAZIRLIK EĞİTİMİ

Nur Kuban Torun*
Bilecik Şeyh Edebali Üniversitesi, Yönetim Bilişim Sistemleri, Bilecik, Türkiye
**(İletişimden sorumlu yazar) nurkuban.akdemir@bilecik.edu.tr*

ÖZET

Dijital dönüşüm teknolojilerinin hayatımıza girişiyle birlikte, sanal gerçeklik (VR) teknolojisi hem eğlenceli hem de eğitici yönleriyle dikkat çekmektedir. Sanal gerçeklik, kullanıcıların, görme, duyma ve bazen de dokunma duyularını kullanarak gerçek dünyadaki hisleri yaşadıkları sanal bir deneyim sunar. Kullanılan duyuların sayısı arttıkça, sanal dünya ile gerçek dünya arasındaki fark azalır. Bu denli gerçeklik yaratabilen sanal dünyalar birçok alanda kullanılmaya başlanmıştır. Türkiye, aktif fay hatlarının bulunduğu bir bölgede yer aldığı için depremler, toplum üzerinde önemli etkiler yaratır. Bu nedenle, afet yönetimi kurumları hazırlıklar yapmakta ve eğitimler vermektedir. Ancak, bu eğitimler genellikle yüz yüze ve sınırlı sayıda katılımcı ile gerçekleşmektedir. Bu sorunu çözmek amacıyla, Kurtuluş Deprem Simülasyonu adlı bir oyun geliştirilmiştir. Bu oyun, sanal gerçeklik gözlüğü kullanarak bir simülasyon ortamı oluşturur ve deprem öncesi, deprem anı ve deprem sonrası neler yapılması gerektiğini deneyimleyerek kalıcı bir öğrenme sağlamayı amaçlar. Unity oyun motoru kullanılarak C# programlama dili ile geliştirilen bu oyun, taşınabilir ve ulaşılabilir bir ürün olarak öne çıkmaktadır. Bu şekilde, kullanıcılar deprem olmadan hazırlık ve tedbir yöntemlerini öğrenebilir ve deprem anında ve sonrasında neler yapmaları gerektiğini deneyimleyerek daha bilinçli hale gelebilirler.

Şekil 1: Kurtuluş Deprem Simülasyonu Sahne Örneği

Anahtar Kelimeler: Sanal gerçeklik, eğitim, deprem.

SAVUNMA SANAYİSİNDE MAKİNE ÖĞRENMESİ İLE DESTEKLENEN ÖNLEYİCİ BAKIM PLANLAMASI

Pınar Özel*

Balıkesir Üniversitesi, Fen Bilimleri Enstitüsü, Matematik Bölümü, Balıkesir, Türkiye

Sebahattin İkikardeş

Balıkesir Üniversitesi, Matematik Bölümü, Balıkesir, Türkiye

**(İletişimden sorumlu yazar) piiinarozel@gmail.com*

ÖZET

Günümüzde geleneksel bakım yöntemleri, araçların belirli zaman aralıklarında servise alınmasını ve bazı parçaların periyodik olarak değiştirilmesini esas almaktadır. Ancak, her bir aracın kullanım şekli ve koşulları farklılık gösterdiğinden bu yöntemler, bakım süreçlerinde hem maliyet hem de verimlilik açısından yetersiz kalabilmektedir. Bazı bileşenlerin gereksiz yere değiştirilmesi veya bakımın gecikmesi nedeniyle oluşabilecek beklenmedik arızalar, sistem güvenliğini ve operasyonel sürekliliği olumsuz etkilemektedir. Bu bağlamda, önleyici bakım tahmini; araçlardan elde edilen gerçek zamanlı veriler aracılığıyla, bakım ihtiyacını doğru şekilde belirlemeyi amaçlayan yenilikçi bir yaklaşımdır. Bu teknoloji, yalnızca maliyetleri azaltmakla kalmayıp aynı zamanda araçların kullanım ömrünü uzatmakta, operasyonel verimliliği artırmakta ve güvenlik seviyesini yükseltmektedir. Geliştirilen sistem, özellikle savunma sanayi başta olmak üzere birçok sektörde bireysel ve ticari kullanıcılara yönelik önemli avantajlar sunmaktadır. Model ve sistem, test edilip geri bildirimlere dayanarak iyileştirirerek sonuçlar doğrultusunda, sistem savunma sanayiinde kullanılabilir hale getirebilmektedir. Ayrıca, modelin hata payı ve riskleri izlenerek sürekli iyileştirme için önlemler alabilmektedir. Makine öğrenmesi ile desteklenen bir bakım planlaması amaçlanmıştır. Projede araç bakımı ile ilgili bazı verilere ulaşılmıştır. Verilerde akü durumu, fren durumu, lastik durumu, geçmiş bakım sayısı gibi özellikler yer almaktadır. Eksik veriler analiz edilmiştir, veri temizliği ve görselleştirme yapılmıştır. 9 farklı regresyon modeli (örneğin XGBoost, Lineer Regresyon, Rastgele Orman) test edilmiştir. Modeller hata metriklerine göre karşılaştırılmıştır ve çapraz doğrulama uygulanmıştır.

Anahtar Kelimeler: Savunma sanayisi, araç bakım veri kümesi, önleyici bakım planlaması, makine öğrenmesi.

Savunma Sanayii Başkanlığı, SAYZEK, Akademik Tez Programı kapsamında bu çalışmalar yapılmaktadır.

SİBER GÜVENLİK KONUSUNDA ÖĞRETMEN ADAYLARININ FARKINDALIKLARININ İNCELENMESİ

Ayşegül Ayyıldız Asil*
Hakkari Üniversitesi, Eğitim Fakültesi, Hakkari, Türkiye

Abdurrahman Asil
Selahaddin Eyyubi Ortaokulu, Hakkari, Türkiye
(İletişimden sorumlu yazar) aysegulkadi@hakkari.edu.tr

ÖZET

Araştırmanın amacı, öğretmen adaylarının siber güvenliğe ilişkin görüşlerini incelemektir.

Araştırmada nitel yöntem tercih edilmiştir. Bir problemin çözümüne ilişkin gözlem, görüşme ve doküman analizi gibi nitel veri toplama yöntemlerini kullanan nitel araştırma, daha önceden bilinen veya fark edilmemiş problemlerin algılanmasına, probleme ilişkin doğal olguların gerçekçi bir şekilde ele alınmasına yönelik öznel-yorumlayıcı bir süreci ifade etmektedir [1].

Araştırmada nitel araştırma yöntemlerinden olgu bilim yöntemine başvurulmuştur. Olgu bilim araştırmalarında olgulara ilişkin yaşantıları ve bunlara yüklenen anlamları ortaya çıkarmak için araştırmanın odaklandığı olguyu yaşayan veya bu olguyu dışa vurabilecek bireylerle çalışılır [2].

Araştırmada kolay ulaşılabilir örnekleme yöntemi tercih edilmiştir. Yıldırım ve Şimşek'e göre (2008) "kolay ulaşılabilir durum örneklemesi" olarak da adlandırılır [3]. Bu örnekleme türü, araştırma yapılacak birey ya da grupların araştırma sürecine dâhil edilmesinin daha kolay ya da bunlara erişilmesinin daha kolay olmasıyla ilişkilidir. Çalışma Hakkâri Üniversitesi öğretmen adayları ile yürütülecektir. 15 öğretmen adayı araştırmaya dahil edilecektir.

Araştırmada "Siber Güvenliğe İlişkin Yarı Yapılandırılmış Görüşme Formu" kullanılmıştır.

Verilerin analizinde içerik analizi kullanılacaktır. İçerik analizi, sosyal gerçekliği araştıran nesnel, sistematik, tümdengelime dayalı okuma aracı olarak önceden belirlenen ölçütlere göre kavramlardan, metinlerden, sözlü veya yazılı materyallerden anlamlar çıkarmayı amaçlayan metodolojik araç ve teknikler bütünüdür [4].

Öğretmen adaylarının siber güvenliğe ilişkin görüşlerine dair yapılan nitel çalışmada, katılımcılar siber güvenliği en çok dijitalde koruma sistemi olarak tanımlamışlardır. Siber güvenlik ihlalleri arasında en çok kişisel veri suistimali ve kimlik avı ön plana çıkarken, ihlallerin en önemli nedeni olarak zayıf şifre kullanımı gösterilmiştir. İhlallerin kurumsal etkileri arasında maddi kayıplar ve kurumsal güvenin zedelenmesi sıkça dile getirilmiştir. Güçlü şifre kullanımı, antivirüs yazılımları ve yazılım güncellemeleri gibi önlemler önerilmiş; yöneticiler

ve çalışanlar içinse en çok eğitimlere katılımın önemli olduğu vurgulanmıştır. Siber güvenlik çalışmalarının kuruma katkısı olarak kurumsal güven ve itibar artışı öne çıkmıştır.

Anahtar Kelimeler: Siber güvenlik, nitel, öğretmen adayı.

REFERANSLAR

[1] Seale, C. (1999). Quality in qualitative research. Qualitative Inquiry, 5(4), 465-478.

[2] Büyüköztürk, Ş., Kılıç Çakmak, E., Akgün, Ö.E., Karadeniz, Ş. ve Demirel, F. (2008). *Bilimsel araştırma yöntemleri.* Ankara: Pegem Akademi.

[3] Yıldırım, A., ve Şimşek, H. (2008). Sosyal Bilimlerde Nitel Araştırma Yöntemleri. Ankara: Seçkin.

[4] Tavşancıl, E. ve Aslan, E. A. (2001). Sözel, yazılı ve diğer materyaller için içerik analizi ve uygulama örnekleri. İstanbul: Epilon.

SİBER SAVAŞTA YAPAY ZEKÂ: ASKERİ YAPAY ZEKÂ UYGULAMALARINDA EVRENSEL ETİK VE HUKUKUN DİJİTAL ÇATIŞMADAKİ ROLÜ

Erdoğan Kahraman*
Millî Savunma Bakanlığı, KKK, İzmir, Türkiye
**(İletişimden sorumlu yazar) erdogankahraman@yahoo.com*

ÖZET

Yapay zekâ (YZ) ve makine öğrenimi (MÖ), savaşın ve siber savaşın yeni aktörleri haline gelirken, insanlığın etik sınırlarını zorlamaktadır. Teknolojinin bu alanlarda "savaş odaklı" dönüşümü Otonom silahların "akıllı" hedef seçimi veya yapay zekâyla yönetilen siber saldırıların hastane, enerji şebekesi gibi hayati sistemleri hedef alması, teknolojinin karanlık olarak nitelendirilebilecek bir yönünü ortaya çıkarabilmektedir. İnsanlık, geliştirdiği bu gücü kontrol edebilecek ahlaki olgunluğa sahip midir? Yoksa teknoloji, hukukun ve vicdanın önüne geçerek yeni bir çağın tehlikeli oyuncusu mu olacaktır?

Uluslararası İnsan Hakları Hukuku (İHİH) ve Cenevre Sözleşmeleri, çatışmalarda sivil kayıpların minimize edilmesini, orantılılık ilkesini ve "insani vicdan"ı merkeze almaktadır. Ancak YZ tabanlı otonom silahlar, hedef seçimini algoritmalara devrederek bu ilkeleri aşındırmaktadır. Örneğin, 2021'de Libya iç savaşında insansız hava araçlarının (İHA'lar) otonom modda insan hedeflere saldırdığı iddiaları, BM raporlarında kayıtlara geçmiştir. Bu olay, İHL'nin 36. Protokolü'nün ihlali anlamına gelmektedir: Bir silahın kullanımı, hukuka uygunluk sağlayamıyorsa yasaklanmalıdır. Benzer şekilde, Avrupa İnsan Hakları Mahkemesi (AİHM)'nin, Yaşam Hakkı (Madde 2) ve İşkence Yasağı (Madde 3) ilkeleri, algoritmaların "ölüm kararı" vermesini açıkça reddetmektedir.

YZ'nin savaş alanına entegrasyonu, insanın "ahlaki yargı" yetisini teknolojiye devretme riskini doğurmaktadır. Örneğin, bir MÖ modeli, düşmanın konumunu tahmin ederken, sivil yerleşim alanlarını "yanlışlıkla" hedef alabilir. Bu durumda, sorumluluk kimde olacaktır? Algoritmayı kodlayan mühendiste mi, operasyonu onaylayan komutanda mı? Felsefi etik bağlamında, Kant'ın "Araçsallık İlkesi" (insanların araçsallaştırılmaması) ve Bentham'ın "Faydacılık" teorisi (en çok sayıda insanın iyiliği) burada çatışmaktadır: Otonom sistemler, sivil kayıpları "kabul edilebilir" bir maliyet olarak hesaplarken, insan onurunu metalaştırmaktadır.

Öte yandan, BM İnsan Hakları Yüksek Komiserliği, 2023›te yayınladığı raporda, «ölüm algoritmaları»nın demokratik denetimden uzak geliştirilmesini, insan hakları ihlali olarak nitelendirmiştir. Bu durum, devletlerin «ulusal güvenlik» adı altında şeffaflıktan kaçınmasıyla daha da karmaşıklaşmaktadır.

Günümüzde YZ'nin askeri kullanımını düzenleyen bağlayıcı bir uluslararası sözleşme bulunmamaktadır. AB'nin Yapay Zekâ Yasası, "yüksek riskli" sistemleri sınıflandırarak askeri

YZ'yi kısmen dışarıda bırakmaktadır. Buna karşın, Birleşmiş Milletler bünyesindeki LAWS (Ölümcül Otonom Silah Sistemleri) müzakereleri, 9 yıldır sonuçsuz kalmaktadır. Çin ve ABD gibi devletler, "insan kontrolü"nün gevşek tanımlanması üzerinde ısrar ederken, Avusturya ve Brezilya gibi ülkeler, tamamen yasaklanmasını savunmaktadır.

Siber alanda ise Budapeşte Sözleşmesi gibi mevcut metinler, YZ destekli saldırıların dinamiklerine ayak uyduramamaktadır. Örneğin, derin öğrenme ile üretilen deepfake propagandası, toplumsal kutuplaşmayı derinleştirerek Birleşmiş Milletler Medeni ve Siyasi Haklar Sözleşmesi'nin (BMSHS) 19. ve 20. maddelerini (bilgiye erişim ve nefret söylemi yasağı) ihlal etmektedir.

İnsan-merkezli bir gelecek için yol haritası hazırlanması, teknolojinin insan onuru, hukukun üstünlüğü ve etik hesap verilebilirlik üçgenine entegre edilmesinden geçmektedir.

Sonuç: YZ ve MÖ, savaşın doğasını değiştirirken, insanlığın kolektif vicdanını da sınamaktadır. Teknolojinin "yapabiliriz, çünkü yapabiliyoruz" mantığıyla Teknolojinin, hukukun ve etiğin önünde ilerlemesi yerine, Otonom silahların ve siber savaşın yarattığı tehditler karşısında, devletlerin tek taraflı hamleleri yerine küresel işbirliği şarttır. Evrensel İnsan Hakları Beyannamesi'nin 1. Maddesi ("Tüm insanlar özgür ve onur bakımından eşit doğar") ile uyumlu bir yaklaşım benimsenmelidir. Çünkü bir algoritmanın "yanlış" kararı, yalnızca veri hatası değil, insanlığın ahlaki çöküşünün de göstergesi olabilir.

Anahtar Kelimeler: Askeri yapay zeka, insan hakları, siber savaş.

SMART OPTIMIZATION IN LOGISTICS: A CONCEPTUAL REVIEW

İrem Karakaya*
Bartın University, Department of Management and Organization, Bartın, Türkiye
**(Corresponding author) isahmutoglu@bartin.edu.tr*

ABSTRACT

Machine learning and ensemble learning methods have emerged as powerful tools in critical areas of modern logistics and retail sectors, such as delivery time estimation, production process planning, and enhancing customer satisfaction. Studies have shown that processing data obtained from production and distribution processes and analyzing it through predictive models both increases operational efficiency and accelerates decision-making processes [1].

In delivery time prediction research, algorithms such as support vector machines, artificial neural networks, XGBoost, CatBoost, and gradient boosting are frequently preferred and achieve high accuracy rates. In addition, time series analyses that take historical data into account significantly improve prediction performance. In last-mile deliveries, route optimization and resource-constrained models reduce delivery times and optimize resource utilization [2].

Some studies have experimented with hybrid models that combine machine learning and human expertise, but purely machine learning-based models generally provide higher accuracy. On the other hand, analyses conducted with large datasets yield more accurate predictions and higher customer satisfaction compared to traditional methods [3].

Overall, the studies in this field demonstrate that the correct selection of algorithms, data preprocessing techniques, and appropriate modeling strategies play a critical role in digitizing logistics processes and gaining a competitive advantage.

Keywords: Machine learning, ensemble learning, logistics optimization.

REFERENCES

[1] Sagi, Omer, and Lior Rokach. 2018. "Ensemble Learning: A Survey." Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery 8(4):1–18. doi: 10.1002/widm.1249.

[2] Tsolaki, Kalliopi, Thanasis Vafeiadis, Alexandros Nizamis, Dimosthenis Ioannidis, and Dimitrios Tzovaras. 2023. "Utilizing Machine Learning on Freight Transportation and Logistics Applications: A Review." ICT Express 9(3):284–95.

[3] Zaghloul, Maha, Sherif Barakat, and Amira Rezk. 2024. "Predicting E-Commerce Customer Satisfaction: Traditional Machine Learning vs. Deep Learning Approaches." Journal of Retailing and Consumer Services 79(March):103865. doi: 10.1016/j.jretconser.2024.103865.

SUALTI GÖRÜNTÜLERİNDE DERİN ÖĞRENME İLE NESNE TESPİTİ

Kadriye Filiz Balbal
Dokuz Eylül Üniversitesi, Fen Fakültesi, Bilgisayar Bilimleri, İzmir, Türkiye

Özge Çelik*
Balıkesir Üniversitesi, Fen Bilimleri Enstitüsü, Matematik, Balıkesir, Türkiye

Sebahattin İkikardeş
Balıkesir Üniversitesi, Fen Fakültesi, Matematik, Balıkesir, Türkiye
**(İletişimden sorumlu yazar) ozgephdcelik@gmail.com*

ÖZET

Dünyayı ekonomik ve sosyal açıdan etkisi altına alan küresel ısınma su altı ekosistemini de derinden etkilemektedir. Gerek iklim değişikliği gerekse bilinçsiz avlanma sorunları, üç tarafı denizlerle çevrili olan ülkemizin de sualtı ekosisteminde geri dönüştürülemez değişimlere yol açmaktadır. Deniz suyundaki ısınma ve kirlilik özellikle sualtı canlılarının neslini tehlikeye atmaktadır. Hayatta kalma ve devamlılığını sürdürme içgüdüsüyle birçok tür uygun yaşama koşullarına ulaşabilmek amacıyla daha soğuk sulara gitme eğilimi göstermektedir. Bu da Türkiye deniz ve iç sularındaki istilacı ve tehlikeli deniz canlılarının görülme sıklığını arttırmaktadır. Özellikle ülkemizdeki koruma altındaki türlerin görünme sıklığını büyük ölçüde azaltmıştır. Ekosistemdeki bu dejenerasyon ve değişimin sonuçlarına teknolojik çözümlerle katkı sağlamak amacıyla ülkemiz sualtı canlılarının fotoğraflarından oluşan bir veri kümesi toplanmıştır. Toplanan bu özgün veri kümesi üzerinde, sualtındaki ışık miktarının yetersizliği nedeniyle karşılaşılan bulanıklık ve renk solması gibi etkenlerin elimine edilmesi amacıyla gerekli görüntü ön işleme tekniklerinden geçirilmiştir. Eğitim, doğrulama ve test olmak üzere üç parçaya ayrılan görüntü kümesi üzerinde gerekli sınıf etiketlemeleri yapılmıştır. Ön işleme sonrasında veri kümesi uygun derin öğrenme mimarileri kullanılarak nesne tespitinde kullanılmıştır. Model performansı, farklı doğruluk hesaplama metrikleri yardımıyla kıyaslanarak en yüksek performansa ulaşılmaya çalışılmıştır. Bu çalışma Balıkesir Üniversitesi Araştırma Hibesi No.:2025/004 tarafından desteklenmiştir.

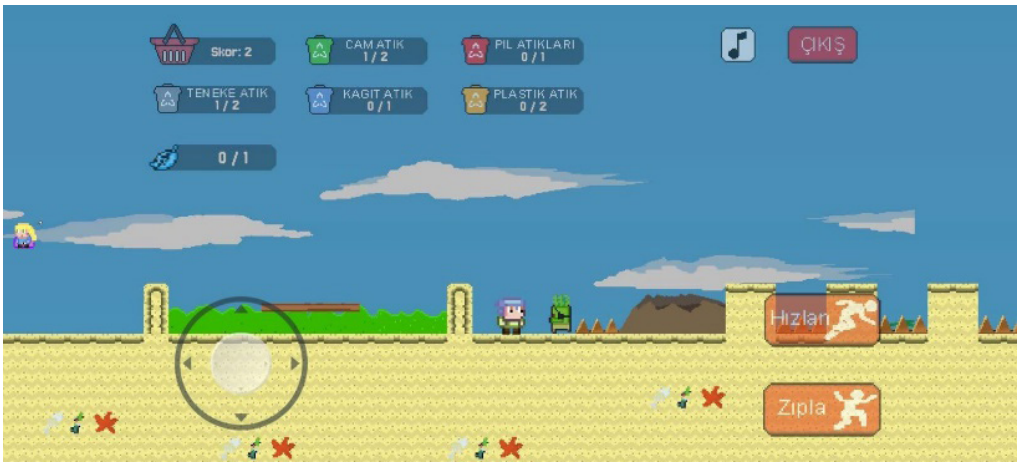
Anahtar Kelimeler: Görüntü işleme, sualtı veri kümesi, derin öğrenme.

SÜRDÜRÜLEBİLİR KALKINMA HEDEFLERİNİN ÇOCUKLARA MOBİL OYUN UYGULAMASI İLE ÖĞRETİMİ

Nur Kuban Torun*
Bilecik Şeyh Edebali Üniversitesi, Yönetim Bilişim Sistemleri, Bilecik, Türkiye
**(İletişimden sorumlu yazar) nurkuban.akdemir@bilecik.edu.tr*

ÖZET

Sürdürülebilir Kalkınma Hedefleri (SKH), daha iyi bir gelecek için belirlenmiş 17 hedeften oluşur. Bu hedefler, yoksulluk ve açlığı sona erdirmekten sağlıklı ve kaliteli yaşamı teşvik etmeye, eğitimde eşitsizlikleri azaltmaktan temiz suya erişimi sağlamaya kadar geniş bir yelpazeyi kapsar. Aynı zamanda, ekonomik büyüme, sanayileşme ve altyapı, sürdürülebilir şehirler ve topluluklar, sorumlu üretim ve tüketim, iklim değişikliği, sucul ve karasal yaşam, barış, adalet ve iş birliklerini de içerir. Bu önemli hedeflerin gelecek nesiller tarafından da sürdürülebilmesi için, çocukların küçük yaşlardan itibaren bu konularda bilinçlendirilmesi gerekmektedir. Bu amaçla, 8-14 yaş grubundaki öğrencilere Sürdürülebilir Kalkınma Hedeflerini öğretmek için bir mobil oyun geliştirilmiştir. Bu oyun, her bir hedefi ele alan 17 bölümden oluşmaktadır ve Unity platformu üzerinde C# programlama dili kullanılarak yazılmıştır. Oyun, çocuklara sürdürülebilirlik konularını eğlenceli ve interaktif bir şekilde öğretmeyi amaçlar. Oyunu oynayan öğrencilere sonrasında anket çalışması yapılarak hedefleri ne kadar anladıkları nitel yöntemlerle tespit edilmiştir. Sonuçlar, projenin öğretim amaçlarına ulaşmada etkili olduğunu göstermiştir. Bu mobil oyun, öğrencilerin öğrenirken eğlenmelerini sağlarken, sürdürülebilirlik kalkınma hedefleri konusunda farkındalıklarını artırarak onları daha bilinçli bireyler olmaya teşvik eder.



Şekil 1: Oyun içi örnek görsel örneği

Anahtar Kelimeler: Sürdürülebilir kalkınma hedefleri, mobil oyun, nitelikli eğitim.

SÜRDÜRÜLEBİLİRLİK YÖNETİMİNDE YEŞİL TEKNOLOJİLERİN ÖNEMİ

Yaprak Akçay Zileli*
Ankara, Türkiye
**(İletişimden sorumlu yazar) yaprakzileli@gmail.com*

ÖZET

Sürdürülebilirlik, günümüz ihtiyaçlarını karşılarken gelecek nesillerin kendi ihtiyaçlarını karşılama yeteneğini tehlikeye atmayan bir kalkınma modelidir. Bu doğrultuda şekillenen kurumsal sürdürülebilirlik, işletmelerin çalışanlar, hissedarlar, müşteriler ve toplum gibi iç ve dış paydaşlarının ihtiyaçlarını, gelecekteki faaliyetlerinden ödün vermeden karşılamasını amaçlamaktadır. Kurumsal sürdürülebilirlik; ekonomik, çevresel ve sosyal olmak üzere üç boyutu dengeli biçimde gÖZETmeyi gerektirir.

Ekonomik sürdürülebilirlik, işletmelerin kârlılığını sürdürmesi, yenilikçi modeller geliştirmesi ve büyümenin çevresel ve sosyal etkilerini dikkate almasını içerir. Çevresel sürdürülebilirlik ise doğal kaynakların korunmasını, emisyon ve atıkların azaltılmasını, yenilenebilir enerji kaynaklarının kullanımını ve ekosistemlerin sürdürülebilirliğini hedefler. Sosyal sürdürülebilirlik ise toplumsal refah, insan hakları, adalet, kültürel çeşitlilik ve eşitliği destekler. İşletmelerin bu alanda sorumluluk alması; adil ücret, iyi çalışma koşulları ve etik değerlere bağlılık gibi unsurları kapsar.

Yeşil teknolojiler, çevre dostu üretim süreçlerini benimseyerek enerji verimliliği sağlar, kaynak kullanımını optimize eder ve çevresel etkileri en aza indirir. Bu teknolojiler; yenilenebilir enerji, akıllı şehir sistemleri, atık ve su yönetimi, karbon yakalama ve tarım teknolojilerini içermektedir.

Sonuç olarak, yeşil teknolojilerin iş süreçlerine entegrasyonu, sürdürülebilirlik yönetiminde temel bir rol üstlenmektedir. Doğal kaynakların korunması, çevresel etkilerin azaltılması ve toplumsal refahın artırılması gibi hedefler doğrultusunda işletmelere stratejik avantaj sağlamaktadır. Bu çerçevede, çalışma kapsamında sürdürülebilirlik ve kurumsal sürdürülebilirlik kavramlarına, Birleşmiş Milletler Kalkınma Hedeflerine ve yeşil teknolojilerin sürdürülebilirlik yönetimine entegrasyonuna odaklanılmıştır.

Anahtar Kelimeler: Sürdürülebilirlik, Kurumsal Sürdürülebilirlik, Yeşil Teknolojiler.

TRENDS IN IOT SECURITY: A POST-QUANTUM PERSPECTIVE

Aykut Karakaya*
Zonguldak Bulent Ecevit University, Department of Computer Engineering, Zonguldak, Türkiye
**(Corresponding author) aykut.karakaya@beun.edu.tr*

ABSTRACT

The Internet of Things (IoT) systems constitute critical infrastructures that are widely used across various sectors today, comprising distributed and interconnected structures. As these systems increasingly operate in mobile, resource-constrained, and heterogeneous environments, ensuring their security has become even more complex [1]. Although traditional cryptographic solutions can address some of the current needs to a certain extent, the advancement of quantum computing technologies has introduced new types of threats. In this context, the literature reveals a wide range of approaches developed to balance concerns of both security and efficiency.

Existing studies focus on fundamental security requirements such as secure data transmission, authentication, privacy preservation, and integrity assurance. Proposed methods include a variety of solutions, ranging from blockchain-based approaches for distributed architectures and edge computing-supported frameworks to privacy-focused techniques like federated learning and software-defined network architectures. Moreover, there is a growing interest in lightweight cryptographic techniques and quantum-resistant algorithms, particularly regarding their applicability in resource-constrained devices [2]. Efforts to ensure IoT security in the post-quantum era focus not only on developing new algorithms but also on holistic solutions that take into account system-level criteria such as architectural scalability, latency, energy consumption, and interoperability [3].

This study provides a synthesis of key themes in the IoT security literature, particularly focusing on post-quantum security, and offers a comparative analysis of solution trends.

Keywords: Internet of things, cybersecurity, post-quantum security.

REFERENCES

[1] Karakaya and S. Akleylek, "A survey on security threats and authentication approaches in wireless sensor networks", 2018 6th International Symposium on Digital Forensic and Security (ISDFS), Antalya, Turkey, 2018, 1-4.

[2] J. Wen, C. Zhou, L. Zhang, "Efficient anti-quantum attack batch authentication and key agreement protocol based on ntru in vanet", Peer-to-Peer Networking and Applications, 18 (1), 2025, 1–21.

[3] Karakaya and F. Arat, "A Survey on Security Requirements, Threats and Protocols in Industrial Internet of Things", IJISS, 10 (4), 2021, 138–152.

WHEAT GERM OIL AS A POTENTIAL DNA-PROTECTIVE AGENT: INSIGHTS FROM SMART ASSAY AND AI-ASSISTED ANALYSIS

Melih Can Demirbek*
Balıkesir University, Molecular Biology and Genetics, Balıkesir, Türkiye

Serap Doğan
Balıkesir University, Molecular Biology and Genetics, Balıkesir, Türkiye

Begümhan Yılmaz Kardaş
Balıkesir University, Molecular Biology and Genetics, Balıkesir, Türkiye
**(Corresponding author) melihcdemirbek@gmail.com*

ABSTRACT

Wheat germ oil is a nutrient-rich oil containing potent antioxidant and anti-inflammatory compounds. Wheat germ is rich in carotenoids, polyphenols, and tocopherols. The oil, obtained by milling, is widely used in food and cosmetics. Additionally, due to its methionine, threonine, and lysine content, it has medical relevance [1].

Exposure to genotoxic substances is inevitable due to the widespread use of plants, agricultural chemicals, and cosmetics. The somatic mutation and recombination test (SMART) is used to detect such substances. Thanks to its metabolism and high genomic similarity to humans, Drosophila melanogaster is widely used in genetic toxicology. The SMART test detects genotoxicity by observing genetic changes such as mutations, deletions, point mutations, nondisjunction, and recombination in fly wing hairs [2, 3].

This study investigated the antigenotoxic potential of wheat germ oil using the SMART assay. Data were statistically analyzed using three AI tools: ChatGPT, Deepseek, and Microsoft Copilot.

The results showed that wheat germ oil is non-genotoxic and significantly reduced H_2O_2-induced genetic damage. These findings were confirmed via AI-based statistical analysis. The AI tools efficiently calculated mutation frequencies. Minor differences appeared in percentage inhibition calculations. In Student's t-test, p-values varied slightly across the three AIs. Yet, all results aligned with Excel-based manual calculations when using $p<0.05$ as the significance threshold.

Wheat germ oil emerges as a potential natural agent against DNA damage. Furthermore, AI-assisted statistics enhance speed and reliability in molecular biology and genetics.

Keywords: Wheat, genotoxic, artificial intelligence.

REFERENCES

[1] G. S. Kumar and A. G. G. Krishna, "Studies on the nutraceuticals composition of wheat derived oils wheat bran oil and wheat germ oil," J Food Sci Technol, vol. 52, no. 2, pp. 1145–1151, Feb. 2015, doi: 10.1007/S13197-013-1119-3.

[2] I. Gaivão, J. Ferreira, L. M. Sierra, I. Gaivão, J. Ferreira, and L. M. Sierra, "The Somatic Mutation and Recombination Test (SMART) of Drosophila melanogaster for Detecting Antigenotoxic Activity," Genotoxicity and Mutagenicity - Mechanisms and Test Methods, Mar. 2020, doi: 10.5772/INTECHOPEN.91630.

[3] P. Pitchakarn, W. Inthachat, J. Karinchai, and P. Temviriyanukul, "Human Hazard Assessment Using Drosophila Wing Spot Test as an Alternative In Vivo Model for Genotoxicity Testing—A Review," Int J Mol Sci, vol. 22, no. 18, p. 9932, Sep. 2021, doi: 10.3390/IJMS22189932.

YAPAY ZEKÂ VE GAZETECİLİK: LİSANSÜSTÜ TEZLERE İLİŞKİN BİR İNCELEME

Mutluhan Çamur*
Zonguldak Bülent Ecevit Üniversitesi, İletişim Fakültesi, Zonguldak, Türkiye
**(İletişimden sorumlu yazar) mutluhan.camur@beun.edu.tr*

ÖZET

Bu araştırmada yapay zekâ ve gazetecilik ilişkisi lisansüstü tezler bağlamında değerlendirilmiştir. Araştırmada gazetecilik alanında yapay zekâyı ele alan tezlerin bu teknolojiye hangi perspektiflerden odaklandığını tespit etmek ve gelecekteki çalışmalara yol göstermek amaçlanmıştır. Tezlere YÖK Ulusal Tez Merkezi veri tabanından erişilmiştir. Erişilen tezler araştırma soruları bağlamında nitel ve karma yöntem analiz programı Maxqda 24'te kodlanarak içerik analizi yöntemiyle incelemiştir. Bulgular, gazetecilik alanında yapay zekâyı irdeleyen tezlerin sınırlı sayıda olmakla birlikte büyük çoğunluğunun 2024 yılında yazıldığını ve yüksek lisans düzeyinde olduğunu ortaya koymuştur. Yine incelenen tezlerin çoğunda nitel yöntem kullanıldığı ve yapay zekânın gazeteciliğe olan etkisine odaklanıldığı saptanmıştır. Ayrıca gazetecilik ve yapay zekâ ilişkisinin farklı perspektiflerden değerlendiren kapsamlı tezlere ihtiyaç olduğu tespit edilmiştir.

Anahtar Kelimeler: Yapay zekâ, yapay zekâ ve gazetecilik, içerik analizi.

2020-2024 YILLARI ARASINDA WEB OF SCİNECE VERİ TABANINDA ÜRETKEN YAPAY ZEKA VE PROGRAMLAMA EĞİTİMİ ALANLARINDA YAYINLANMIŞ YAYINLARIN BİBLİYOMETRİK ANALİZİ

Erdinç Karaman*
Bartın Üniversitesi, Bartın, Türkiye

Mesut Erdönmez
Bartın Üniversitesi, Bartın, Türkiye
(İletişimden sorumlu yazar) erdinckaraman@bartin.edu.tr

ÖZET

Bu çalışma, 2020-2024 yılları arasında Web of Science (WoS) veri tabanında "Generative Artificial Intelligence" (GAI), "Artificial Intelligence" (AI) [Topic], "Programming", "Programming Education" ve "Education" [All Feild] anahtar kelimeleri kullanılarak yapılan tarama sonucunda elde edilen 2.497 belgenin bibliyometrik analizini içermektedir. Analizler, bu alandaki araştırmaların küresel ölçekte nasıl geliştiğini, hangi ülkelerin ve kurumların öne çıktığını ve hangi temaların öncelikli olduğunu ortaya koymayı amaçlamaktadır [1][2]. Bu çalışmada, yapay zeka ve programlama alanlarındaki bilimsel yayınların bibliyometrik analizi için R programı kullanılmıştır. R, özellikle bibliyometrik analizler için geliştirilmiş "bibliometrix" paketi ile bu tür çalışmalar için oldukça uygun bir araçtır [1][2]. Analiz süreci, Web of Science (WoS) veri tabanından elde edilen verilerin R ortamında işlenmesi, analiz edilmesi ve görselleştirilmesi aşamalarını içermektedir. Analiz sonuçlarının görselleştirilmesi için R'de bulunan "biblioshiny" gibi grafik paketleri kullanılmıştır [1]. Bu sayede, yayınların yıllara göre dağılımı, en çok atıf alan makalelerin görselleştirilmesi, yazar işbirliklerinin ağ grafikleri ve anahtar kelime bulutları gibi görsel çıktılar oluşturulmuştur [3][4].

Elde ettiğimiz verilerden bazı analizlerimize bu ÖZETte değineceğiz. 2020-2024 yılları arasında Web of Science (WoS) veri tabanında "Üretken Yapay Zeka" ve "Programlama Eğitimi" alanlarında yayınlanan makale sayısı, 2020'de 55 iken 2024'te 1.661'e ulaşarak belirgin bir artış göstermiştir. Bu sıçramanın, küresel teknolojik gelişmelerin ivme kazanması ve akademik araştırmalardaki disiplinler arası boşlukların doldurulma çabalarıyla ilişkili olduğu değerlendirilmektedir. 2023 yılı, makale başına ortalama 20,88 atıf ve yıllık 6,96 atıf ortalamasıyla nitelikli çalışmaların yoğunlaştığı bir dönem olarak öne çıkmıştır. 2024 yılında ise yayın sayısının 1.661'e ulaşmasına rağmen, makalelerin yeniliği nedeniyle atıf ortalaması 3,68 ile sınırlı kalmıştır. WANG Yinying (19 makale) ve CHIU Thomas K.F.(16 Makale) gibi yazarların çok yazarlı ekiplerle geniş işbirliklerine odaklandığını, ancak bireysel katkı oranlarının düşük olduğunu ortaya koymuştur. Buna karşılık, SU Jiahong (fraksiyonel katkı: 6,87) ve BOZKURT Aras (fraksiyonel katkı: 5,04) gibi araştırmacılar, bağımsız veya küçük ekiplerle yürüttükleri çalışmalarda yüksek bireysel etki sağlamıştır. KIM Jinhee (11 makale) ve ZHANG Yumei (14 makale) ise hem üretkenlik hem de dengeli katkılarıyla

alanda öne çıkan isimlerdir. Makalelerin 420 farklı kaynakta yayınlandığı görülmüştür. "Education and Information Technologies" (241 makale, 2.450 atıf) teknoloji entegrasyonu ve yapay zeka tabanlı eğitim araştırmalarında lider konumdadır. Anahtar kelime analizinde ise "education" (%32), "artificial-intelligence" (%28) ve "students" (%25) terimlerinin dominant olduğunu görülmüştür. Özellikle 2024'te "artificial-intelligence" kelimesinin kullanımının zirve yapması, teknolojinin eğitimdeki rolüne yönelik artan vurguyu yansıtmaktadır. Ülkelerin yayın performansını incelediğimizde; Çin (5.455 atıf) ve ABD (3.602 atıf), en yüksek atıf performansına sahip ülkeler olarak öne çıkmıştır. Bu durum, söz konusu ülkelerin araştırma altyapısı ve kaynak yatırımlarıyla ilişkilendirilebilir. En çok atıf alan makaleleri incelediğimizde, Debby R.E. Cotton (541 atıf) küresel ölçekte yüksek etkili yazar olarak karşımıza çıkmaktadır. BradfordLaw, Atıf ağı ve Ülke, yazar işbirlikleri makalemizde inceleyeceğimiz diğer konular olacaktır.

Anahtar Kelimeler: Üretken yapay zeka, programlama eğitimi, bibliyometrik analiz.

REFERANSLAR

[1] Aria, M., & Cuccurullo, C. (2017). "Bibliometrix: An R-Tool For Comprehensive Science Mapping Analysis." Journal of Informetrics, 11(4), 959-975.

[2] Maulani, A. (2024). "Bibliometric analysis: Adoption of big data analytics in financial auditing." Business Informatics, 2(2), 78-89.

[3] Önder, H., & Tirink, C. (2022). "Bibliometric analysis for genomic selection studies in animal science." Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi.

[4] Purwanto, E. (2023). "Bibliometrics analysis of digital leadership research." Ijaeb, 1(4), 2216-2230.

FULL PAPERS

(Tam Metinler)

A COMPARATIVE ANALYSIS OF DEEP LEARNING MODELS FOR PLANT LEAF DISEASE CLASSIFICATION

Buse Calisir*
Department of Software Engineering, Technology Faculty, Fırat University, Elazığ, Türkiye

Bihter Das
Department of Software Engineering, Technology Faculty, Fırat University, Elazığ, Türkiye
(Corresponding author) bcalsr2709@outlook.com

ABSTRACT

Accurate classification of plant diseases is of great importance to prevent data loss in agricultural production and to ensure sustainable agriculture. In this study, we present an approach that detects and classifies plant leaves as healthy or unhealthy. In this study, a comprehensive performance comparison between different deep learning models is performed using a self-created dataset of healthy and unhealthy leaf images. In this context, AlexNet, VGG16, ResNet, EfficientNet, DenseNet, Inception and Xception models, which are among the common CNN architectures based on deep learning, are analyzed and examined. The study is analyzed in two classes as healthy and unhealthy leaf types and tested with 20 different plant species data. According to the results obtained, it is seen that DenseNet is the model with the highest performance with 99% precision, 98% recall, 98% F1 score and 98% accuracy. The study reveals that deep learning provides an effective and reliable solution for plant health detection and provides a strong basis for automated systems to be developed in this field, both with the specific data set and with multiple model comparison.

Keywords: Deep Learning, plant disease classification, CNN.

Introduction

The agricultural sector is of great importance both in terms of meeting the basic food needs of people and contributing to the economic development of countries. Today, changes and developments on a global scale require agricultural production to be carried out both sustainably and efficiently, which plays a critical role in ensuring food security [1]. Agriculture is not only limited to food supply, but is also a fundamental building block in many areas such as economic growth, social balance and improvement of rural living conditions. In this respect, the protection of plant health is a crucial element to ensure the continuity of agricultural production, increase productivity and ensure food supply security at the global level [2].

Early detection of plant diseases has a direct impact on both product quality and yield. Leaf deterioration, in particular, provides important clues about the overall health of the plant. Since such assessments usually require expert knowledge, they are both time-consuming and error-prone when performed manually; however, with the development of artificial intelligence-based image processing methods in recent years, it has become possible to perform such analyses automatically and more reliably.

Deep learning approaches, especially convolutional neural networks (CNNs), have shown highly effective and successful results in image classification. Therefore, classification applications for determining the health status of plant leaves have become an important component of agricultural decision support systems. However, it is noteworthy that many studies in the literature have been conducted with data sets generated in prepared or controlled conditions, which may limit the applicability of the models in real field environments.

This study aims to detect plant diseases using deep learning methods and classify them into two main groups as healthy and unhealthy by analyzing leaf samples of 20 different plant species. The developed models are tested with a specially created dataset and their performances are comparatively evaluated.

The main contributions of this research are as follows:

- The study analyzes general plant leaf images to identify disease types.
- Seven different deep learning-based methods were applied and performance comparisons were made.
- A comprehensive review is presented by evaluating disease detection on more than twenty different plant species.
- Leaves were classified into two basic classes: healthy and unhealthy.
- All experimental studies were carried out using the original dataset created within the scope of the research.

Related Works

In this section, the studies in the literature are reviewed in detail and the artificial intelligence-based methods used in the detection and classification of plant diseases are comprehensively evaluated.

Sujatha et al. [3] comparatively used both machine learning and deep learning based models to detect diseases on the leaves of citrus plants. Their results show that deep learning methods achieve higher accuracy rates in disease detection and perform better than other methods. Too et al. [4] comparatively evaluated various deep learning models with fine-tuning method on PlantVillage dataset. As a result of the analysis, the highest success is obtained with the DenseNet121 model. Barbedo [5], in his study, he examined disease detection on small datasets and trained the model with both original and background removed data. In some cases, it is observed that training with background sub-

tracted data increases the accuracy rates. However, due to the small datasets used, it is stated that the model accuracy is largely affected by the number of data. Atila et al. [6] in their study first performed comparisons with various deep learning models without data augmentation. Then, they re-analyzed the same models by increasing the data sets. In both cases, the highest accuracy rate is obtained with the EfficientNet-B5 model. Sachdeva et al. [7] in their study aimed to increase the accuracy and prevent overlearning by integrating a Bayesian learning approach to the CNN model to classify plant diseases. With the CNN-Bayesian model they developed, they show that they achieved 40% higher accuracy compared to traditional machine learning models using fewer parameters. Thaiyalnayaki and Joseph [8] performed a comparative classification with SVM and MLP models using soybean plant data in the UCI dataset. As a result of the analysis, it is observed that SVM model gives more successful results in disease classification. Panchal et al. [9] classify plant diseases as early, intermediate and advanced. In this classification process, various deep learning models were compared and as a result of the analysis, the highest accuracy rate is obtained with the VGG16 model. Kotwal et al. [10] compare methods such as SVM, K-Means, LBP with deep learning architectures. In order to better interpret the model performances, they used visualization techniques such as saliency maps, Grad-CAM and feature maps. The deep learning models they evaluated with these visualization approaches give successful results with high accuracy rates. K.P. [11] compares the results of CNN and transfer learning based models using grape leaf images from a ready-made dataset. The training process is evaluated with four different accuracy metrics and the highest performance is obtained with the DenseNet201 model. Kundu et al. [12] in their study, performs analysis with the AlexNet model by applying various image processing steps such as grayscale conversion, Canny edge detection, thresholding and noise removal in the classification of sorghum plant diseases. Albattah et al. [13], in their study, they develop a special CenterNet architecture based on DenseNet-77. The proposed model is tested on the PlantVillage dataset with various multiple loss functions. Comparisons are also made with different deep learning architectures and it is shown that the proposed model provides the highest accuracy rate compared to other models. Reddy and Neeraja [14] develop a method for the detection of leaf diseases and damage. In the study, DenseNet model is used for classification and 1D CNN-based semantic segmentation method based on pixels is used for damage detection. These models give more successful results in classification and segmentation processes compared to other methods. Sholihati et al. [15] classify potato leaf diseases under five different classes in their study. For this purpose, VGG16 and VGG19 deep learning architectures are used and successful classification results are obtained with high accuracy rates.

In this study, a unique dataset consisting of healthy and unhealthy leaves of different plant species was created and classification was performed with various deep learning models. The data were collected in real agricultural environments to increase the generalization capability of the model. Unlike the studies in the literature, which are commonly conducted in laboratory conditions or with ready-made data sets, this study aims to contribute to the literature with the natural data set and comparative analysis approach used in this study.

Materials and Methods

In this study, different deep learning models are examined and their performances are evaluated in order to detect and classify healthy or unhealthy plant leaves. Preprocessing steps such as data augmentation, image resizing, conversion of labels into one-hot vectors and calculation of class weights to reduce data imbalance were applied in the implementation process. Classification was performed with different deep learning architectures such as ResNet, VGG16, VGG19, AlexNet, InceptionV3, Xception, DenseNet and EfficientNet. These methods were used to detect diseases in plant leaves and to distinguish healthy leaves. Figure 1 shows the overall architecture of the system. The proposed approach contributed to decision-making processes by automatically detecting plant diseases and achieved successful classification results.

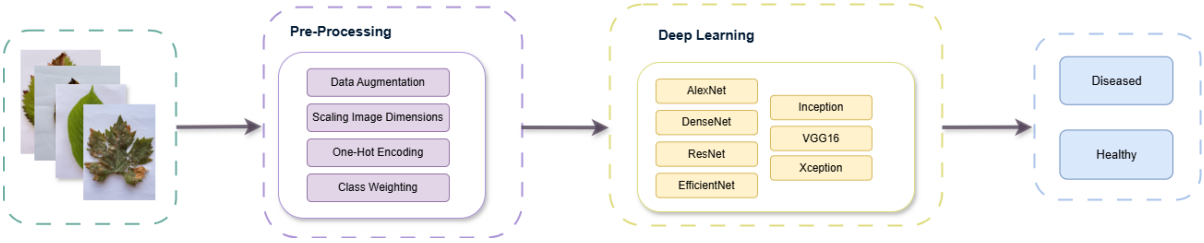


Figure 1: System Flowchart

Datasets

The dataset used in this study consists of leaf images of more than twenty plant species. The high-resolution images collected from different gardens and various plant growers were obtained with a camera with a resolution of 4284 x 5712 pixels and an aperture of f/1.6. Within the scope of the study, leaves were categorized into two different classes: healthy and unhealthy. In total, a balanced dataset of 498 healthy and 498 unhealthy leaf images was used. Figure 2 shows sample images of this dataset.

Figure 2: Plant Leaves Dataset

Preprocessing and Data Augmentation

In this study, data preprocessing and data augmentation steps are included in order to train the developed deep learning models properly. All images were resized to fit the input dimensions required by the deep learning models and pixel values were normalized between 0 and 1. To make the model more robust and generalizable to different situations, various data augmentation methods such as rotation, scaling, panning, brightness adjustment and horizontal translation were used. This makes the models more robust to different image variations and reduces the risk of over-learning.

Materials and Methods

This study uses these deep learning methods for classification and disease detection.

ResNet: Developed in 2015 by Kaiming He and his team at Microsoft Research Asia, it is widely used in image classification and feature extraction [16]. This model is designed with building blocks based on residual learning instead of direct learning, and deep architectures are created by stacking these blocks on top of each other. These deep structures demonstrate the effectiveness of the residual learning approach by providing high accuracy rates in image classification tasks [17].

AlexNet: AlexNet was developed by Krizhevsky and his team in 2012 and is considered a groundbreaking development in the field of deep learning with its success in the ImageNet competition [18]. This model, which is widely used especially in object detection applications, owes its success to ImageNet, one of the large and comprehensive datasets. Dropout method is applied to prevent overlearning during the training process; this method increases the generalization ability of the model by temporarily disabling neurons at certain rates [19].

Inception: The Inception architecture was designed by Szegedy and his team to balance the problems of multi-scale feature extraction, computational cost and network depth. In this architecture, convolution filters of different sizes work in parallel at the same time, enabling both detailed feature extraction and computational efficiency. Many studies have shown that the Inception architecture provides successful results and high accuracy rates in tasks such as the detection of plant diseases and pests in agricultural fields [20].

VGG16: VGG16 was developed by Simonyan and Zisserman from Oxford University and stands out as a convolutional neural network architecture that draws attention with its deep layer structure. This model, which consists of 13 convolutional and 3 fully connected layers in total, gives successful results in many classification problems thanks to its simple and regular structure [21].

DenseNet: Developed in 2016 by Huang and his team, DenseNet is an innovative deep learning architecture based on the principle of dense connections between layers. In this structure, each layer directly receives the output of all previous layers, thus increasing information sharing and reuse of features. This approach increases the parameter efficiency of the network and improves the learning process by enabling the gradient flow to be transmitted more efficiently. Thanks to these features, DenseNet provides successful results in reducing information loss and learning difficulties that occur especially in deep networks [22].

EfficientNet: It is an advanced convolutional neural network architecture developed by Google AI that stands out with its high accuracy, efficient training process and simple structure [23]. In this study, the model is implemented on the ImageNet dataset using a transfer learning approach with pre-trained weights. In order to preserve the basic features in the early layers, only the last layers are made suitable for retraining. At the output of the model, multiclass prediction is performed with softmax activation, and the risk of overlearning is reduced by adding Dropout and Global Average Pooling layers.

Xception: Xception is a modern convolutional neural network architecture developed to improve efficiency in deep learning. Instead of traditional convolution layers, it uses a depthwise separable convolution structure where channels are processed separately, requiring less computation and reducing the number of parameters. With its architectural structure, it is observed to produce successful results with high accuracy and strong performance, especially on large datasets [24].

Experimental Results

This study aims to identify and classify the health status of plant leaves. In line with this goal, a comprehensive analysis was performed using various deep learning models. The applied methods include CNN (Convolutional Neural Network) based models such as VGG16, DenseNet, AlexNet, ResNet, EfficientNet, Xception and Inception. The performance of each model has been comprehensively evaluated on the basis of criteria such as precision, recall, f1-score and accuracy, and the findings are shown in Table 1.

Table 1: Performance Metrics for Each Models

Method	Recall	Precision	F1 Score	Accuracy
VGG16	0.93	0.93	0.93	0.93
DenseNet	0.98	0.99	0.98	0.98
AlexNet	0.87	0.87	0.87	0.87
EfficientNet	0.97	0.97	0.97	0.97
Xception	0.86	0.87	0.86	0.86
Inception	0.86	0.86	0.85	0.86
ResNet	0.95	0.95	0.95	0.95

Table 1 compares the recall, precision, F1 score and accuracy values of the plant classification process by developing different deep learning architectures. As a result of the comparison, DenseNet architecture shows the highest accuracy rate in all metrics and stands out as the most successful model with 98%. This success can be attributed to DenseNet's strong feature sharing, where each layer directly utilizes information from previous layers. ResNet and EfficientNet are the next most successful models after DenseNet with 95% and 97% accuracy. While ResNet facilitates learning in deep networks through residual connections, EfficientNet produced successful results by providing an optimal balance between resource utilization and accuracy. Another method, VGG16, with an accuracy rate of 93%, is lower than the other models, but still performs well. However, AlexNet, Xception and Inception architectures achieved lower success compared to the other models, staying in the 86%-87% accuracy range. This result shows that simpler or older architectures are not as effective as deep and connected architectures, especially in classifying complex plant leaf images. As a result, DenseNet achieved the best performance

among the tested models. For datasets with detailed information such as agricultural leaf images, DenseNet, ResNet and EfficientNet, which are characterized by their deep and connected structures, are more effective.

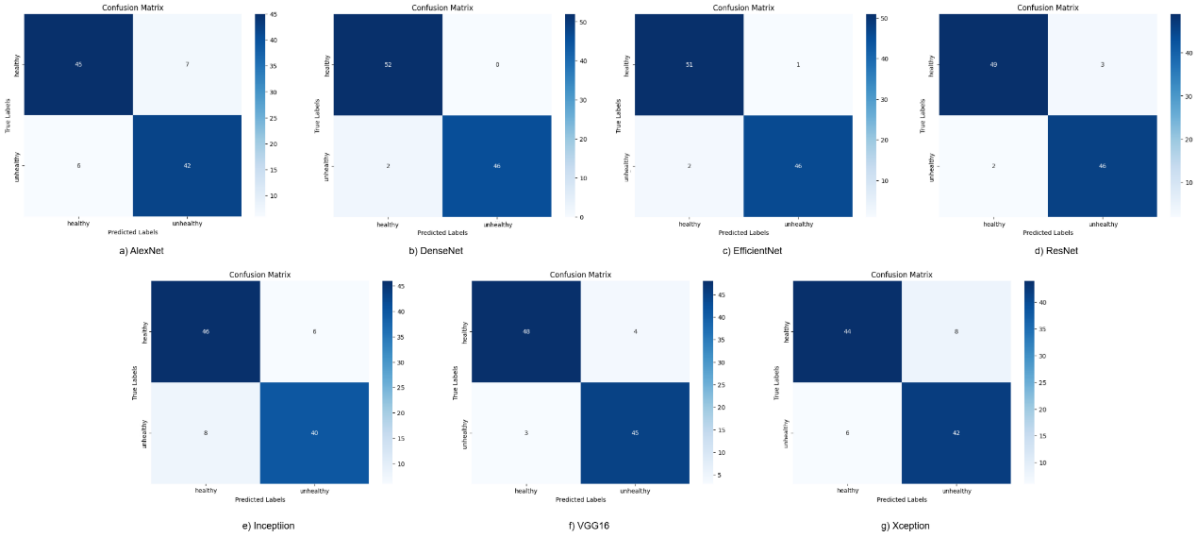


Figure 3: Deep Learning Models Confusion Matrices

The confusion matrices presented in Figure 1 show the comparative performance of the evaluated deep learning architectures in classifying healthy and unhealthy leaves. According to these results, the DenseNet architecture achieved the highest accuracy rate by making almost error-free predictions for both classes and demonstrated a prominent success in the classification task. EfficientNet and ResNet models also produced successful results with low misclassification rates, showing a fairly balanced distribution. VGG16 provided a stable and reliable performance with a reasonable level of error. On the other hand, AlexNet, Inception and Xception architectures showed relatively more errors, especially when unhealthy instances were assigned to the healthy class. These findings suggest that deeper and more structurally advanced architectures provide higher performance in tasks that require detailed analysis, such as agricultural image classification.

Table 2: Comparative Analysis of the Proposed Method with Methods in the Literature

Author	Method	Dataset	Accuracy
Demilie [25]	ResNet50	Custom Plant Dataset	%91,25
Reddy and Neeraja [14]	1D-CNN	PlantVillage	%97
Sholihati et al. [15]	VGG16	PlantVillage(Potato), Google Images, Malang	%91,31
The Propesed	DenseNet	New Dataset(20 Plants Type)	%98

Table 2 compares the methods, datasets and accuracy rates obtained in different studies. While most studies used ready-made datasets such as PlantVillage and Google Images for classification, the proposed method was tested on a new dataset that we created ourselves and included 20 different plant species. Our model with DenseNet achieved 98% accuracy and outperformed other models. The results show that unlike existing methods in the literature, the proposed model achieves better results on a new dataset with more plant species and offers a more flexible approach to plant disease classification.

Conclusion

In the agricultural sector, plant health is a fundamental factor that significantly affects product yield and quality. Early detection of plant diseases contributes to the prevention of economic losses in agricultural production and is critical for sustainable agricultural practices. In this study, healthy and unhealthy leaves are analyzed under two classes using various deep learning-based models to classify the health status of plant leaves. Since traditional methods require expert knowledge, are time-consuming and costly, the speed and accuracy provided by artificial intelligence-supported approaches are of great importance in this field. Accordingly, image processing based deep learning models were used for classification and their performances were compared in detail. According to the experimental results, the DenseNet model was the most successful method in the study with 98% accuracy, 98% F1 score, 99% precision and 98% recall. DenseNet is followed by EfficientNet with 97% accuracy and ResNet with 95% accuracy. Among the other models, VGG16 with 93%, AlexNet with 87%, Xception and Inception with 86% accuracy were less successful. The results show that deep learning models can be used effectively in leaf health classification. Especially the DenseNet architecture stands out as a good method for applications in this field with its high success rate. In future studies, the generalization capability of the model can be increased by expanding the dataset and including different plant species and disease symptoms in the classification. Furthermore, testing these models under real field conditions and different environmental factors will provide valuable outputs for field applications.

Declaration of Interest: The author declares that there are no known competing financial interests or personal relationships that could appear to influence the work reported in this paper.

Author Contributions: Buse Calisir: Writing – original draft, Software, Methodology, Bihter Das: Supervision, Conceptualization.

REFERENCES

[1] M.H. Saleem, J. Potgieter, and K.M. Arif, "Plant disease detection and classification by deep learning", Plants, 8(11), 468, 2019.

[2] J. Liu and X. Wang, "Plant diseases and pests detection based on deep learning: a review", Plant Methods, 17, 1–18, 2021.

[3] R. Sujatha, J.M. Chatterjee, N.Z. Jhanjhi, and S.N. Brohi, "Performance of deep learning vs machine learning in plant leaf disease detection", Microprocessors and Microsystems, 80, 103615, 2021.

[4] E.C. Too, L. Yujian, S. Njuki, and L. Yingchun, "A comparative study of fine-tuning deep learning models for plant disease identification", Computers and Electronics in Agriculture, 161, 272–279, 2019.

[5] J.G.A. Barbedo, "Impact of dataset size and variety on the effectiveness of deep learning and transfer learning for plant disease classification", Computers and Electronics in Agriculture, 153, 46–53, 2018.

[6] Ü. Atila, M. Uçar, K. Akyol, and E. Uçar, "Plant leaf disease classification using EfficientNet deep learning model", Ecological Informatics, 61, 101182, 2021.

[7] G. Sachdeva, P. Singh, and P. Kaur, "Plant leaf disease classification using deep convolutional neural network with Bayesian learning", Materials Today: Proceedings, 45, 5584–5590, 2021.

[8] K. Thaiyalnayaki and C. Joseph, "Classification of plant disease using SVM and deep learning", Materials Today: Proceedings, 47, 468–470, 2021.

[9] A.V. Panchal, S.C. Patel, K. Bagyalakshmi, P. Kumar, I.R. Khan, and M. Soni, "Image-based plant diseases detection using deep learning", Materials Today: Proceedings, 80, 3500–3506, 2023.

[10] J. Kotwal, R. Kashyap, and S. Pathan, "Agricultural plant diseases identification: From traditional approach to deep learning", Materials Today: Proceedings, 80, 344–356, 2023.

[11] K.P. Akshai and J. Anitha, "Plant disease classification using deep learning", in Proc. 2021 3rd International Conference on Signal Processing and Communication (ICPSC), pp. 407–411, May 2021, IEEE.

[12] R. Kundu, U. Chauhan, and S.P.S. Chauhan, "Plant leaf disease detection using image processing", in Proc. 2022 2nd International Conference on Innovative Practices in Technology and Management (ICIPTM), vol. 2, pp. 393–396, February 2022, IEEE.

[13] W. Albattah, M. Nawaz, A. Javed, M. Masood, and S. Albahli, "A novel deep learning method for detection and classification of plant diseases", Complex & Intelligent Systems, 1–18, 2022.

[14] B. Sai Reddy and S. Neeraja, "Plant leaf disease classification and damage detection system using deep learning models", Multimedia Tools and Applications, 81(17), 24021–24040, 2022.

15] R.A. Sholihati, I.A. Sulistijono, A. Risnumawan, and E. Kusumawati, "Potato leaf disease classification using deep learning approach", in Proc. 2020 International Electronics Symposium (IES), pp. 392–397, September 2020, IEEE.

[16] M. Arun, M. Deenadhayalan, R. Deepak, S. Balaraman, and D. Deepak, "Lung cancer detection using ResNet-50 CNN architecture", in Proc. 2024 Second International Conference on Advances in Information Technology (ICAIT), vol. 1, pp. 1–6, July 2024, IEEE.

[17] J. Xu, Y. Pan, X. Pan, S. Hoi, Z. Yi, and Z. Xu, "RegNet: self-regulated network for image classification", IEEE Transactions on Neural Networks and Learning Systems, 34(11), 9562–9567, 2022.

[18] H. Cheng, "Advancements in image classification: From machine learning to deep learning", in ITM Web of Conferences, vol. 70, p. 02016, 2025, EDP Sciences.

[19] T. Sivakumari, and R. Vani, "Performance Analysis of Alexnet for Classification of Knee Osteoarthritis," Current Medical Imaging, vol. 20, no. 1, pp. e090823219574, 2024.

[20] J.P. Nyakuri, C. Nkundineza, O. Gatera, and K. Nkurikiyeyezu, "State-of-the-art deep learning algorithms for Internet of Things-based detection of crop pests and diseases: A comprehensive review", IEEE Access, 2024.

[21] M. Hameed, A. Zameer, and M.A.Z. Raja, "A comprehensive systematic review: Advancements in skin cancer classification and segmentation using the ISIC dataset", CMES-Computer Modeling in Engineering & Sciences, 140(3), 2024.

[22] V. Jadeja, A.L.N. Rao, A. Srivastava, S. Singh, P. Chaturvedi, and G. Bhardwaj, "Convolutional neural networks: A comprehensive review of architectures and application", in Proc. 2023 6th International Conference on Contemporary Computing and Informatics (IC3I), vol. 6, pp. 460–467, September 2023, IEEE.

[23] E. Hassan, "Enhancing coffee bean classification: A comparative analysis of pre-trained deep learning models", Neural Computing and Applications, 36(16), 9023–9052, 2024.

[24] A. Thakur, S.K. Biswas, S. Majumdar, and D.M. Thounaojam, "A comprehensive review on different CNN architectures", in Proc. 2025 3rd International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT), pp. 1997–2004, February 2025, IEEE.

[25] W.B. Demilie, "Plant Disease Detection and Classification Techniques: A Comparative Study of the Performances," Journal of Big Data, vol. 11, no. 1, p. 5, 2024.

A COMPARATIVE STUDY OF MACHINE LEARNING MODELS FOR DYSLEXIA PREDICTION WITH EXPLAINABLE ARTIFICIAL INTELLIGENCE ANALYSIS

Şule Akça
Bahçeşehir University, Graduate School, İstanbul, Turkey

Tamer Uçar*
Bahçeşehir University, Department of Software Engineering, İstanbul, Turkey
**(Corresponding author) tamer.ucar@bau.edu.tr*

ABSTRACT

Early diagnosis of dyslexia, a specific learning disability that affects individuals' reading and writing skills, is of great importance. In addition to the traditional methods applied for the diagnosis of dyslexia, the inclusion of artificial intelligence in the process in recent years has greatly enabled more detailed analysis. In this study, seven different machine learning models (Random Forest, XGBoost, LightGBM, SVM, Naive Bayes, Neural Network, and k-NN) were applied to predict dyslexia based on linguistic skills assessed through 32 different tasks. Since the data set created with the real data was unbalanced, the data was balanced with the SMOTE method and analyzes were performed. The results showed that the Random Forest model was the most successful model with 96.47% accuracy and 98% recall. However, when analyzed using Explainable Artificial Intelligence (XAI) techniques, different models were found to be prominent for different language skills. XAI analyses identified variables that play a critical role in dyslexia detection and showed that questions Q24-Q32 are the most important factors that stand out in many models. In particular, LIME results indicated that k-NN and SVM exhibited stronger predictive power for specific questions. This study determines the most effective method for dyslexia detection by comparing different machine learning models, improves interpretability by explaining model decisions with XAI techniques, and provides important findings on the relationship between language skills and dyslexia. These results provide guidance for future studies on dyslexia diagnosis and may contribute to the development of more transparent and effective diagnostic systems.

Keywords: Dyslexia detection, machine learning, explainable artificial ıntelligence.

Introduction

Today, the rapid development of technology has brought many innovations in the fields of economy, education, services, etc. as well as health. Among these advancements, Artificial Intelligence (AI) has emerged as a powerful tool. AI techniques, ranging from

machine learning to deep learning, play a crucial role in various health-related fields, such as enhancing new medical systems, managing patient data and records, and treating a wide range of diseases [1]. Medical practice is gradually evolving due to AI. There are numerous AI applications in various medical areas, including clinical, diagnostic, rehabilitative, surgical, and predictive practices. Another significant area where AI is making a difference is in clinical decision-making and the diagnosis of diseases [2].

Researchers have applied different AI-driven methods, including machine learning and deep learning models, to identify diseases like skin, liver, heart conditions, and Alzheimer, which require early diagnosis [1]. Although AI has shown tremendous potential in health care, one area where its application has yet to be fully explored is in the field of learning disabilities, particularly dyslexia. Dyslexia is a prevalent learning challenge that individuals face during their education, impacting their ability to read, write, spell, read fluently, decode words, and transcribe dictation [3]. While not a disease in the traditional sense, dyslexia can significantly impact educational outcomes. As a result, it is critical to recognize and address dyslexic children's learning challenges as early as possible [4]. Just as AI has revolutionized disease diagnosis in the medical field, it holds the promise of improving the early detection and intervention of learning difficulties like dyslexia. While AI has shown significant potential in various fields, its application in dyslexia detection has been largely driven by machine learning techniques.

Machine learning techniques have gained popularity in dyslexia detection recently due to their ability to deliver improved accuracy in identification and more reliable predictive results [4]. While machine learning techniques play a significant role in dyslexia detection, the lack of transparency in these methods poses certain challenges. Current methods for detecting dyslexia, like AI analysis of handwritten text, also face challenges related to explainability. This issue highlights the need for an explainable AI system that allows users to interpret and understand the AI's decision-making process, ensuring transparency and trustworthiness [5]. This is where Explainable AI (XAI) comes into play.

Explainable Artificial Intelligence (XAI) refers to a domain within Artificial Intelligence (AI) focused on developing tools, methods, and algorithms that provide clear, understandable, and intuitive explanations of AI decision-making processes [6]. This lack of explainability can be problematic in sensitive areas such as dyslexia detection, where understanding the rationale behind a diagnosis is crucial. Traditional dyslexia assessment methods depend on subjective evaluations, leading to delays and potential errors. Although deep learning models provide high accuracy, their black-box nature raises concerns about transparency, which makes their acceptance in educational and clinical settings difficult [7]. XAI techniques improve transparency in AI applications by providing clear explanations of how models make decisions, and this ability to interpret the models helps build trust in AI-assisted diagnostics, such as dyslexia detection, tackling the "black-box" issue and promoting their acceptance and integration in clinical environments [8]. By identifying the most influential linguistic features in dyslexia prediction, XAI allows researchers and practitioners to understand which cognitive skills are most affected by dyslexia and how different machine learning models make their predictions. This study aims to bridge this gap by integrating various XAI methods into dyslexia detection to enhance both accuracy and interpretability.

In this study, various machine learning techniques such as Random Forest, XGBoost, LightGBM, SVM, Naïve Bayes, Artificial Neural Networks and k-NN are applied to predict dyslexia based on language features extracted from a dataset. Furthermore, Explainable Artificial Intelligence (XAI) techniques such as LIME, SHAP and PDP have been used to interpret model decisions, identify critical linguistic features and increase the transparency of the dyslexia prediction process. The integration of these methods provides a comparative analysis of different machine learning models, allowing to understand which models are stronger or weaker in which aspects of dyslexia detection.

The aim of this study is to explore the use of machine learning techniques for dyslexia detection, with a particular focus on enhancing the interpretability and transparency of the decision-making process through the integration of Explainable Artificial Intelligence (XAI).

Methodology

Dataset

This study analyzes a dataset used for dyslexia detection. The dataset consists of 32 different tasks designed to assess linguistic skills, including alphabetic awareness, phonological awareness, visual discrimination, working memory, morphological awareness, and syntactic awareness.

A total of 3,644 Spanish-speaking individuals aged 7-17 participated in the study. Among them, 392 (45.2% female, 54.8% male) had a dyslexia diagnosis, while 3,252 (49.7% female, 50.3% male) did not. Participants completed various tests through an online game, and the following dependent variables were collected:

- Number of clicks (may reflect hesitation in processing information)
- Number of correct responses (Hits, indicating success in cognitive tasks)
- Number of incorrect responses (Misses, reflecting attention and perception errors)
- Score (total number of correct responses, measuring overall performance)
- Accuracy rate (Correct responses / Clicks, representing cognitive accuracy)
- Miss rate (Incorrect responses / Clicks, showing the tendency to make errors)

Task Categories

The 32 questions used in the study are categorized based on the cognitive skills they assess:

- **Q1-Q4:** Alphabetic awareness, phonological awareness, visual discrimination
- **Q5-Q9:** Phonological awareness, syllabic awareness, auditory discrimination
- **Q10-Q13:** Matching written and spoken words
- **Q14-Q17:** Visual discrimination and categorization
- **Q18-Q21:** Visual and auditory working memory

- **Q22-Q23:** Phonological and orthographic awareness
- **Q24-Q25:** Morphological and syntactic awareness
- **Q26-Q32:** Word correction, word formation, sentence formation, dictation

Methods

In this study, seven machine learning models were evaluated for dyslexia prediction. They are Random Forest, XGBoost, LightGBM, SVM, Naïve Bayes, Artificial Neural Networks and k-NN and these methods is explained below.

Random forest is an ensemble learning algorithm that enhances prediction accuracy and stability by combining multiple decision trees using bagging and random subspace methods [9]. XGBoost is a scalable, accurate ensemble method using gradient boosting, while LightGBM optimizes speed by selectively sampling high-gradient instances without losing accuracy [10]. Support Vector Machines (SVMs) are machine learning models that optimize loss functions for accurate classification and regression across continuous, binary, and categorical outcomes [11]. Naïve Bayes is a probabilistic classifier based on Bayes' theorem, assuming feature independence, and is widely used for tasks like spam filtering and medical diagnosis despite its simplifying assumptions [12]. Additionally, Neural networks are machine learning models that learn autonomously and are used in applications like search engines, OCR, and computer vision to improve learning processes [13]. Lastly, k-Nearest Neighbors (k-NN) is a classification algorithm that predicts the class label by majority voting from the k closest training instances, based on distance metrics, and is widely used in machine learning [14].

Explainable Artificial Intelligence (XAI) techniques, including SHAP, Grad-CAM, and LIME, improve the interpretability of AI-driven decisions in healthcare, strengthening trust and promoting collaboration [15], and in this study XAI techniques such as LIME, SHAP and PDP have been used to interpret model decisions. The LIME model uses feature importance and partial dependency plots to illustrate which features impact predictions and how they affect the model's decision-making process [16]. Orange color means this feature favors the prediction of dyslexia (increases the probability of having dyslexia). Blue color means this feature favors the non-dyslexia prediction (decreases the likelihood of dyslexia). The value of a feature can be negative or positive, but what matters is how the model reacts to this value.

SHAP values shows the impact of each feature on the model's predictions, allowing for both global and local interpretability [17] and, they measure the influence of each feature on a model's output, improving both interpretability and transparency [18].

X-axis: SHAP value and the impact of the feature on the model prediction.

Y-axis: Features (ranked from most important to least important).

Color: Value of the feature. Red (High) = High value of the feature. Blue (Low) = The value of the attribute is low.

A Partial Dependence Plot (PDP) illustrates how a specific feature influences a model's predictions, providing insights into the model's behavior regarding that variable [19]. The X-Axis shows the values of the selected independent variable. The Y-Axis shows the model's predicted probability or partial dependence value on the dependent variable.

Based on all this, below are the results of the ML and XAI techniques applied to the models. For XAI part, LIME is applied to all models. SHAP is applied to Random Forest, SVM, XGBoost, LightGBM models. Lastly, PDP is applied to Naïve Base, Neural Network and k-NN models. Additionally, there was a serious imbalance in our data set; the number of individuals without dyslexia was much higher than the number of individuals with dyslexia. Considering that this could negatively affect the learning process of the model, SMOTE (Synthetic Minority Over-sampling Technique) was applied to make the data more balanced. It is a popular and widely used algorithm for both over-sampling and under-sampling [20]. SMOTE increased the number of samples of the minority class, making the model fairer to both classes. After the application, positive changes were observed in the model's metrics such as accuracy, sensitivity and F1 score.

Results

Results of machine learning and explainable AI

Performance comparisons were performed using metrics such as accuracy, precision, recall, F1-score and ROC-AUC. In addition, Explainable Artificial Intelligence (XAI) techniques such as LIME and SHAP were applied to better understand the decision processes of the models. In this way, the features on which the model makes predictions and how reliable its decisions are analyzed. In this section, the results obtained are discussed in detail and evaluated in terms of model performance as well as explainability.

The following conclusions are drawn in line with the relevant metrics.

Machine learning analysis

The effective use of machine learning in dyslexia screening depends on precise numerical data, efficient algorithms (such as Random Forest), and performance metrics like recall, accuracy, and precision to guarantee reliable identification of individuals with dyslexia [21].

Table 1 shows classification reports of Dyslexia class (Class 1) and Non-Dyslexia class (Class 0) for following machine learning models.

The **Random Forest** model demonstrated the highest performance among all models, achieving **96.47% accuracy** and an impressive **98% recall**. This indicates that the model effectively identifies individuals with dyslexia while maintaining a strong overall classification capability, making it the most reliable choice for this task. As a result, Random Forest was the algorithm that gave the best results on this dataset and the use of this model is recommended.

Table 1: Classification Report of Classes

Model	Class 1			Class 0		
	Precision	**Recall**	**F1-Score**	**Precision**	**Recall**	**F1-Score**
Random Forest	0.94	0.98	0.96	0.98	0.94	0.96
XGBoost	0.55	0.44	0.49	0.93	0.95	0.94
k-NN	0.17	0.73	0.28	0.94	0.52	0.67
LightGBM	0.54	0.50	0.52	0.93	0.94	0.94
SVM	0.36	0.68	0.47	0.95	0.84	0.89
Naive Bayes	0.15	0.75	0.25	0.93	0.43	0.59
Neural Network	0.75	0.10	0.17	0.89	1.00	0.94

XAI analysis

The explainable artificial intelligence (XAI) techniques applied in this study provide insights into how different machine learning models make predictions for dyslexia detection. The two main techniques used, LIME and SHAP, revealed key variables influencing model decisions. Across multiple models, questions Q24-Q32 consistently emerged as critical features, highlighting the importance of morphological awareness (Q24), syntactic awareness (Q25), letter and syllable rearrangement (Q27-Q28), and word separation (Q29). These results suggest that phonological, orthographic, and executive function skills play a crucial role in distinguishing dyslexic individuals. SHAP and LIME images are shown following for k-NN. Additionally, key features of prediction of Dyslexia and Non-dyslexia classes are shown in Table 2.

Explainable AI (XAI) results for methods

LIME, SHAP, and PDP techniques were applied to interpret model predictions. In the following, LIME analysis is shown for k-NN in Figure 1 and SHAP analysis is shown for Random Forest in Figure 2.

Random Forest

LIME analysis showed that the model predicted a high probability of non-dyslexia (80%) and a lower probability of dyslexia (20%). The key influencing factors were Q24 (Morphological Awareness), Q25 (Syntactic Awareness), and Q31 (Dictation Task). Higher Clicks24 and Clicks25 values were associated with an increased probability of dyslexia. SHAP analysis revealed that Clicks28, Clicks25, and Accuracy29 were the most significant predictors. These features, related to letter/syllable rearrangement (Q27-Q28) and sentence formation (Q29), highlighted the cognitive difficulties experienced by dyslexic individuals.

XGBoost

LIME results indicated that the model predicted a high probability of non-dyslexia but recognized some dyslexia-related features. Specifically, Misses12 (Lexical Awareness, Auditory Memory) was identified as an important factor, and the absence of errors in this task contributed to the non-dyslexia classification. SHAP analysis identified the most influential

factors as age, mother tongue, clicks on specific tasks (Clicks24, Clicks25, Clicks28), and accuracy scores (Accuracy29, Accuracy30, Accuracy32). Older participants showed lower dyslexia probabilities, likely due to compensatory strategies.

k-NN

LIME analysis showed that the model assigned a 90% probability to dyslexia. The most influential features included Hits29 (sentence formation skills), Misses31 (Lexical, Orthographic Awareness, and Auditory Memory), and Hits6 (syllable recognition). PDP results indicated that strong visual and auditory working memory reduced dyslexia probability, while high error rates in phonological awareness tasks correlated with increased dyslexia predictions.

LightGBM

LIME results showed that the model predicted non-dyslexia with 96% probability but highlighted phonological awareness errors (Missrate8, Accuracy12) as contributors to dyslexia risk. SHAP analysis revealed that Missrate8, Missrate31, and Missrate32 were the strongest predictors, reinforcing the significance of phonological and orthographic awareness in dyslexia detection.

SVM

LIME analysis indicated that the model leaned toward dyslexia predictions (75%). Low Score14 (Visual Discrimination and Executive Functions) strongly influenced dyslexia classification, while high Score29 (sentence formation skills) reduced dyslexia probability. SHAP results showed that Score10 (Phonetic and Orthographic Discrimination) and Score14 (Visual Discrimination) were key factors, along with age and native language effects.

Naïve Bayes

LIME analysis revealed that the model consistently predicted non-dyslexia but still highlighted variables typically associated with dyslexia. The independence assumptions of Naïve Bayes may explain its poor sensitivity to dyslexia patterns. PDP results identified key predictors such as syntactic awareness (Q25), phonetic spelling accuracy (Q16), and alphabetic awareness (Q4). Strong performance in these tasks lowered dyslexia risk.

Neural Network

LIME results showed that the model struggled with dyslexia prediction, likely due to feature prioritization. While Missrate26 and Age favored dyslexia classification, they were not strong enough predictors. PDP analysis indicated that Missrate26 had a significant impact on dyslexia probability, reinforcing the role of phonological and orthographic awareness in dyslexia assessment.

The findings suggest that linguistic abilities, particularly morphological and phonological awareness, auditory and visual working memory, and syntactic awareness, play a crucial role in dyslexia detection. Among the models evaluated, Random Forest achieved the highest overall accuracy; however, LIME analysis indicated that individual predictions

sometimes leaned towards non-dyslexia, highlighting potential limitations in capturing dyslexia-related features. Feature importance analysis using SHAP and LIME consistently emphasized the significance of tasks Q24-Q32 in distinguishing dyslexic individuals. The k-NN model exhibited weaker general performance, but LIME showed high confidence in certain predictions. LIME analysis highlighted Hits29 (Word Separation Task), Misses31 (Dictation Errors), and Clicks32 (Interaction in Dictation Tasks) as major factors in predicting dyslexia.

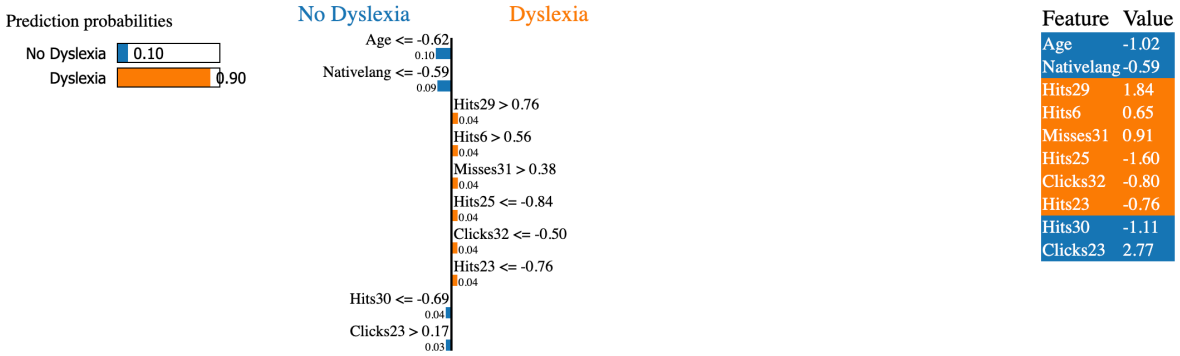


Figure 1: LIME result for k-NN

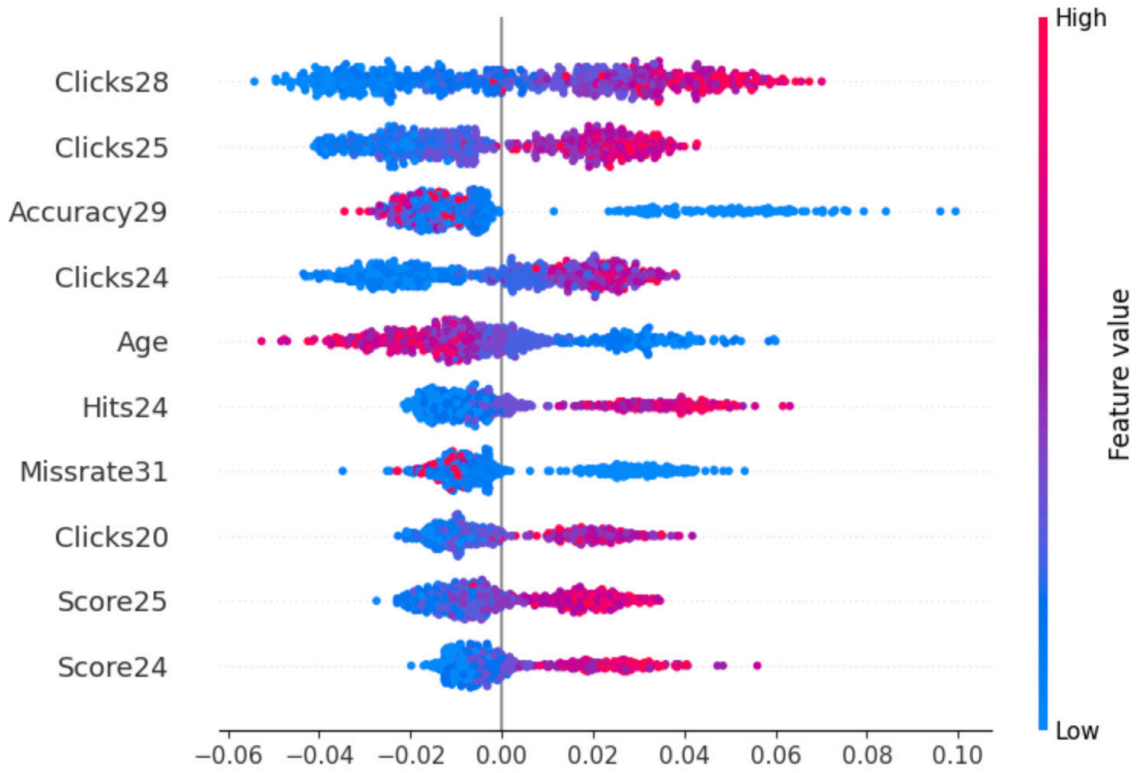


Figure 2: SHAP result for Random Forest

Table 2: Key features of prediction of Dyslexia and Non-dyslexia

Machine Learning Model	XAI Techniques Used	Key Features Influencing Dyslexia Prediction (Increase Probability)	Key Features Influencing Non-Dyslexia Prediction (Decrease Probability)
Random Forest	LIME, SHAP	Higher Clicks24 (Morphological Awareness), Higher Clicks25 (Syntactic Awareness), Clicks28 (Letter/Syllable Rearrangement), Clicks25	Q24 (Morphological Awareness), Q25 (Syntactic Awareness), Q31 (Dictation Task), Accuracy29 (Sentence Formation)
XGBoost	LIME, SHAP	Misses12 (Lexical Awareness, Auditory Memory) (absence of errors led to non-dyslexia), Clicks24 (Morphological Awareness), Clicks25 (Syntactic Awareness), Clicks28 (Letter/Syllable Rearrangement), Age (younger)	Accuracy29 (Sentence Formation), Accuracy30, Accuracy32, Age (older)
k-NN	LIME, PDP	Hits29 (Sentence Formation Skills), Misses31 (Lexical, Orthographic Awareness, and Auditory Memory), Hits6 (Syllable Recognition), High error rates in phonological awareness tasks	Strong visual and auditory working memory
LightGBM	LIME, SHAP	Missrate8 (Phonological Awareness Errors), Accuracy12 (Phonological Awareness) (errors), Missrate31 (Orthographic Awareness), Missrate32	
SVM	LIME, SHAP	Low Score14 (Visual Discrimination and Executive Functions), Score10 (Phonetic and Orthographic Discrimination), Age (younger), Native language effects	High Score29 (Sentence Formation Skills)
Naïve Bayes	LIME, PDP	Syntactic awareness (Q25) (poor performance), Phonetic spelling accuracy (Q16) (poor performance), Alphabetic awareness (Q4) (poor performance)	Syntactic awareness (Q25) (strong performance), Phonetic spelling accuracy (Q16) (strong performance), Alphabetic awareness (Q4) (strong performance)
Neural Network	LIME, PDP	Missrate26 (Phonological and Orthographic Awareness), Age (older)	

Discussion and Conclusion

This study demonstrated the effectiveness of machine learning in dyslexia detection, with Random Forest emerging as the most accurate and reliable model. The model's high recall (98%) indicates its strong ability to identify individuals with dyslexia, making it a valuable tool for early detection. In contrast, models such as SVM and k-NN showed moderate to weak performance, highlighting the importance of selecting robust algorithms for dyslexia classification. These findings support the idea that machine learning, when applied correctly, can significantly improve the accuracy and efficiency of dyslexia screening.

Beyond model performance, the explainable artificial intelligence (XAI) techniques used in this study provided key insights into how these models make decisions. LIME and SHAP analyses revealed that Q24-Q32 questions played a central role in classification, emphasizing the significance of morphological awareness (Q24), syntactic awareness (Q25), letter and syllable rearrangement (Q27-Q28), and word separation (Q29) in dyslexia diagnosis.

These linguistic features were consistently ranked as important across different models, reinforcing their relevance in identifying dyslexic individuals. By using XAI, we were able to not only validate model decisions but also enhance the interpretability of machine learning in educational and clinical settings.

In another study [22], the same dataset was used, and the Random Forest model was also the best-performing model there. For balancing the Dyslexia and non-dyslexia classes, they used weighted attributes. In contrast, this study used the SMOTE technique for balancing. In addition to this study, XAI techniques were applied in our work, and it was determined which variables play a more significant role in dyslexia diagnosis.

The findings of this study contribute to dyslexia diagnosis by providing a data-driven approach to screening. The identification of critical linguistic features can assist in developing more transparent, explainable, and efficient diagnostic tools. Additionally, understanding the decision-making processes of models can help researchers and practitioners refine assessment methods, leading to better-targeted interventions.

Future Directions

While the study successfully demonstrated the potential of machine learning and XAI in dyslexia detection, several areas require further exploration. Future research should focus on enhancing model fairness and generalizability by testing on more diverse datasets, integrating additional cognitive and behavioral markers to refine dyslexia screening and developing more interpretable AI models that can be easily understood by educators and clinicians. By advancing these areas, dyslexia screening can become more accurate, accessible, and interpretable, ultimately benefiting both individuals with dyslexia and the professionals working to support them.

Declaration of Interest: The authors declare that there is no conflict of interest.

Acknowledgements: This paper presents research conducted for the MSc thesis of Şule Akça under the supervision of Tamer Uçar at Bahçeşehir University.

Author Contributions: Şule Akça: Conceptualization, Methodology, Software, Investigation, Writing – Original Draft. Tamer Uçar: Conceptualization, Methodology, Supervision, Writing – Review & Editing

REFERENCES

[1] Kumar, Yogesh, et al. "Artificial intelligence in disease diagnosis: a systematic literature review, synthesizing framework and future research agenda." *Journal of ambient intelligence and humanized computing* 14.7 (2023): 8459-8486.

[2] Secinaro, Silvana, et al. "The role of artificial intelligence in healthcare: a structured literature review." *BMC medical informatics and decision making* 21 (2021): 1-23.

[3] Alqahtani, Norah Dhafer, Bander Alzahrani, and Muhammad Sher Ramzan. "Deep learning applications for dyslexia prediction." *Applied Sciences* 13.5 (2023): 2804.

[4] Kaisar, Shahriar. "Developmental dyslexia detection using machine learning techniques: A survey." *ICT Express* 6.3 (2020): 181-184.

[5] Robaa, Mahmoud, et al. "Explainable AI in Handwriting Detection for Dyslexia Using Transfer Learning." *arXiv preprint arXiv:2410.19821* (2024).

[6] Das, Arun, and Paul Rad. "Opportunities and challenges in explainable artificial intelligence (xai): A survey." *arXiv preprint arXiv:2006.11371* (2020).

[7] Kothapalli, Pavan Kumar Varma, Cheepurupalli Raghuram, and Boddu LV Siva Rama Krishna. "Enhancing Dyslexia Detection and Intervention through Deep Learning: A Comprehensive Review and Future Directions." *Algorithms in Advanced Artificial Intelligence* (2024): 249-256.

[8] Kumar, Anuj, et al. "Explainable Artificial Intelligence (XAI) for Healthcare: Enhancing Transparency and Trust." *International Conference on Cognitive Computing and Cyber Physical Systems*. Springer, Singapore, 2025.

[9] Dudek, Grzegorz. "A comprehensive study of random forest for short-term load forecasting." *Energies* 15.20 (2022): 7547.

[10] Bentéjac, Candice, Anna Csörgő, and Gonzalo Martínez-Muñoz. "A comparative analysis of gradient boosting algorithms." *Artificial Intelligence Review* 54 (2021): 1937-1967.

[11] Guenther, Nick, and Matthias Schonlau. "Support vector machines." *The Stata Journal* 16.4 (2016): 917-937.

[12] Wickramasinghe, Indika, and Harsha Kalutarage. "Naive Bayes: applications, variations and vulnerabilities: a review of literature with code snippets for implementation." *Soft Computing* 25.3 (2021): 2277-2293.

[13] Parashar, Parul. "Neural networks in machine learning." *International Journal of Computer Applications* 105.14 (2014).

[14] Zardini, Enrico, Enrico Blanzieri, and Davide Pastorello. "A quantum k-nearest neighbors algorithm based on the Euclidean distance estimation." *Quantum Machine Intelligence* 6.1 (2024): 23.

[15] Tuan, Dang Anh. "Bridging the Gap Between Black Box AI and Clinical Practice: Advancing Explainable AI for Trust, Ethics, and Personalized Healthcare Diagnostics." (2024).

[16] An, Junkang, Yiwan Zhang, and Inwhee Joe. "Specific-input lime explanations for tabular data based on deep learning models." *Applied Sciences* 13.15 (2023): 8782.

[17] Aromsuk, Tinnarat, Siriwut Kongnakorn, and Wasinee Noonpakdee. "Classification covid-19 data using XGBoost and SHAP (shapley additive exPlainations)." *2023 8th International Conference on Business and Industrial Research (ICBIR)*. IEEE, 2023.

[18] Gebreyesus, Yibrah, et al. "AI for Automating Data Center Operations: Model Explainability in the Data Centre Context Using Shapley Additive Explanations (SHAP)." *Electronics* 13.9 (2024): 1628.

[19] Greenwell, Brandon M. "pdp: An R package for constructing partial dependence plots." (2017).

[20] Kavya, Ramisetty, et al. "Machine learning and XAI approaches for allergy diagnosis." *Biomedical Signal Processing and Control* 69 (2021): 102681.

[21] Gaikwad, Aaishwarya Ashish, and Roshani Raut. "Computer-Aided Diagnosis of Dyslexia Using Machine Learning." *International Conference on Advances in Data-driven Computing and Intelligent Systems*. Singapore: Springer Nature Singapore, 2023.

[22] Rello, Luz, et al. "Predicting risk of dyslexia with an online gamified test." *Plos one* 15.12 (2020): e0241687.

A DYNAMIC TRAJECTORY OPTIMIZATION MODEL PROPOSAL IN THE CONTEXT OF INTELLIGENT AUTOMATION

Muhammet Raşit Cesur

Endüstri Mühendisliği, Mühendislik ve Doğa Bilimleri, İstanbul Medeniyet Üniversitesi, Türkiye

Elif Cesur*

Endüstri Mühendisliği, Mühendislik ve Doğa Bilimleri, İstanbul Medeniyet Üniversitesi, Türkiye

**(Corresponding author) elif.karakaya@medeniyet.edu.tr*

ABSTRACT

This study focuses on robotic trajectory optimization for a four-axis page-turning system in a robotic book scanner operating along Cartesian coordinates. The optimization process is centered on minimizing both page-turning time and total energy consumption. As a result, for a book measuring 16x24 cm, the system reduced the scanned area from 826 cm² to 727 cm², which can now be scanned in just 1.13 seconds, wheras it was previously scanned in 1.5 seconds. To enable the developed model to adapt to dynamic environmental conditions, physical parameters affecting energy consumption, such as motor temperature, speed, and displacement were examined. It was observed that under room temperature, the motor reached a maximum temperature of 42°C. A regression model was developed to analyze the relationship between displacement, motor speed, and current consumption, and this relationship was further evaluated using an ANOVA test. The results indicated that motor speed and displacement are significantly correlated with current consumption. However, motor temperature, with a t-value low relative to the degrees of freedom and a p-value well above the 5% significance threshold, showed no significant effect on current consumption. This indicates that even if the motor heats up, its performance remains stable and is not adversely affected.

Keywords: Robotic trajectory optimization, energy consumption, sustainability.

Introduction

Robotics offer numerous opportunities to advance the Sustainable Development Goals (SDGs), including improvements in ecosystem monitoring infrastructure, disaster response, healthcare, responsible production and consumption, and even the promotion of gender equality [1]. Although robotics offers advantages in achieving the SDGs, the sustainability of robots is critically important for the efficiency of this process. To this end, it becomes essential for robots to perform the maximum amount of work with minimal energy consumption and maintenance. Since collaborative robots (cobots) developed for Industry 4.0 environments make dynamic decisions, sustainability constraints must also be integrated into this process dynamically. This study aims to develop an energy-efficient robotic trajectory optimization model that operates within defined cycle time constraints.

Literature Review

Energy-driven optimization models for robots vary in terms of the specific problems addressed, model complexity, and the algorithms employed. Optimization process is performed for single robot, cyclic multi-robot [2] or autonomous mobile robots [3], targeting various trajectory shapes in a 3D environment [4]. Both single-objective models (e.g., minimizing cycle time or energy) and multi-objective models (incorporating cycle time, energy, stability, deformation, adaptivity, and throughput) [5-7] are solved while accounting for kinodynamic constraints, collision avoidance, robot configuration changes, acceleration limit, torque limit and cyclic task constraints [6], [8]. Energy consumption is typically predicted using intelligent models such as Long Short-Term Memory (LSTM) networks [9], [10] and Artificial Neural Networks (ANNs) . The models are optimized using various algorithms, including Genetic Algorithm (GA) [11], Ant Colony Optimization [12], Model Predictive Control (MPC)-based methods [13], Particle Swarm Optimization (PSO) [3], the Improved Seagull Optimization Algorithm (iSOA), and the Modified Douglas–Peucker (mDP) algorithm [14].

Method And Case Study

This paper presents the optimization of a page-flipper designed for a robotic book scanner. The page-flipping mechanism performs both linear and circular motions to turn a page, as shown in Figure 1. Trajectory optimization was carried out by modeling the joints responsible for these independent movements. The page-fliper can execute linear and circular motions either sequentially or in parallel. It is assumed that, in order to reliably flip a book page measuring 16×24 cm, the page-flipper must cover an area of at least 718 cm².

Figure 1: Linear and circular displacement of page-flipper

Stepper motors were tested over a 24-hour period at ambient temperatures of 10°C and 25°C by executing combinations of displacements ranging from 10 cm to 300 cm and feed rates between 125 mm/min and 1000 mm/min. The resulting temperature variations under these conditions are shown in Figure 2. Additionally, the motors were operated at -20°C for one hour, demonstrating similar performance characteristics to those observed at higher temperatures.

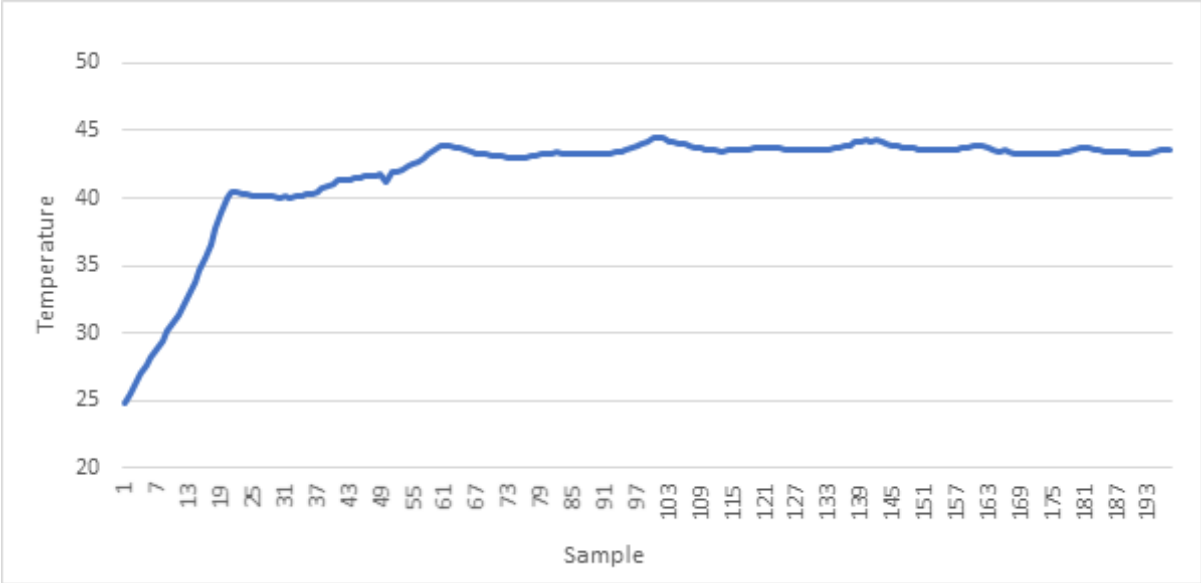


Figure 2: Temperature variation of the stepper motor during operation.

When the collected parameters, such as feed rate, displacement, and motor temperature, were correlated with energy consumption using a multi-parameter regression model, the resulting ANOVA test is presented in Table 1. The model achieved an R^2 value of 0.9146 and a Mean Absolute Deviation (MAD) of 0.1464. According to the ANOVA results, energy consumption is significantly influenced by feed rate and displacement, while stepper motor temperature does not show a statistically significant effect. This conclusion is based on the low t-value at 114 degrees of freedom and a p-value greater than 5%.

Table 1. Results of ANOVA

	Coefficients	Standard Error	t-Value	p-Value	Lower 95.0%	Upper 95.0%
Constant	6.76519	0.238757	28.33507	1.24E-52	6.292078	7.238303
Feed Rate	0.001461	4.28E-05	34.13252	1.09E-60	0.001376	0.001546
Displacement	0.000695	0.00016	4.354828	2.98E-05	0.000379	0.001011
Temperature	0.003852	0.005524	0.69734	0.487048	-0.00709	0.014797

The optimization model and its constraints are presented at Equations 1 – 3. The objective function aims to minimize total consumed energy (C_e). The cycle time (T_c) of the system must between 0 and 2 seconds and total area coverege (A_c) must be bigger than 712 cm^2. The energy consumption model developed by Taphasanoğlu, Cesur and Cesur [15] is employed as the objective function in the proposed optimization framework. The model is solved using a Genetic Algorithm (GA) with an iteration limit of 1000.

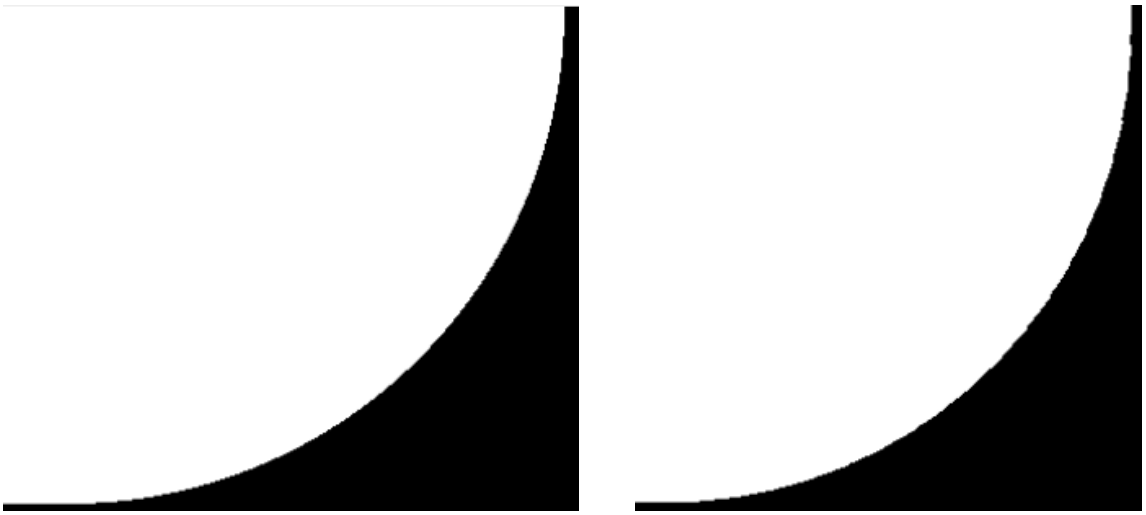
$$\text{MINIMIZE } C_e \quad (1)$$

$$0 \leq T_c \leq 2 \text{ sec} \quad (2)$$

$$A_c \geq 712 \text{ cm}^2 \quad (3)$$

Discussion and Conclusion

Robotic systems offer numerous advantages in achieving the SDGs. Therefore, ensuring the sustainability of robotic systems has become increasingly important for efficiently reaching these goals. To achieve sustainability in robotic systems, energy consumption and system wear must be minimized. A robotic trajectory optimization model designed to achieve this has been developed for a robotic page-flipper. The proposed model performs trajectory optimization to scan an area of 727 cm² in 1.13 seconds, in accordance with the constraints required to reliably flip a book page measuring 16×24 cm. The area scanned by the page-flipper during the page-flipping motion before optimization is shown in Figure 3a, and the area scanned after optimization is shown in Figure 3b.



a. The area scanned by the page-flipper system before optimization

b. The area scanned by the page-flipper system after optimization

Figure 3:

REFERENCES

[1] V. Mai et al., "The Role of Robotics in Achieving the United Nations Sustainable Development Goals - The Experts' Meeting at the 2021 IEEE/RSJ IROS Workshop [Industry Activities]," IEEE Robot Autom Mag, vol. 29, no. 1, pp. 92-98+107, 2022, doi: 10.1109/MRA.2022.3143409.

[2] E. Glorieux, S. Riazi, and B. Lennartson, "Productivity/energy optimisation of trajectories and coordination for cyclic multi-robot systems," Robot Comput Integr Manuf, vol. 49, pp. 152–161, Feb. 2018, doi: 10.1016/J.RCIM.2017.06.012.

[3] P. B. Fernandes, R. C. L. Oliveira, and J. V. Fonseca Neto, "Trajectory planning of autonomous mobile robots applying a particle swarm optimization algorithm with peaks of diversity," Appl Soft Comput, vol. 116, p. 108108, Feb. 2022, doi: 10.1016/J.ASOC.2021.108108.

[4] M. Sarkar, S. Nandy, and S. N. Shome, "Energy Efficient Trajectory Tracking Controller for Underwater Applications: ARobust Approach," Aquat Procedia, vol. 4, pp. 571–578, Jan. 2015, doi: 10.1016/J.AQPRO.2015.02.074.

[5] H. Feng et al., "Multi-objective time-energy-impact optimization for robotic excavator trajectory planning," Autom Constr, vol. 156, p. 105094, Dec. 2023, doi: 10.1016/J.AUTCON.2023.105094.

[6] S. Liu, Y. Wang, X. V. Wang, and L. Wang, "Energy-efficient trajectory planning for an industrial robot using a multi-objective optimisation approach," Procedia Manuf, vol. 25, pp. 517–525, Jan. 2018, doi: 10.1016/J.PROMFG.2018.06.122.

[7] Z. Wang, Y. Li, P. Sun, Y. Luo, B. Chen, and W. Zhu, "A multi-objective approach for the trajectory planning of a 7-DOF serial-parallel hybrid humanoid arm," Mech Mach Theory, vol. 165, p. 104423, Nov. 2021, doi: 10.1016/J.MECHMACHTHEORY.2021.104423.

[8] A. Wachter, C. Hartl-Nesic, and A. Kugi, "Robot Base Placement Optimization for Pick-and-Place Sequences in Industrial Environments," IFAC-PapersOnLine, vol. 58, no. 19, pp. 19–24, Jan. 2024, doi: 10.1016/J.IFACOL.2024.09.080.

[9] P. Jiang, Z. Wang, X. Li, X. V. Wang, B. Yang, and J. Zheng, "Energy consumption prediction and optimization of industrial robots based on LSTM," J Manuf Syst, vol. 70, pp. 137–148, Oct. 2023, doi: 10.1016/J.JMSY.2023.07.009.

[10] Z. Wang, X. Li, P. Jiang, X. V. Wang, and H. Yuan, "Energy consumption modeling based on operation mechanisms of industrial robots," Robot Comput Integr Manuf, vol. 94, p. 102971, Aug. 2025, doi: 10.1016/J.RCIM.2025.102971.

[11] M. Zhang and J. Yan, "A data-driven method for optimizing the energy consumption of industrial robots," J Clean Prod, vol. 285, p. 124862, Feb. 2021, doi: 10.1016/J.JCLEPRO.2020.124862.

[12] Z. Dong, X. Zhang, W. Yang, M. Lei, C. Zhang, and J. Wan, "Ant colony optimization-based method for energy-efficient cutting trajectory planning in axial robotic roadheader," Appl Soft Comput, vol. 163, p. 111965, Sep. 2024, doi: 10.1016/J.ASOC.2024.111965.

[13] K. Wu, G. Zheng, and J. Zhang, "FEM-based trajectory tracking control of a soft trunk robot," Rob Auton Syst, vol. 150, p. 103961, Apr. 2022, doi: 10.1016/J.ROBOT.2021.103961.

[14] T. Lei, T. Sellers, C. Luo, D. W. Carruth, and Z. Bi, "Graph-based robot optimal path planning with bio-inspired algorithms," Biomimetic Intelligence and Robotics, vol. 3, no. 3, p. 100119, Sep. 2023, doi: 10.1016/J.BIROB.2023.100119.

[15] S. Taphasano, ˘ Glu, M. Ra¸sit, R. Cesur, and E. Cesur, "A Precise Energy Consumption Model for Computer Numerical Control Machines: A Hybrid Approach," 2024, doi: 10.3390/su162310659.

ANALYZING CURRICULUM EVOLUTION IN A MORPHOLOGICALLY RICH LANGUAGE USING UNSUPERVISED NLP: A CASE STUDY IN TURKISH MATHEMATICS EDUCATION

Batın Örene*
University of London, Department of Computer Science, London, United Kingdom
Ghislaine Le Grange
University of London, Department of Computer Science, London, United Kingdom
(Corresponding author) batinorene@gmail.com

ABSTRACT

Traditional methods for curriculum analysis often struggle with scalability and subjectivity. This study introduces a computational framework built in Python that utilizes unsupervised natural language processing (NLP) techniques to analyze curriculum evolution in a morphologically rich language. Focusing on Turkish mathematics education, we compare two curricula—one from 2018 (302 learning objectives) and one from 2024 (801 learning objectives). Our approach integrates topic modeling to identify thematic shifts and cognitive complexity assessments based on Bloom's taxonomy to evaluate changes in pedagogical emphasis. The results revealed a profound reorganization of themes, as evidenced by a low cross-curriculum similarity of 0.04, indicating the emergence of new curricular priorities. Furthermore, cognitive complexity analysis based on Bloom's taxonomy demonstrates a notable increase in higher-order "Create" objectives (+19.3%) alongside a decline in "Evaluate" objectives (–24.4%). To validate these findings, we propose a multi-tiered validation framework that triangulates results across various analytical approaches. Overall, our study not only provides data-driven insights into curricular evolution but also establishes a replicable methodology for computational curriculum analysis in morphologically rich languages.

Keywords: Curriculum evolution, unsupervised NLP, morphologically rich languages, turkish mathematics education, topic modeling, cognitive complexity, bloom's taxonomy.

Introduction

Curriculum documents represent the formal educational intentions of educational systems and serve as the blueprint for teaching and learning processes. Traditional curriculum analysis typically relies on manual methods such as content analysis, expert reviews, qualitative coding [1], and established frameworks like the Surveys of Enacted Curriculum (SEC) [2] or TIMSS assessments [3]. While these approaches offer valuable insights into curriculum alignment and implementation contexts, they often face limitations in terms of scalability, reproducibility, and potential subjective biases. These challenges become more pronounced when examining large-scale curriculums or tracking changes over time.

Mathematics education faces unique challenges in curriculum analysis due to the specialized terminology, symbolic notations, and cognitive progression patterns specific to the discipline. In the Turkish context, these challenges are compounded by the language's agglutinative morphology and rich inflectional system, making automated analysis particularly difficult [4]. Effectively processing Turkish requires specialized tools and techniques to handle complexities like morphological analysis and parsing [5], [6], [7]. Despite growing interest in applying computational methods to educational research and advancements in NLP techniques, there remains a significant gap in the literature regarding NLP applications to mathematics curriculum analysis, especially within the challenging Turkish language context. Furthermore, validation methods for unsupervised NLP findings in this specific domain are currently underdeveloped, limiting the confident application of these powerful techniques.

Our study addresses the question:

- How can unsupervised NLP techniques reveal substantive curricular shifts in Turkish mathematics education between 2018 and 2024, and what frameworks can validate these findings in the absence of established baselines?

This research is particularly timely given the significant reform introduced with the 2024 curriculum, known as the "Türkiye Yüzyılı Maarif Modeli"[8] (The Turkish Century Education Model). This reform substantially increased the number of learning objectives from 302 to 801, reversing previous simplification trends observed in comparisons between earlier curricula like 2005, 2013, and 2018 [9], [10]. By using specialized Turkish NLP preprocessing techniques tailored to the language's structure [4], [6], [7], established unsupervised methods like topic modeling (often using Latent Dirichlet Allocation - LDA) [11], [12], and cognitive complexity analysis based on established frameworks like Bloom's taxonomy [13], [14], [15], our approach provides a multi-dimensional view of curricular evolution. It aims to establish a framework for validating such computational analyses, addressing the previously noted gap in methodology for this specific context.

Related Work

Curriculum Analysis Methods in Mathematics Education

Traditional approaches to curriculum analysis have relied on manual content analysis, expert reviews, and frameworks such as the Surveys of Enacted Curriculum (SEC)[2] and TIMSS. These methods offer in-depth insights into curriculum alignment and the interplay between standards and assessments. TIMSS not only provides an international benchmark but also provides rich contextual data on teaching practices, classroom environments and student attitudes. The TIMSS 2023 [3] report for Türkiye indicated that students who perceived mathematics instruction as highly understandable achieved higher average scores, underscoring the importance of instructional clarity. Additionally, frequent disruptive behaviors negatively correlated with achievement, highlighting the critical role of classroom management in curriculum implementation.

Curricula evolve through educational reforms and shifting priorities. Over the years, Turkish mathematics curricula have historically adapted to changing educational aims, contemporary trends, and technological advancements [16]. Comparative analyses between

2005 and 2013 primary school math curricula revealed systematic simplification, with a clear reduction in objectives [9]. Similarly, the 2018 mathematics curriculum further reduced the number of objectives compared to 2013 [10]. However, the recent reform The Turkish Century Education Model has reversed this simplification trend, emphasizing expansive changes that prioritize higher-order thinking skills [8].

Researchers typically employ both qualitative and quantitative approaches to analyze curriculum evolution. Qualitative methods in Turkish studies have examined curricular goals, scope, pedagogical orientation, and evolutionary aspects, highlighting themes such as STEM integration and student-centered approaches [1]. Quantitatively, structural modeling techniques such as topic modeling, notably Latent Dirichlet Allocation (LDA), are increasingly utilized. However, research applying LDA to curricula has noted occasional mismatches between intended curriculum topics and actual content [11].

Topic modeling has emerged as a powerful unsupervised method to analyze large educational documents, providing insights into thematic patterns and content evolution. Applications include mapping course topics to institutional competencies in medical school curricula [12]. and summarizing K-12 learning outcomes into thematic clusters to identify curriculum gaps or overlaps. Determining the optimal number of topics is crucial, with topic coherence metrics such as c_v frequently guiding this decision. For example, in one study, an 80-topic LDA model achieved the highest coherence score, effectively balancing model complexity and thematic clarity [12]. However, since unsupervised models rely solely on pattern recognition, human validation remains essential for assigning pedagogically meaningful interpretations.

Challenges of NLP in Morphologically Rich Languages

Turkish presents specific challenges for NLP applications due to its agglutinative structure, causing significant vocabulary sparsity and tokenization complexity. For example, the Turkish word “gidebiliyoruz” consists of multiple morphemes (gid-ebil-iyor-uz), illustrating how extensive word forms complicate NLP tasks like POS tagging and parsing [4]. Turkish NLP tools, such as Zemberek for morphological analysis and transformer-based models like BERTurk, have been developed to address these challenges but still lag behind their English counterparts in accuracy [5], [6], [7]. Previous research emphasizes the necessity of morphological preprocessing like lemmatization and specialized stopword removal to enhance topic modeling effectiveness in Turkish texts, particularly with specialized mathematical terminology [6], [7].

Despite advancements in NLP techniques and the particular linguistic challenges posed by Turkish, applying these approaches to curriculum analysis remains limited. Given recent substantial reforms in Turkish mathematics education, the gap in NLP-based curriculum analysis methodologies is particularly notable. The validation methods for unsupervised NLP findings within curriculum analysis are currently underdeveloped due to a lack of standardized baselines. This research addresses these gaps by proposing a novel NLP-based methodology specifically tailored to Turkish mathematics curricula and establishing a robust validation framework.

Cognitive Complexity and Bloom's Taxonomy in NLP

Bloom's Taxonomy is a cognitive complexity level of cognitive skills that consists of 6 levels in its updated version: Remember, Understand, Apply, Analyze, Evaluate, Create [13], [14]. Classifying content according to these levels involves looking at the action verb in an outcome, as Bloom's taxonomy provides well-known verb lists for each cognitive level. These verb mappings provide rules for manual classification. Turkish curriculum studies often manually code each outcome's verb phrase to a Bloom category to analyze the cognitive demand distribution [15]. A recent comparison of Turkey's and Germany's math curricula classified every learning outcome by both cognitive process dimension and knowledge type (factual, conceptual, procedural and metacognitive) using subject matter experts in the process [15]. That analysis revealed that Turkey's middle school algebra outcomes focus on the Analyze level, whereas its German counterpart included more Evaluate and Create level objectives. Verb-based taxonomic analysis is essential to understand how rigorous the curricula are.

Automating Bloom's level classification is challenging, especially in Turkish. A research study mentioning dictionary mapping cue words shows achieving 75% accuracy in recall questions, but lower accuracy in other levels [17]. In that research, when compared with other classification models, BERT performed higher compared to binary and multi-class multi-label classifiers such as support vector machine, naïve bayes, logistic regression, and random forests. A Turkish rule-based approach may flag any outcome containing "çıkar" as analysis level, but many verbs appear in multiple levels depending on the context(e.g. "çıkar" can mean to deduct in the remembering level, or infer in the Analysis level). This overlap causes low precision. In research on autonomous Bloom's taxonomy classification based on keywords, 55% accuracy was achieved [18]. On the other hand, research using supervised learning and BERT models achieved 0.93 inter-annotator agreement and 0.95 F-1 score [17], achieving human-level performance.

Another study, using an LSTM model with pre-trained word embeddings, classified outcomes with 87% accuracy [18]. While these efforts are mostly in English, it is evident that with sufficient training data and methodology, NLP models can accurately classify cognitive-level learning objectives. Another approach is using a semi-supervised approach, using the known Turkish Bloom verb list as a starting point to label a set of outcomes and train a model to generalize them. There are a few BERT models trained for English language Bloom's taxonomy classification that have high accuracy (97%) [19], [20], but they lack Turkish language support. A POS tagger or parser must be able to deal with morphologically rich language to correctly identify action verbs. Other than morphological classification, some studies go further and classify learning outcome targets as conceptual understanding, procedural skill, creative thinking or critical thinking, which require expert review and require nuanced modeling techniques [17].

Data and Methodology

All analyses were implemented using Python 3.9. Our computational framework integrates widely-used open-source libraries such as spaCy, networkx, scikit-learn.

Dataset Description

Our analysis focuses on two versions of the Turkish national mathematics curriculum: the 2018 curriculum [21] and the 2024 curriculum [22]. These documents represent official educational policy issued by the Turkish Ministry of National Education and define the learning objectives, content coverage, and pedagogical approaches for mathematics education across all grade levels.

The dataset exhibits significant differences in size and structure between the two curriculum versions:

- The 2018 curriculum

• Contains 302 learning objectives

• Comprises 15,732 words arranged in 1,937 sentences

- The 2024 curriculum

• Contains 801 learning objectives

• Comprises 97,735 words arranged in 6,114 sentences.

The 2024 curriculum exhibits an increase of 82,003 words and 499 additional objectives. This expansion not only reflects a higher level of curricular specificity but also suggests a substantial shift in educational priorities. Moreover, the average word length increased from 4.63 to 5.89 characters and the average sentence length from 6.29 to 13.65 words, indicating enhanced linguistic complexity and detail in the newer curriculum.

Preprocessing Pipeline

Given the challenges of processing Turkish mathematical content, we developed a specialized preprocessing pipeline:

1. PDF Extraction: We employed PDFMiner and PyPDF2 with custom parameters to preserve mathematical notation and Turkish special characters.

2. Turkish-specific Preprocessing: Using the spaCy `tr_core_news_trf` transformer-based model that has high accuracy in morphological analysis (0.915), POS tagging (0.909), and lemmatization (0.878) [23].

3. Mathematical Notation Normalization: We standardized mathematical notation (function notation, symbols, and operators) to ensure consistent tokenization and processing rather than to analyze symbolic relationships. This normalization focused on maintaining textual consistency for subsequent text-based analyses rather than mathematical semantic analysis.

4. Lemmatization and Stopword Removal: Our custom process addressed Turkish morphology while preserving domain-specific terminology.

Analysis Methods

Our analytical approach combined multiple NLP techniques to develop a comprehensive understanding of curricular shifts:

- **Topic Modeling Implementation**

We employed Latent Dirichlet Allocation (LDA) using Gensim, a common technique for analyzing curriculum structure [11], [12], to identify latent topics in each curriculum. The optimal number of topics was determined using C_v coherence scores, a standard evaluation metric in topic modeling applications [12], resulting in 5 topics for the 2018 curriculum and 8 for the 2024 curriculum. Key implementation steps included document preprocessing, term-document matrix construction, LDA model building with optimized parameters, and coherence evaluation.

- **Topic Coherence Metrics and Selection**

We employed multiple topic coherence metrics to validate our topic models, primarily utilizing the C_v coherence score, which assesses semantic similarity within topics using external word embeddings, similar to approaches used in other large-scale curriculum analyses [12]. The optimal topic models achieved coherence scores of 0.5335 for the 2018 curriculum and 0.4235 for the 2024 curriculum. We also calculated topic stability by running multiple iterations with different random seeds and measuring the consistency of resulting topic structures using Jensen-Shannon divergence.

- **Cognitive Complexity Analysis Using Bloom's Taxonomy**

We developed a computational approach to assess the cognitive complexity of learning objectives based on the revised Bloom's taxonomy [13], [14], focusing on the six cognitive levels: Remember, Understand, Apply, Analyze, Evaluate, and Create. For each level, we compiled comprehensive lists of Turkish action verbs that correspond to the cognitive operations in Bloom's framework. While Turkish curriculum studies often manually code outcomes to Bloom categories (e.g., [15]), and automated approaches using NLP face challenges due to linguistic ambiguity and context dependence, often requiring sophisticated models like BERT [17] or LSTM [18] with varying success, our rule-based approach utilizes comprehensive verb lists derived from the taxonomy's principles [13], [14].

- **Word Frequency and Co-occurrence Networks**

We analyzed changes in terminology by calculating relative frequency differences between curricula. After lemmatization and stopword removal [4], [6], we computed term frequencies within each curriculum and normalized these to account for differences in document length. The relative frequency differences between curricula highlighted terms with the largest increases and decreases, providing insights into shifting curricular priorities. Additionally, we constructed word co-occurrence networks to visualize relationships between key mathematical concepts in each curriculum. These networks were created by identifying terms that frequently appeared together within the same learning objectives, with edge weights reflecting co-occurrence strength. This network analysis revealed changes in conceptual relationships between the two curricula.

• **Mathematical Domain Classification**

We developed a classification approach to categorize curriculum objectives into mathematical domains based on domain-specific terminology rather than symbolic analysis. For each mathematical domain (algebra, geometry, calculus, statistics, number theory, and logic), we compiled comprehensive lists of domain-specific terms in Turkish, informed by standard domain definitions and previous curriculum comparisons like [15]. Learning objectives were then classified into domains based on the presence of these terms, with objectives potentially belonging to multiple domains when they integrated concepts across mathematical areas.

Rather than attempting to parse or interpret mathematical notation symbolically, our approach focused on the linguistic markers that signal different mathematical domains. This term-based classification allowed us to track shifts in domain coverage between the 2018 and 2024 curricula, identifying areas of expanded or reduced emphasis and providing insights into changes in mathematical content priorities, similar in goal to the comparative analysis performed [15] but using computational linguistic methods.

Validation Framework

We utilized a three-tiered validation framework to address the challenge of validating unsupervised NLP findings without established benchmarks, a noted gap in applying NLP to curriculum analysis, particularly in morphologically rich languages like Turkish [4].

1. Internal Validation: We employed multiple coherence metrics (C_v and NPMI), common in topic modeling evaluation [12], and topic stability analysis using Jensen-Shannon divergence across multiple runs.

2. Methodological Triangulation: Findings were cross-validated across different analytical approaches, ensuring consistency between topic modeling [11], [12], frequency analysis, cognitive complexity assessment (using Bloom's [13], [14]), and domain classification (informed by prior work like [15]).

3. Theoretical Validation: Findings were cross-validated across different analytical approaches, ensuring consistency between topic modeling [11], [12], frequency analysis, cognitive complexity assessment (using Bloom's [13], [14]), and domain classification (informed by prior work like [15]).

Alignment Between Topic Modeling and Frequency Analysis

We validated topic modeling results by comparing them with direct word frequency analysis, confirming that key terms identified in topics also showed significant frequency differences between curricula:

• Terms identified as characteristic of new topics in the 2024 curriculum (e.g., "belirle", "ilişkin", "varsayım") also appeared among the terms with the largest frequency increases.

• Terms associated with topics more prominent in the 2018 curriculum (e.g., "ver", "yer", "iletişim") also appeared among terms with the largest frequency decreases.

Consistency Between Different Analytical Lenses

We examined the consistency of findings across different analytical approaches:

- Cognitive complexity analysis through verb usage patterns

- Mathematical domain classification

- Topic modeling

- Term frequency analysis

The convergence of findings across these different lenses provided strong evidence for the validity of our identified curricular shifts.

Results

Topic Modeling

Our topic modeling approach uncovered a marked reorganization of themes between the 2018 and 2024 curricula. The optimal topic solution identified five distinct topics for the 2018 curriculum and eight for the 2024 curriculum, reflecting a greater thematic diversification in the newer version.

Comparison of Topics Between Curricula

The 2024 curriculum topics revealed a restructured emphasis:

1. Topic 0 (11.1%): Function representations (fonksiyon - function, temsil - representation, grafik - graph, ifade - expression, matematiksel - mathematical, referans - reference, nitel - attribute, cebirsel - algebraic, denklem - equation, doğrusal - linear)

2. Topic 1 (9.2%): Algebraic reasoning (cebirsel - algebraic, öner – suggest/propose, nokta - point, sözel - verbal, grafiksel – graphical, uygun – appropriate/suitable, fonksiyon - function, ilişki - relation, belirli - specific, bileşen - component)

3. Topic 2 (6.7%): Coordinate geometry (incele - examine, doğru - line, nokta - point, kavramsal - conceptual, üzer – on/over, limit - limit, eğim - slope, dönüşüm - transformation, ödev - homework, sistem - system)

4. Topic 3 (8.2%): Mathematical logic (yönelik – pertaining to/toward, çıkarım - inference, matematiksel - mathematical, önerme - proposition, ulaş – obtain/reach, araştır – investigate/research, doğrula – verify/validate, istatistiksel - statistical, değerlen – evaluate/assess , ilgili - related)

5. Topic 4 (9.7%): Function properties (tanım - definition, fonksiyon - function, şekil – shape/form, üçgen - triangle, nitel – quality/attribute, değer - value, küme - set, karesel – square, polinom - polynomial, gerçek - real)

6. Topic 5 (9.3%): Real-world problems (gerçek - real, sayı - number, öğrenci - student, problem - problem, sağla – provide/ensure, durum – condition/state, hareket - movement, yaşam – life, içer – include/contain, değişken - variable)

7. Topic 6 (30.9%): Pattern recognition (varsayım – assumption/hypothesis, kullan - use, ilişki - relation, fark - difference, karşıla – compare, örüntü - pattern, Genelle - generalize, dayalı – based on, temsil - representation, belirle - determine)

8. Topic 7 (15.0%): Problem-solving strategies (çözüm - solution, strateji - strategy, belirle - determine, problem - problem, kulan – use, veri - data, öğrenme - learning, araştır -investigate/research, problemin - of the problem, soru - question)

These emerging topics in the 2024 curriculum indicate not only an increased number of themes but also the introduction of new emphases, particularly in mathematical logic (Topic 3) and problem-solving strategies (Topic 7).

The 2018 curriculum topics centered around:

1. Topic 0 (15.2%): Geometric measurement (uzunluk – length, çember – circle, kenar – edge/side, bulun – find, denklem – equation, nokta – point, toplam – sum, alan – area, üçgen – triangle, ölçü – measurement)

2. Topic 1 (22.8%): Technology and functions (iletişim – communication, bilgi – information, teknoloji – technology, fonksiyon – function, grafik – graph, yararlan – benefit/use, çalış – work/operate, doğru – line/correct, yardım – help, parça – part)

3. Topic 2 (20.2%): Numbers and expressions (sayı – number, kenar – edge/side, ifade – expression, kavram – concept, polinom – polynomial, üçgen – triangle, fark – difference, veri – data, oluş – formation, uygulama – application)

4. Topic 3 (29.5%): Real-world applications (fonksiyon – function, gerçek – real, hayat(i) – life, problem – problem, küme – set, nokta – point, üzer – on, grafik – graph, türev – derivative, eksen – axis)

5. Topic 4 (12.3%): Probability and measurement (üçgen – triangle, değer – value, olay – event, ayrık – discrete, kavram – concept, alan – area, çarpan – factor, açık – open, taban – base, hesapla – calculate)

These topics from 2018, characterized by terms related to geometry, technology, and real-world applications, appear to have either shifted in emphasis or been integrated into new thematic structures in the 2024 curriculum.

Topic Similarity Metrics

A cross-curriculum similarity analysis reveals a very low average similarity (approximately 0.04 on a 0–1 scale), with the highest observed similarity being 0.176 (between 2018's Topic 3 and 2024's Topic 4). Several topic pairs exhibit near-zero similarity, highlighting that the 2024 curriculum is not merely an expansion of the 2018 content but rather a substantive reorganization of mathematical themes.

Topic Prevalence Changes

The distribution of learning objectives across topics also shifted significantly. In the 2018 curriculum, the most prevalent topic (Topic 3: "Real-world applications") covered 29.5% of

objectives. In the 2024 curriculum, the most prevalent topic (Topic 6: "Pattern recognition") covered 30.9% of objectives, but the remaining objectives were more evenly distributed across other topics, suggesting a more balanced curricular structure.

The emergence of new topics such as "Mathematical logic" (Topic 3, 8.2%) and "Problem-solving strategies" (Topic 7, 15.0%) in the 2024 curriculum reveals new curricular priorities that were not explicitly emphasized in the 2018 curriculum.

Content Evolution

Our analysis revealed significant shifts in both terminology and mathematical domain coverage between the two curricula.

Mathematical Domain Coverage Shifts

The classification of learning objectives by mathematical domain showed substantial changes in domain emphasis:

The classification of learning objectives by mathematical domain reveals significant changes:

- **Algebra**: Increased from 75 to 195 objectives (+160%)
- **Geometry**: Increased from 90 to 134 objectives (+48%)
- **Calculus**: Increased from 11 to 30 objectives (+172%)
- **Statistics**: Increased from 7 to 19 objectives (+171%)
- **Number Theory**: Increased from 23 to 93 objectives (+304%)
- **Logic**: Increased from 1 to 54 objectives (+5300%)

These shifts suggest a major expansion in algebra, number theory, and particularly logic, indicating an enhanced focus on analytical and deductive reasoning skills.

Terminology Changes

Analysis of term frequency changes revealed significant shifts in mathematical terminology:

- **Increases**: Terms associated with reasoning and problem-solving (e.g., "belirle," "ilişkin," "varsayım," "matematiksel," "çözüm") show significant increases. For instance, "belirle" rises from 0.1148% in 2018 to 1.7089% in 2024.

- **Decreases**: Conversely, terms related to technology integration and general instructional directives (e.g., "ver," "yer," "iletişim," "hayat(i)," "teknoloji") have diminished considerably.

This terminological shift reflects a move from a technology-focused and directive approach toward an emphasis on mathematical reasoning and conceptual understanding.

Cognitive Complexity

Our analysis of cognitive complexity based on Bloom's taxonomy revealed significant shifts in the distribution of learning objectives across cognitive levels.

Table 1: Distribution Changes Across Bloom's Taxonomy Levels

Cognitive Level	2018 Curriculum	2024 Curriculum	Change
Remember	5.5%	8.3%	+2.8%
Understand	5.5%	1.4%	-4.1%
Apply	13.8%	16.2%	+2.3%
Analyze	7.7%	11.9%	+4.1%
Evaluate	37.0%	12.6%	-24.4%
Create	30.4%	49.6%	+19.3%

The data indicates:

A dramatic rise in "Create" level objectives, suggesting an increased emphasis on original thought, design, and argumentation.

A moderate increase in "Analyze" level objectives, reflecting a growing focus on deconstructing problems.

A significant decrease in "Evaluate" level objectives, implying a shift away from judgment-based activities.

Overall, the average cognitive complexity marginally increased from 4.56 to 4.68 (out of 6.0), reinforcing that the 2024 curriculum places slightly higher cognitive demands on students while redistributing focus across the cognitive hierarchy.

Discussion

Our findings demonstrate that the 2024 Turkish mathematics curriculum represents not merely an expansion but a profound reorganization of educational content and priorities. This contrasts sharply with previous simplification trends observed in comparisons between the 2005, 2013, and 2018 curricula[9], [10]. The shift from 5 topics in 2018 to 8 in 2024—with a very low average cross-curriculum topic similarity (0.04)—confirms that the new curriculum is designed to reconceptualize how mathematical knowledge is structured and delivered, reversing the earlier pattern of objective reduction.

Reconceptualizing Curricular Themes

The emergence of new topics—particularly those emphasizing mathematical logic (Topic 3), algebraic reasoning (Topic 1), and problem-solving strategies (Topic 7)—suggests that the 2024 curriculum is shifting toward more ABSTRACT, higher-order cognitive tasks, aligning with the stated goals of the recent reform [8]. This increased focus on higher-order skills potentially addresses findings from earlier comparative studies, such as Kuzu et al. [15], which noted that the previous Turkish middle school algebra curriculum had a stronger focus on 'Analyze' level objectives compared to the German curriculum's inclusion of more 'Evaluate' and 'Create' objectives. The near disappearance of the previously dominant "technology and functions" theme from 2018 implies that technology integration

may now be embedded across all content rather than isolated as a distinct focus. Additionally, the significant increase in pattern recognition (Topic 6) aligns with contemporary views that prioritize the identification of relationships and conceptual patterns over mere procedural computation. These changes reflect an evolving educational philosophy that values metacognitive and creative approaches to learning mathematics, echoing themes like student-centeredness sometimes identified in qualitative Turkish curriculum analyses [1].

Implications for Instruction and Assessment

The dramatic redistribution of cognitive complexity—most notably, the 19.3% increase in "Create" level objectives coupled with a 24.4% decrease in "Evaluate" level objectives—indicates that the 2024 curriculum prioritizes student-generated reasoning and creative problem-solving. This marks a significant departure from the cognitive profile suggested by prior comparative analyses [15] and underscores the reform's emphasis on higher cognitive functions [8]. In practice, this shift will likely require teachers to foster more open-ended, inquiry-based learning environments. Instruction may need to incorporate project-based activities and innovative assessment techniques that capture students' ability to generate and articulate original mathematical ideas, moving beyond traditional evaluative methods that focus on rote or judgment-based tasks, aligning with the need for instructional clarity highlighted by TIMSS findings [3].

Terminological and Domain Shifts as Indicators of Pedagogical Change

Our terminology analysis further supports a reorientation in teaching approaches. The significant increases in terms associated with reasoning and problem-solving (e.g., "belirle," "varsayım," "matematiksel," "çözüm") indicate a deliberate move away from teacher-directed, procedural instruction toward a more student-centered, inquiry-based pedagogy, a trend also observed qualitatively in Turkish educational studies [1]. Conversely, the reduced presence of terms linked to technology and direct instructional directives (e.g., "iletişim," "teknoloji") suggests that such elements have either been integrated into the fabric of the curriculum or deprioritized in favor of fostering deeper mathematical understanding.

Methodological Contributions and Broader Implications

In addition to identifying curricular shifts, this study demonstrates the efficacy of unsupervised NLP techniques in analyzing complex, morphologically rich educational documents, despite the known challenges posed by the Turkish language [4]. The successful application of topic modeling, aligning with its increasing use in curriculum analysis [11], [12], supported by frequency and cognitive complexity analyses, underscores the value of methodological triangulation. Our computational approach to Bloom's taxonomy builds upon established manual coding practices [15] while navigating the complexities of automated classification, which previous studies have shown can vary in accuracy depending on the method (e.g., keyword matching vs. advanced models) [17], [18]. Our robust three-tiered validation framework provides a replicable model for curriculum analysis in contexts lacking established baselines, addressing a noted gap in the literature, particularly for morphologically complex languages.

By revealing these substantive shifts in mathematical content organization and cognitive demands, our research contributes to both educational practice and methodological development. The restructured curriculum, with its emphasis on creative and analytical thinking may inform changes in how mathematics is taught and better prepare students for a future where innovative problem-solving is paramount.

Limitations

While our approach yields useful insights, several limitations should be noted. First, the agglutinative nature of Turkish may introduce segmentation and lemmatization errors. Second, unsupervised methods like LDA inherently involve some randomness, which may result in minor variations in topic composition despite our efforts to ensure stability. Third, the absence of expert validation from mathematics educators limits the depth of our interpretative conclusions. Finally, our analysis primarily focuses on textual elements, potentially overlooking nuances in mathematical notation that might require more advanced symbolic analysis.

Conclusion

This research successfully employed unsupervised NLP techniques to reveal substantial evolutionary shifts between the 2018 and 2024 Turkish mathematics curricula. Our analysis identified a profound reorganization of content, characterized by new thematic priorities like mathematical logic and problem-solving strategies, alongside significantly expanded domain coverage. Critically, we observed a dramatic shift in cognitive demands, with a marked increase in higher-order 'Create' objectives and a decrease in 'Evaluate' objectives, signaling a pedagogical move towards student-generated reasoning and creative problem-solving.

Methodologically, this study contributes a validated computational framework, including a specialized preprocessing pipeline, for analyzing curricula in morphologically complex languages like Turkish. These techniques provide practical tools for examining curriculum evolution, especially in under-resourced or linguistically complex settings.

Declaration of Interest: The authors declare that there is no conflict of interest.

Acknowledgements: This research did not receive any specific grant or external funding from public, commercial, or not-for-profit sectors. All resources and analyses presented were independently developed and executed by the authors.

Author Contributions: Batın ÖRENE: Conceptualization, Methodology, Data Curation, Software Development, Formal Analysis, Visualization, Validation, Writing – Original Draft, Writing – Review & Editing. Ghislaine LE GRANGE: Validation, Conceptual Proofreading, Writing – Review & Editing.

Reproducibility Statement: All analyses were conducted in Python using open-source libraries including spaCy, Gensim, and scikit-learn. To support reproducibility, the complete preprocessing and analysis code—encompassing LDA topic modeling, Bloom's taxonomy classification, and co-occurrence network construction— is publicly available [24],[25].

REFERENCES

[1] Z. Kıryak, T. Kozaklı Ülger, B. B. Ülger, I. Bozkurt, and S. Çepni, "2018 ve 2024 İlk ve Ortaokul Fen Bilimleri ve Matematik Dersleri Öğretim Programları Öğrenme Çıktılarının Karşılaştırılması ve Beceriler Açısından İncelenmesi," *Bayburt Eğitim Fakültesi Derg.*, vol. 19, no. 44, pp. 3054–3089, Dec. 2024, doi: 10.35675/befdergi.1507283.

[2] J. Atuhurra, R. Chaudhry, and M. Kaffenberger, "Conducting Surveys of Enacted Curriculum Studies in Low- and Middle-Income Countries: A Toolkit for Policymakers, Researchers, and Education Practitioners," Research on Improving Systems of Education (RISE), Mar. 2023. doi: 10.35489/BSG-RISE-Misc_2023/13.

[3] "TIMSS 2023 TÜRKİYE RAPORU." Milli Eğitim Bakanlığı.

[4] K. Oflazer, "Turkish and its challenges for language processing," *Lang. Resour. Eval.*, vol. 48, no. 4, pp. 639–653, Dec. 2014, doi: 10.1007/s10579-014-9267-2.

[5] G. Eryiğit and K. Oflazer, "Statistical Dependency Parsing for Turkish," in *11th Conference of the European Chapter of the Association for Computational Linguistics*, D. McCarthy and S. Wintner, Eds., Trento, Italy: Association for Computational Linguistics, Apr. 2006, pp. 89–96. Accessed: Apr. 01, 2025. [Online]. Available: https://aclanthology.org/E06-1012/

[6] A. A. Akın and M. D. Akın, "Zemberek, an open source NLP framework for Turkic Languages".

S. Schweter, *BERTurk - BERT models for Turkish*. (Apr. 27, 2020). Zenodo. doi: 10.5281/ZENODO.3770924.

[7] "Türkiye Yüzyılı Maarif Modeli Öğretim Programları Ortak Metni." Milli Eğitim Bakanlığı, 2024.

[8] Ş. Danişman and E. Karadağ, "Comparison of 2005 and 2013 5th Grade Mathematics Curriculums in the context of Learning Domains and Objectives," *Turk. J. Comput. Math. Educ. TURCOMAT*, vol. 6, no. 3, p. 380, Dec. 2015, doi: 10.16949/turcomat.03423.

[9] R. Uysal and L. İNciKabi, "Son Dönem Matematik Dersi Öğretim Programlarının Genel Amaçları Üzerine Bir Araştırma".

[10] T. Sekiya, Y. Matsuda, and K. Yamaguchi, "Analysis of Curriculum Structure Based on LDA," 2009.

[11] P. A. Takizawa, "Using a topic model to map and analyze a large curriculum," *PLOS ONE*, vol. 18, no. 4, p. e0284513, Apr. 2023, doi: 10.1371/journal.pone.0284513.

[12] P. W. Airasian and K. A. Cruikshank, "A Revision of Bloom's Taxonomy of Educational Objectives," p. 333, 2001.

[13] L. Anderson, "Bloom ' s Revised Taxonomy : Cognitive , Affective , and Psychomotor," 2000.

[14] O. Kuzu, Y. Çiçek, and Z. İğdeli, "A Comparison of the Mathematics Curriculums in Turkey and Germany in the Context of Algebra Learning Domain," vol. 5, no. 1, 2023.

[15] H. Taş, "2018 MEB Ortaokul Matematik Programındaki Beş, Altı ve Yedinci Sınıfa Ait Kazanımların Zorluk Düzeylerinin Öğrenciler Açısından Değerlendirilmesi," 2020.

[16] Yuheng Li, M. Rakovic, Boon Xin Poh, D. Gasevic, and Guanliang Chen, "Automatic Classification of Learning Objectives Based on Bloom's Taxonomy," Jul. 2022, doi: 10.5281/ZENODO.6853191.

[17] S. Shaikh, S. M. Daudpotta, and A. S. Imran, "Bloom's Learning Outcomes' Automatic Classification Using LSTM and Pretrained Word Embeddings," *IEEE Access*, vol. 9, pp. 117887–117909, 2021, doi: 10.1109/ACCESS.2021.3106443.

[18] K. Banujan, S. Kumara, and S. Prasanth, "Revolutionising Educational Assessment: Automated Question Classification using Bloom's Taxonomy and Deep Learning Techniques – A Case Study on Undergraduate Examination Questions".

[19] H. Sebbaq and N. El Faddouli, "MTBERT-Attention: An Explainable BERT Model based on Multi-Task Learning for Cognitive Text Classification," *Sci. Afr.*, vol. 21, p. e01799, Sep. 2023, doi: 10.1016/j.sciaf.2023.e01799.

[20] "Ortaöğretim Matematik Dersi (9, 10, 11 ve 12. Sınıflar) Öğretim Programı." Milli Eğitim Bakanlığı, 2018.

[21] "Ortaöğretim Matematik Dersi Öğretim Programı (Hazırlık, 9, 10, 11 ve 12. Sınıflar)." 2024.

[22] Turkish NLP Suite, *tr_core_news_trf: Turkish transformer pipeline for TrSpaCy*. (2023). Hugging Face. [Online]. Available: https://huggingface.co/turkish-nlp-suite/tr_core_news_trf

[23] B. Örene and G. Le Grange, *Turkish Mathematics Curriculum NLP Analysis*. GitHub repository. [Online]. Available: https://github.com/batinium/turkish-math-curriculum-nlp

[24] B. Örene and G. Le Grange, *batinium/turkish-math-curriculum-nlp: Initial Release*. (Mar. 29, 2025). Zenodo. doi: 10.5281/ZENODO.15106063.

ARTIFICIAL INTELLIGENCE IN MULTILINGUAL ITINERARY PLANNING: A CASE STUDY OF GROK IN DIYARBAKIR

Vedat Kızıl*

Batman University, Graduate Education Institute, Department of Tourism Management, Batman, Turkey

(Corresponding author) vekizil@gmail.com

ABSTRACT

Tourism is an old pursuit that enables individuals to explore cultural heritage and natural wonders and to broaden their perspectives while keeping and sharing regional, and cultural values [1]. Globalization has made travel planning more complicated day by day, with the chance of seeing multiple different destinations and forcing us to get used to fast information flows. Artificial intelligence (AI) is made to address such challenges by offering multilingual, personalized solutions. This study assessed Grok, an AI tool by xAI, in creating a one-day travel route in Diyarbakır in Turkish, English, and French languages. On March 19, 2025, we entered this query in Grok: "I am at Diyarbakır Airport. Tomorrow, Wednesday, March 19, 2025, I will explore Diyarbakır's historical and cultural sites. Can you prepare an itinerary based on current data?" Responses were analyzed within the Tourist Experience Theory framework [1]. The Turkish itinerary gave responses in cultural depth and accuracy, featuring local drinks like menengiç coffee, while English and French answers showed deficiencies in linguistic and contextual meaning. This paper shows conversational travel planning in an AI system program that provides travel recommendations in an individual-based way to assist tourists in diverse languages and cultures and suggests improvements for its multilingual capabilities.

Keywords: Artificial intelligence, tourism, itinerary planning, Diyarbakır, multilingual guidance, Grok.

Introduction

Traveling is a term that comes from people's curiosity about different cultures and beautiful places which connects people while they explore[1]. This activity is not only a great personal and group adventure but also helps keep culture alive. [8]. Globalization has made planning a lot difficult, as there are many different places and info sharing gets faster this forces us the change the usual ways we use like guidebooks and travel guides. [2].The importance of AI technologies is getting expanded day by day which offers more flexible and quick solutions for our personal needs [2]. The journey of artificial intelligence in tourism unfolds in the literature, illustrating its transformation from basic reservation systems to sophisticated chatbots that offer real-time assistance. [5]. Grok, created by xAI, demonstrates advancements in multilingual functionality.[7]. We chose Diyarbakır for this

experiment—a city in Southeastern Anatolia where 5.5 km of UNESCO-listed basalt walls and the ancient Ulu Camii whisper history at every turn (UNESCO, 2015). It's a gem that deserves more spotlight, yet there's not much out there on how AI can boost its tourism game. So, we set out to see if Grok could craft itineraries in Turkish, English, and French that don't just list places but capture Diyarbakır's soul. We looked at its cultural richness, how spot-on it was, and whether it actually helped you get around.

Methods

This study addresses travel-related itinerary planning from the perspective of the multilingual trip planner provided as an example of the implementation of itinerary planning systems using language processing and artificial intelligence technologies between Turkish, English, and French. The proposed itinerary planner allows its users to plan their travel itineraries by selecting sightseeing locations, localized services, and products in Diyarbakır, Turkey

Data Collection and Analysis

On March 19, 2025, we entered this query on Grok website: "I am at Diyarbakır Airport. Tomorrow, Wednesday, March 19, 2025, I will explore Diyarbakır's historical and cultural sites. Can you prepare an itinerary based on current data?" The question was asked in Turkish, English, and French languages to test chatbot's multilingual performance. Responses were gathered through document analysis, a qualitative approach designed for assessing written content.[4]. Analysis was conducted within the Tourist Experience Theory framework, focusing on cultural depth, accuracy, linguistic harmony, and practicality [1]. We checked out each itinerary to see what worked and what didn't in different languages, comparing our findings to show the differences..

Results

Grok's Turkish itinerary included six places: Diyarbakır Walls, Ulu Camii, Hasan Paşa Hanı, Cahit Sıtkı Tarancı Museum, Ten-Eyed Bridge, and İçkale. It provided detailed transport options such as taxis or buses from the airport and regional recommendations like menengiç coffee and ciğer kebabı, reflecting high cultural depth and accuracy. The itinerary's useful guidance, such as travel durations and weather information and tips, improved its utility. The English itinerary detailed seven locations, including Sülüklü Han and Gazi Köşkü, yet it mistakenly dated the construction of Ulu Camii to the 11th century when, in fact, its origins trace back to the 7th century. Furthermore, it failed to include any recommendations for local cuisine, leaving a gap in culinary exploration. The transport information provided was also lacking, which limited its usefulness for travelers. In contrast, the French itinerary, which focused on five sites such as Dört Ayaklı Minare and İç Kale, inaccurately identified Ulu Camii as "Islam's fifth holiest site." Like its English counterpart, it offered scant transport details, which diminished its cultural significance and practical application. These findings reveal a significant distinction in performance, as the Turkish itinerary outperformed both the suggestions in English and French languages.

Discussion

There's something special about how the Turkish itinerary pulls you into Diyarbakır's world—it's the kind of connection Cohen talks about in his Tourist Experience Theory, where local touches make the place come alive [1]. The menengiç coffee suggestion? That's not just a drink; it's a taste of the streets here. It's clear Grok's got a solid grip on Turkish data, which makes sense given its depth[1]. The level of precision and thoroughness implied a well-developed dataset for Turkish language training. [3]. On the other hand But flip to English or French, and it's like the magic fades—historical slip-ups and missing details dim the experience. It's a bit like handing a guidebook to someone who doesn't know the town—useful, but you miss the soul if it's not tuned right. This observation resonates with previous findings that highlight the necessity for AI to align itself with the context in which it operates [5]. Grok's ability to personalize experiences and provide updates in real time stands out as one of its key strengths. However, the platform's limitations in supporting multiple languages reveal specific areas where there is room for improvement and expansion. [2]. This difference emphasizes the potential of AI as a valuable tool in the tourism industry, provided that it is carefully designed for that purpose. However, it also reveals the dangers of producing superficial experiences if the technology is not adequately developed.

Conclusion and Recommendations

Artificial intelligence is capable of supporting flexible itinerary planning services for non-English-speaking areas and can extend the methodology for broader and general purposes [9]. Grok demonstrated significant potential in multilingual itinerary planning, particularly in Turkish, where it enriched tourist experiences with cultural depth and accuracy. However, English and French shortcomings in linguistic harmony and contextual richness indicate that there is a need for improvement. The study suggests Grok's training data may prioritize Turkish contexts, limiting its global applicability. To improve the effectiveness of Grok, it is suggested that the project embrace a wider range of multilingual datasets. In addition, incorporating user feedback will promote ongoing enhancements, while working with local experts may help to maintain cultural accuracy. Looking ahead, it would be beneficial to assess Grok in a variety of destinations, thereby expanding its influence on tourism. By taking these actions, it could be a real bridge between cultures and sustainable travel.

Declaration of Interest: The author declare that there is no conflict of interest.

Acknowledgements: The author thanks Batman University for providing research resources and support

Author Contributions: Vedat Kızıl designed the study, collected and analyzed the data and reviewed the manuscript.

REFERENCES

[1] Cohen, E. "A phenomenology of tourist experiences," Sociology, 13(2), 179-201, 1979.

[2] Gretzel, U., Sigala, M., Xiang, Z., & Koo, C. "Smart tourism: Foundations and developments," Electronic Markets, 25(3), 179-188, 2015.

[3] Tussyadiah, I. P. "Toward a theoretical foundation for experience design in tourism," Journal of Travel Research, 53(5), 543-564, 2014.

[4] Kıral, B. "Nitel araştırma yöntemleri: Teori ve uygulama," Eğitim ve Bilim Dergisi, 45(202), 1-15, 2020.

[5] Ivanov, S. H., Webster, C., & Berezina, K. "Adoption of robots and service automation by tourism and hospitality companies," Revista Turismo & Desenvolvimento, 27/28, 1501-1517, 2017.

[6] UNESCO. "Diyarbakır Fortress and Hevsel Gardens Cultural Landscape," UNESCO World Heritage Centre, https://whc.unesco.org/en/list/1488/, 2015.

[7] xAI. "Grok: An AI companion for understanding the universe," xAI Official Website, https://xai.ai/grok, 2023.

[8] Buhalis, D., & Law, R. "Progress in information technology and tourism management," Tourism Management, 29(4), 609-623, 2008.

[9] David N. AI-powered virtual assistant solution for supporting international students. 2024. theseus. fi

ARTIFICIAL INTELLIGENCE PERFORMANCE IN EMBEDDED SYSTEMS: COMPARISONS OF INFERENCE TIME, COST, AND ENERGY CONSUMPTION

Taner Erol*

Bolu Abant İzzet Baysal University, Faculty of Engineering, Department of Electrical and Electronics Engineering, Bolu, Türkiye

Kemal Polat

Bolu Abant İzzet Baysal University, Faculty of Engineering, Department of Electrical and Electronics Engineering, Bolu, Türkiye

**(Corresponding author) kpolat@ibu.edu.tr*

ABSTRACT

Artificial Neural Network (ANN) algorithms have increasingly begun to show themselves in every field thanks to their efficient results in many fields, especially in health, security, agriculture, education, etc. The fact that these algorithms have entered every area of our lives, including daily life, has increased the interest and tendency in studies in this field due to the built-in capabilities provided by embedded machine learning in response to cloud or remote server needs. Running machine learning (ML) and deep learning (DL) algorithms embedded in wearable or portable devices provides fast results and offers many conveniences. In this article, the focus is on examining and comparing different embedded system cards such as Raspberry Pi, NVIDIA Jetson, Asus Tinker Board, Banana Pi, etc., which are preferred for studies in various fields such as agriculture, robotics, autonomous driving, health, internet of things (IoT), security, drone-based surveillance, etc., and which system can be more efficient within the framework of criteria such as the effect of visual and non-visual sensor data on analysis time, performance and cost, etc. All single-card platforms examined show quite good performance. One of the reasons for performance differences is the complexity of the sensor used in the applications and the targeted purpose. While Raspberry Pi cards stand out in terms of their success and cost advantage in processing non-visual data, Jetson series cards have been shown to be a more appropriate choice in processing visual data. This study provides a good roadmap when considering the operation of ML and DL algorithms on embedded systems within the framework of real-time and edge computing principles.

Keywords: Embedded system, deep learning, machine learning, jetson, raspberry pi, edge computing, single board computer, edge computing, embedded machine learning.

Introduction

Embedded systems (Edge Computing) are specially developed hardware structures to perform artificial intelligence (AI) calculations at the edge of the network and near users, instead of directing the processing of sensor data to a central location such as a cloud service provider's data center [1]. This concept is only one part of the general architecture of "cloud-fog-edge".

Embedded systems, which are the lowest layer of the architecture, can be defined as technological platforms that bring computing processes closer to where data is produced and collected. The place of embedded systems in the architectural structure can be seen in Figure 1.

The proliferation of IoT devices and the importance given to computational efficiency have accelerated the work on embedded systems. When DL algorithms are traditionally executed in a central cloud environment, the cloud computing approach causes large delays, energy and financial costs. It also increases privacy concerns. To reduce these difficulties, the edge computing approach has been proposed, in which the computation on the data is done locally, i.e., the data is produced where the edge devices are used [2] [3].

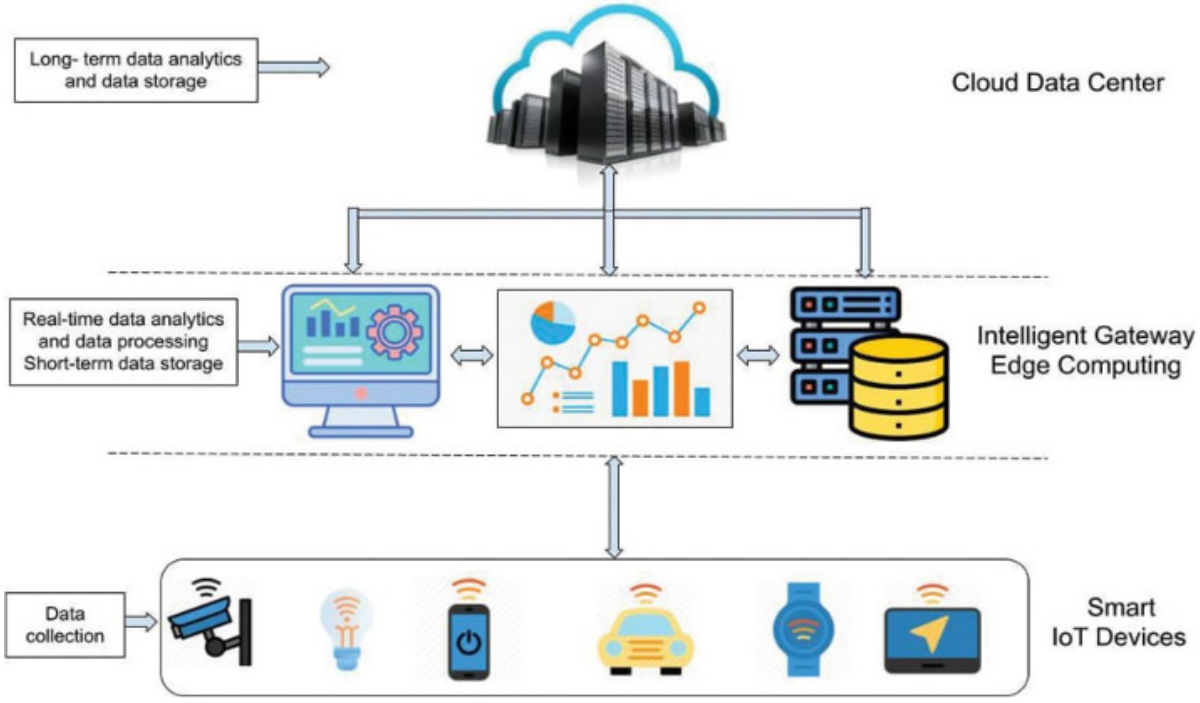


Figure 1: Architectural structure of embedded systems [4] [5].

Single-board computer systems, which have high applicability and low cost, have recently become frequently preferred due to their ability to be combined with other technology areas. However, efficient hardware implementation of ML algorithms, especially for real-time applications, is still a challenge [6]. In this study, embedded platforms that enable the operation of the developed ML or DL model are compared in terms of features and costs, while referring to accuracy, inference time and power consumption comparisons, with the aim of providing a start for studies to be carried out in this field.

Literature Review

Mittal and Vetter [7] concluded in their studies that the Jetson series cards, which emerged as embedded systems, can preload the operating system of the CPU and the CUDA-enabled GPU can be quickly programmed to accelerate complex ML tasks, have a CPU/GPU heterogeneous architecture, and also have a small form factor, low weight and power consumption, making them perfectly suited to weight/power constrained scenarios.

Diab and Villegas [8] stated that machine learning data calculations can be processed on three different bases: cloud, fog and edge, and each has its advantages and disadvantages. They emphasized that cloud and fog-based applications are based on the principle of sending data to remote or local servers, which brings privacy, security and network problems, but since computation is performed close to the sensor level in edge-based (embedded) applications, the need for raw data transfer is eliminated. This principle allows the data to be interpreted directly at the front end of the device. It has been found that while the privacy and security disadvantages are eliminated in embedded systems, the computations are limited by the capability of the preferred system.

Rungsuptaweekoon et al. [9] evaluated the power efficiency of image recognition with YOLO, a real-time object detection algorithm, on NVIDIA embedded GPU systems Jetson TX1 and TX2 to overcome this challenge, as deep learning inference in embedded systems requires not only high efficiency but also low power consumption. Their findings show that it is possible to adjust the trade-off between efficiency and power efficiency in Jetson TX2, and therefore Jetson TX2 has more advantages over Jetson TX1 and a PC server with NVIDIA Tesla P40 in image recognition in embedded systems.

Blanco-Filgueira et al. [10] achieved real-time performance in 10 fps video capture with only 12 W total power consumption, using only the current frame of the live camera at any time, without intermediate storage for delayed processing, in terms of speed and power consumption using the NVIDIA Jetson TX2 development kit. It was shared that memory and computational resources provide the expected requirements to run the entire solution as on a standard platform. It was stated that the latency problems caused by cloud computing will make embedded systems an alternative.

Method

The technological development of single-board computers and their approach to the performance of personal computers have strengthened their use in the field of artificial intelligence. Similarly, the fact that some single-board computers have both CPU and GPU has increased the applicability of artificial intelligence on single-board computers [11]. This article examines the studies that have analyzed platforms such as Jetson and Raspberry Pi, various ML and DL applications applied to these platforms, which of these devices are more widely used for specific applications in different areas, specific ML models applied to the systems and sensor data that provide input. As a result, the hardware feature of the used embedded system, general power consumption, inference times for different models applied and current cost comparisons have been made.

General Hardware

The processor type and speed, memory capacity, storage structure and data path capacity, which are the units that all computer structures have and should be considered when choosing an embedded system for a specific machine learning application, are examined in terms of the embedded systems examined in the article and visualized in Figure 2. Since the article is hardware-focused, a software comparison is not included.

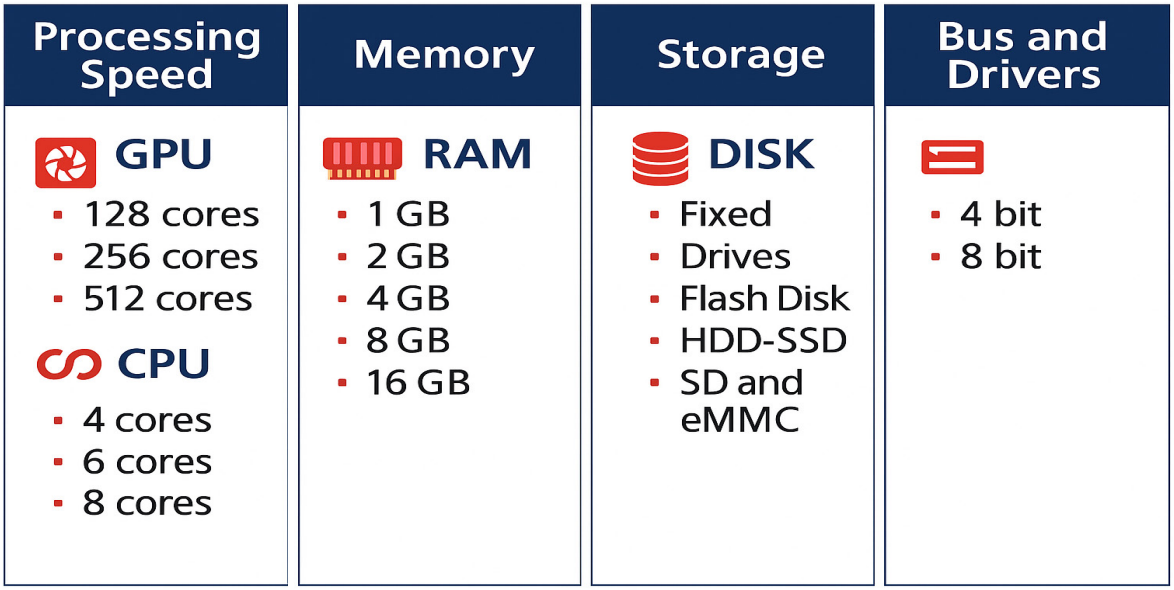


Figure 2: Min-max speed and capacity values of physical units of different existing embedded systems

Special systems

In addition to lower-requirement ML and DL tasks such as image classification, object detection, segmentation and speech processing, creating end-to-end AI robotic applications for manufacturing, delivery, retail and agriculture, NVIDIA Jetson, which can handle even challenging deep learning applications with real-time and edge computing principle, Raspberry Pi and some other embedded systems, which are frequently preferred in education and smart wearable products, etc., are customized according to the requirements given in Figure 1 and classified in Table 1. It is an undeniable fact that one of the biggest challenges a project can face is cost. Most of the time, hardware pREFERENCES that do not meet the requirements due to high costs and related problems may occur. Considering this situation, cost data was included in the study based on the current market value.

NVIDIA Jetson embedded systems

NVIDIA Jetson series is a leader in single-board computing for deep learning applications due to the variety of features it offers in developer kits. Some of the features of the Jetson series are energy efficiency, lightness and portability, ease of use, and high performance per watt. NVIDIA's Jetson series is seen as the most widely used accelerator in the ML extraction stage [11]. TX series cards from Jetson embedded systems are often

preferred over NX and AGX cards due to their ability to handle most applications and their relatively more affordable cost. Since the production of Jetson TX1 model cards has been discontinued, it should be taken into account during the project planning phase that it may be difficult to access these systems.

Raspberry Pi embedded systems

Raspberry Pi represents a versatile system with a valid prototyping base and the potential to be easily incorporated [12]. Raspberry Pi system is a portable and powerful microcontroller board. These boards have been very useful for healthcare, defense, agriculture, education, smart wearables, energy, mining and automation projects [13]. Different models of Raspberry Pi boards have been developed. A table of important features of Raspberry Pi boards, which play important roles in many applications, is given in Table 1.

Some other systems (Asus Tinker Board, Odroid XU4, Banana Pi)

ASUS Tinker Board is a powerful single-board computer developed as an alternative to Raspberry Pi. It offers a capable platform that can provide a variety of projects with its powerful hardware features and wide connectivity options. However, it has limitations in terms of software support and community resources.

Odroid XU4 is powered by the Samsung Exynos 5422 octa-core processor, which consists of four ARM Cortex-A15 cores running at 2 GHz. The remaining cores are ARM Cortex-A7 with a clock speed of 1.4 GHz. The combination of these cores provides a balance between performance and energy efficiency. Since the number of pins is low compared to other platforms, it has lower compatibility with external devices and sensors. In addition, its biggest disadvantage is that it generates a significant amount of heat under heavy computational loads, which requires a cooling utility and poses more difficulties in the face of space limitations [14].

Banana Pi platforms can be used in various electronic projects, educational applications, and as personal computers, similar to Raspberry Pi. Table 1 shows the hardware specifications of the Asus Tinker Board family, Odroid XU4 and Banana pi platforms.

Table 1: Hardware comparisons of embedded platforms

Jetson Nano	Jetson TX1	Jetson TX2	Jetson Xavier NX	Jetson Xavier AGX	Raspberry Pi 3B	Raspberry Pi 3B+	Raspberry Pi 4B	Raspberry Pi 5	Asus Tinker Board S	Asus Tinker Board R	Asus Tinker Board T	Banana PI	Odroid XU4
GPU 128 Core Maxwell GPU	GPU 256 Core Maxwell GPU	GPU NVIDIA Pascal™ Mimarisi GPU	GPU 384 NVIDIA CUDA® ve 48 NVIDIA Volta architecture with tensor cores GPU	GPU Tensor 512 Core Volta GPU	GPU 400 MHz VideoCore IV video processor	GPU 400 MHz VideoCore IV video processor	GPU 400 MHz VideoCore IV video processor	GPU Integrated VideoCore VII 800MHz GPU (2 to 3 times more performance) 60fps rendering	GPU ARM-based Mali™-T764 GPU	GPU ARM® Mali™-T860 MP4 GPU	GPU GC7000 Lite	GPU H3 4 Core Cortex-A7 H.265/HEVC 4K, Mali400MP2 GPU	GPU Mali-T628 MP6 GPU
											TPU Google Edge TPU		
CPU 4 Core ARM® Cortex®- A57 1.4 CPU	CPU 4 Core ARM® Cortex®-A57 CPU	CPU 4 Core ARM® Cortex®-A57 CPU	CPU 6 Core NVIDIA Carmel ARM®v8.2 64 bit CPU	CPU 8 Core ARM v8.2 64 bit CPU	CPU 4 Core 1,2 GHz Broadcom BCM2837 64 bit CPU	CPU 4 Core 1,4 GHz Broadcom BCM2837B0, Cortex-A53 (ARMv8) 64 bit SoC CPU	CPU 4 Core 1,4 GHz Broadcom BCM2837B0, Cortex-A72 (ARMv8) 64 bit SoC CPU	CPU 4 Core 1,5 GHz Broadcom BCM2837B0, Cortex-A72 (ARMv8) 64 bit SoC CPU	CPU 4 Core ARMA76 processor (64 bit - 2,4 GHz)	CPU Rockchip RK3288 4 Core Cortex-A17 -1,8 GHz CPU	CPU Arm® big.LITTLE® A72+A53 6 Core CPU	CPU 4 Core ARM Cortex-A53 CPU	CPU 4 Core ARM Cortex-A55 CPU	CPU ARM Cortex-A15 and Cortex-A7 4x Cortex-A15 Core (high performance) 4x Cortex-A7 Core (energy efficiency)
											NPU 2GB LPDDR3 NPU on Rockchip			
RAM 4 GB 64 bit LPDDR4 25,6 GB/s Memory	RAM 4 GB 64 bit LPDDR4 Memory	RAM 8 GB 128 bit LPDDR4 Memory	RAM 8 GB 128 bit LPDDR4x Memory	RAM 32 GB 256 Bit LPDDR4x Memory	RAM 1 GB LPDDR2 Memory	RAM 1 GB LPDDR2 Memory	RAM 1 GB, 2 GB, 4 GB and 8 GB LPDDR2 Memory	RAM 4 GB and 8 GB LPDDR4X Memory	RAM 2GB LPDDR3 Memory	RAM 4 GB Dual chanel LPDDR4 Memory	RAM 1 GB LPDDR4 Dual chanel Memory	RAM 1 GB DDR3 Bellek	RAM 2 GB LPDDR3 Bellek
CAM 2x MIPI CSI-2 DPHY lane camera	CAM 5MP Fixed Focus MIPI CSI Camera	CAM 5MP Fixed Focus MIPI CSI Camera	CAM 2x MIPI CSI-2 DPHY lane camera	CAM Camera connector	CAM Camera connector	CAM Camera connector	CAM 5MP Fixed Focus MIPI CSI Camera	CAM 2x 4-lane MIPI, DSI/CSI connections	CAM 1x 15-pin MIPI CSI HD	CAM MIPI CSI-2 and MIPI CSI-2/DSI strip camera	CAM 2xMIPI-CSI	CAM CSI camera port	CAM From USB
PORT Ethernet, HDMI ve USB	PORT Ethernet, HDMI and USB tip A and Mikro AB	PORT Ethernet, HDMI and USB tip A and Mikro AB	PORT Ethernet, HDMI and USB tip A and Mikro AB	PORT Ethernet, HDMI and 2xUSB C connector	PORT 100 Base Ethernet, 4 USB 2.0 ve HDMI	PORT 1000 Base Ethernet, 4 USB 2.0 ve HDMI	PORT 1 Gigabit Ethernet, 4 USB 2.0 ve HDMI	PORT 2x USB 3.0 5 GBps, PCIe 2.0 connected , Integrated RTC connector, USB-C, Dual 4K@60 HDMI with HDR support	PORT 1x HDMI, up to 4K/30 fps, MicroSD, RTL HD Codec, 4 x USB 2.0, 1X GB LAN, 192K/24bit HD Audio	PORT USB 3.2 type A and 1xUSB 3.2 Type C, Ethernet and HMDI, Wi-Fi and Bluetooth, Mini PCIe slot (nano-SIM slot), audio jack	PORT HDMI, USB Type-A, USB Type-C, LAN and integrated Wi-Fi and Bluetooth, Mini PCIe slot (nano-SIM slot), audio jack	PORT 2XUSB 2.0 , 1X HDMI, 1XEthernet interface	PORT 2xUSB 3.0 and 1xUSB 2.0, Gigabit Ethernet and HMDI 1.4
DISK MicroSD cart	DISK 16 GB eMMC 5.1 Flash	DISK 32 GB eMMC 5.1 Flash	DISK 32 GB eMMC 5.1 Flash	DISK 32 GB eMMC 5.1 Flash	DISK MicroSD cart	DISK MicroSD cart	DISK MicroSD cart	DISK High speed mikroSD	DISK 16 GB eMMC , SD 3.0	DISK 16GB eMMC Flash and microSD slot	DISK 8GB eMMC memory SD 3.0 interface	DISK 8 GB eMMC	DISK 2 GB eMMC5.0 and microSD
Power Consumption 5-10 Watt	Power Consumption 15 Watt	Power Consumption 25 Watt	Power Consumption 10 Watt	Power Consumption 30 Watt	Power Consumption 1,5 Watt	Power Consumption 2 Watt	Power Consumption 4 Watt	Power Consumption Over 4 Watts	Power Consumption 5 Watt	Power Consumption 65 Watt power supply	Power Consumption 45 Watt power supply	Power Consumption 5 Watt	Power Consumption 5 Watt
Cost (Kit) 340$	Cost (Kit) 800$	Cost (Kit) 1000 $	Cost (Kit) 1100$	Cost (Kit) 1400$	Cost (Kit) 40$	Cost (Kit) 40$	Cost (Kit) 50$	Cost (Kit) 100 $	Cost (Kit) 150$	Cost (Kit) 250$	Cost (Kit) 220$	Cost (Kit) 50$	Cost (Kit) 100 $

Findings and Results

Each of the embedded systems examined in the previous section has been the subject of very different applications. The inference times obtained from studies on image recognition and object detection according to the embedded system and usage area, including agriculture, face recognition, depth estimation, autonomous vehicle obstacle detection, medical diagnosis/service, smart city applications and security applications, are given in Figure 3, while the inference times obtained from studies on non-image signal processing according to the processed signal and the embedded system used are given in Figure 4.

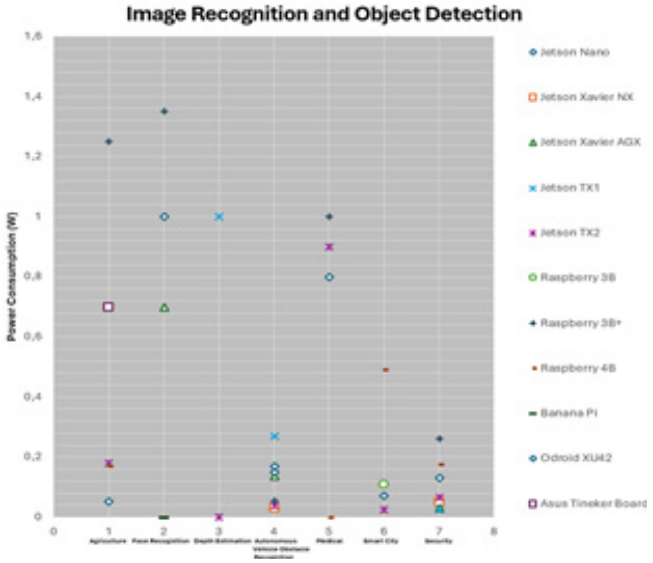


Figure 3: Inference time performance relationships for image recognition and object detection

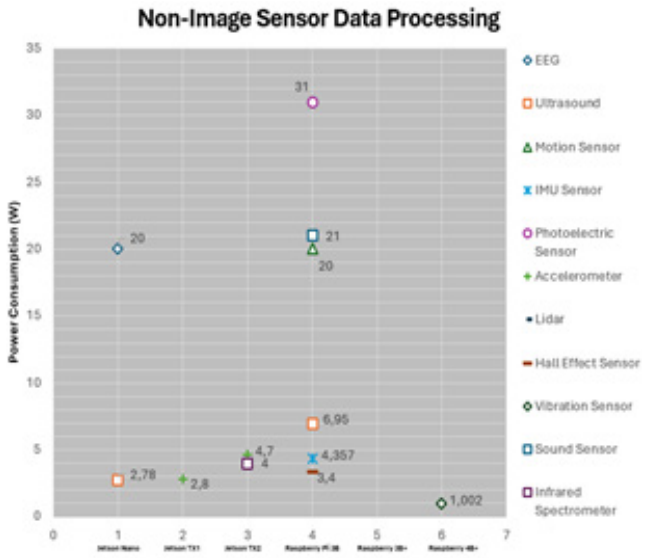


Figure 4: Inference time performance relationships for non-image sensor data processing

All single-board platforms examined in the article perform quite well. One of the reasons for the performance differences is the complexity of the sensor used in the applications and the targeted purpose. As can be seen in Figure 3, Jetson Xavier AGX and Jetson Xavier NX stand out as the most efficient hardware by providing the lowest inference times for scenarios requiring high performance in real-time applications. Jetson TX2 and Jetson Nano offer a balanced option in terms of price-performance. On the other hand, Raspberry Pi 3B, 3B+ and 4B models and Banana Pi show more limited performance in terms of processing power and their inference times are longer.

Differences in hardware performance were observed depending on the task type. In tasks such as face recognition and depth tracking, significantly longer inference times were observed in some hardware. In contrast, lower latency times were generally achieved in security and smart city applications.

When the data obtained from Figure 4 is evaluated, hardware such as Jetson Nano, Jetson TX1 and Jetson TX2 generally work efficiently with low processing times. Raspberry Pi 3B, 3B+ and 4B, on the other hand, exhibit relatively low performance by exhibiting longer inference times. Jetson TX1 and Jetson TX2, in particular, stand out as strong alternatives in terms of data processing, offering average processing times of 2-4 ms.

Significant differences were also observed between sensor types. EEG, IMU and Photoelectric sensors required longer processing times compared to other sensors. In particular, the processing time of the EEG sensor reached a high value of 31 ms, making it one of the slowest processed sensors. In contrast, the Infrared Spectrometer sensor was one of the fastest sensors with a low extraction time of 1.002 ms.

As a result of the general evaluation, it is recommended that Jetson series hardware with more powerful processors be preferred for applications such as motion sensors, accelerometers and Hall effect sensors that require low latency and for sensors that require complex data processing such as EEG. The Raspberry Pi series can be evaluated for low-cost and energy-efficient applications, despite having longer processing times.

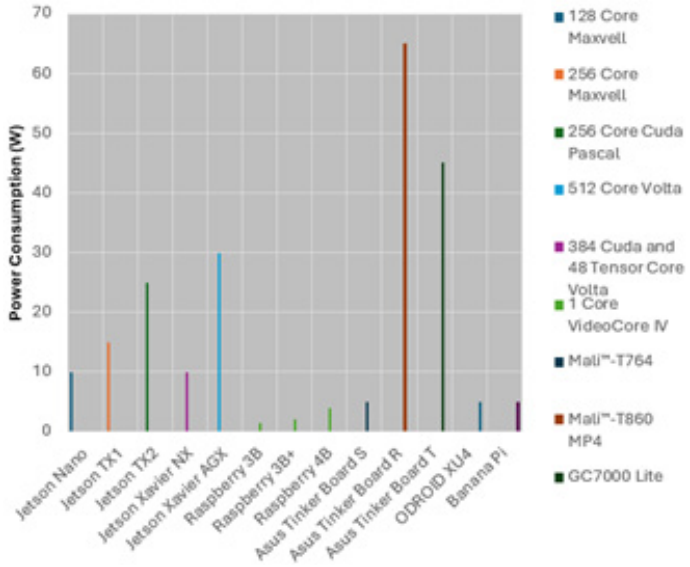


Figure 5: Power consumption of embedded systems depending on GPU features

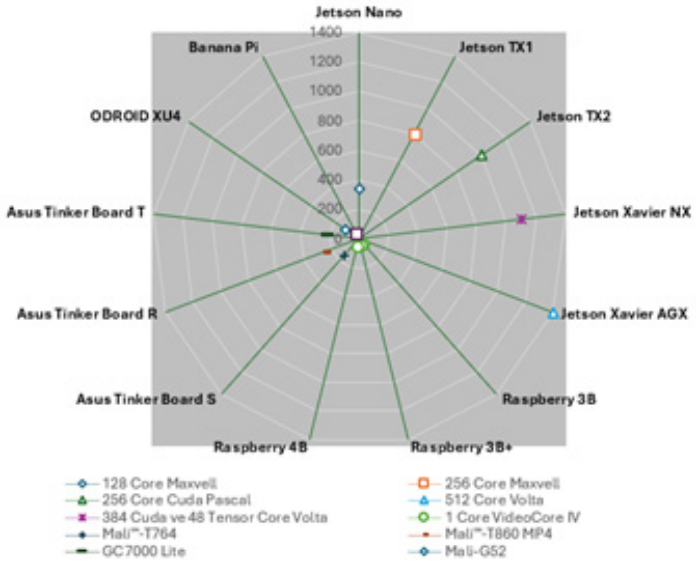


Figure 6: Costs of GPU features in embedded systems

It can be said that performance improves in direct proportion to the improvement of the physical capabilities of the platforms. This situation also brings about cost increases. Figure 5 shows the power consumption and related cost changes when the number of cores in the GPU units of embedded platforms is increased and the architecture is improved, and Figure 6 shows. All machine learning applications aim to consume as little power as possible and to calculate as fast and accurately as possible, but an improvement in one of these areas almost always brings a relative cost compared to the others [15].

In line with the results, the use of the Jetson Xavier series is recommended in applications requiring high accuracy and speed, while the Jetson Nano and TX series can be considered as a suitable alternative in projects requiring low cost and medium performance. When low power consumption is an important criterion, the Raspberry Pi series can be preferred, but it should be taken into account that processing times are longer. Especially in scenarios

that are not sensitive to latency, such as security and smart city applications, this hardware can be considered as a more suitable option.

Discussion

This article is presented by examining the articles accessed as a result of the search using Google Academic. The data examined is classified into two scopes as image and non-image. Accordingly, Jetson embedded platforms are mostly preferred in image recognition and object detection applications, while Raspberry Pi platforms are preferred in non-image data processing applications. While Raspberry Pi provides low power and energy saving performance, it is concluded that NVIDIA Jetson platforms such as Jetson Nano, Jetson TX1, Jetson TX2 have higher performance due to their faster GPUs. In addition, it has been evaluated that the lowest-equipped Jetson Nano developer kit is superior to the Raspberry Pi kit in some DL applications [16]. However, when compared to Jetson TX2 and Jetson Nano in CNN tests among Jetson cards, Jetson Nano shows two times lower performance [17]. Jetson Nano is the platform with the lowest hardware structure examined in this article. Despite this, it has been concluded that it gives successful results in many DL applications. When considered in this context, it can be said that Jetson Nano kits stand out as price/performance products. It can be said that the cost criterion is also effective in choosing the system to be used, in addition to meeting the requirements. As a result, this study is important in terms of guiding future studies when considering the operation of ML and DL algorithms on embedded systems within the framework of real-time and edge computing principles.

Declaration of Interest: The authors declare that there is no conflict of interest in this study.

Thanks: The authors declare that no institution or organization provided direct support in the preparation of this study.

Author Contributions: Taner EROL: Conceptualization of the study, data collection, data analysis, graphic creation, technical arrangements and writing process. Kemal POLAT: Literature review, interpretation of results and correction process.

REFERENCES

[1] Raghubir Singh ve Sukhpal Singh Gill, "Edge AI: A survey," Internet of Things and Cyber-Physical Systems, cilt 3, ss 71-92, 2023, DOI: https://doi.org/10.1016/j.iotcps.2023.02.004

[2] Sparsh Mittal , "A Survey on optimized implementation of deep learning models on the NVIDIA Jetson platform," Journal of Systems Architecture, cilt 97, ss 428-442, 2019, DOI: https://doi.org/10.1016/j.sysarc.2019.01.011

[3] L. Roda-Sanchez, C. Garrido-Hidalgo, F. Royo, J. L. Maté-Gómez, T. Olivares, and A. Fernández-Caballero, "Cloud–edge microservices architecture and service orchestration: An integral solution for a real-world deployment experience," Internet of Things, vol. 22, p. 100777, 2023, DOI: 10.1016/j.iot.2023.100777.

[4] N. Hassan, S. Gilani, E. Ahmed, I. Yaqoob, and M. Imran, "The Role of Edge Computing in Internet of Things," *IEEE Communications Magazine*, vol. PP, 2018, DOI: 10.1109/MCOM.2018.1700906.

[5] S. Naveen and M. R. Kounte, "Key Technologies and challenges in IoT Edge Computing," *Proceedings of I-SMAC (IoT in Social, Mobile, Analytics and Cloud)*, pp. 61-65, 2019, doi: 10.1109/I-SMAC47947.2019.9032541.

[6] Gomez, Agustin, Cuiñas Vázquez, Daniel, Català Calderon, Pau, Xin, Lin, Li, Weidong, Conway, Stewart, ve Lack, Donald , "Use of single board computers as smart sensors in the manufacturing industry", Procedia engineering, cilt 132, ss 153-159, 2015. DOI: https://doi.org/10.1016/j.proeng.2015.12.461

[7] Sparsh Mittal ve Jeffrey S. Vetter, "A Survey of CPU-GPU Heterogeneous Computing Techniques," ACM Journals, cilt 47, sayı 4 ss 1-35, 2015, DOI: https://doi.org/10.1145/278839

[8] M. S. Diab ve E. Rodriguez-Villegas, "Embedded Machine Learning Using Microcontrollers in Wearable and Ambulatory Systems for Health and Care Applications: A Review," in IEEE Access, cilt 10, ss 98450-98474, 2022, DOI: 10.1109/ACCESS.2022.3206782.

[9] K. Rungsuptaweekoon, V. Visoottiviseth and R. Takano, "Evaluating the power efficiency of deep learning inference on embedded GPU systems," 2017 2nd International Conference on Information Technology (INCIT), Nakhonpathom, Thailand, 2017, ss 1-5, DOI: 10.1109/INCIT.2017.8257866.

[10] B. Blanco-Filgueira, D. García-Lesta, M. Fernández-Sanjurjo, V. M. Brea ve P. López, "Deep Learning-Based Multiple Object Visual Tracking on Embedded System for IoT and Mobile Edge Computing Applications," in IEEE Internet of Things Journal, cilt 6, sayı 3, ss 5423-5431, 2019, DOI: 10.1109/JIOT.2019.2902141.

[11] A. A. Süzen, B. Duman ve B. Şen, "Benchmark Analysis of Jetson TX2, Jetson Nano and Raspberry PI using Deep-CNN," 2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), Ankara, Turkey, 2020, ss 1-5, DOI: 10.1109/HORA49412.2020.9152915.

[12] Madina Shayne, Leonardo A. Molina, Bin Hu ve Taylor Chomiak ,"Implementing Gait Kinematic Trajectory Forecasting Models on an Embedded System," Sensors 2024, cilt 24, sayı 8, ss 2649, DOI: https://doi.org/10.3390/s24082649.

[13] S. Karthikeyan, R. A. Raj, M. V. Cruz, L. Chen, J. L. A. Vishal ve V. S. Rohith, "A Systematic Analysis on Raspberry Pi Prototyping: Uses, Challenges, Benefits, and Drawbacks," in IEEE Internet of Things Journal, cilt 10, sayı 16, ss 14397-14417, 2023, DOI: 10.1109/JIOT.2023.3262942.

[14] Amin Biglari ve Wei Tang, "A Review of Embedded Machine Learning Based on Hardware, Application, and Sensing Scheme," Sensors 2023, cilt 23, sayı 4, ss 2131, 2023, DOI: https://doi.org/10.3390/s23042131

[15] Khaled Osmani ve Detlef Schulz, "Comprehensive Investigation of Unmanned Aerial Vehicles (UAVs): An In-Depth Analysis of Avionics Systems," Sensors 2024, cilt 24, sayı 10, ss 3064, 2024, DOI: https://doi.org/10.3390/s24103064

[16] J. Jo, S. Jeong and P. Kang, "Benchmarking GPU-Accelerated Edge Devices," 2020 IEEE International Conference on Big Data and Smart Computing (BigComp), Busan, Korea (South), 2020, ss 117-120, DOI: 10.1109/BigComp48618.2020.00-89.

[17] S. Ullah ve D. -H. Kim, "Benchmarking Jetson Platform for 3D Point-Cloud and Hyper-Spectral Image Classification," 2020 IEEE International Conference on Big Data and Smart Computing (BigComp), Busan, Korea (South), 2020, ss 477-482, DOI: 10.1109/BigComp48618.2020.00-21.

BACKGROUND SUBTRACTION-BASED APPROACH FOR HUMAN ACTIVITY RECOGNITION USING MMWAVE RADAR

Doga Nalci*
Faculty of Engineering, Computer Engineering, Gebze Technical University, Kocaeli, Türkiye

Zehra Bilici
Faculty of Engineering, Computer Engineering, Gebze Technical University, Kocaeli, Türkiye

Saliha Yesilyurt
Faculty of Engineering, Computer Engineering, Gebze Technical University, Kocaeli, Turkey and Sivas Cumhuriyet University, Sivas, Türkiye
**(Corresponding author) doganalci@gtu.edu.tr*

ABSTRACT

Human activity recognition (HAR) using millimeter-wave (mmWave) radar data has gained significant attention due to its robustness under challenging environmental conditions. Building on previous work that employed the VIDAR (Video of Radar) approach—transforming raw radar signals into multidimensional grayscale images for classification via CNN and SVM models—this study proposes a background subtraction-based enhancement technique to further improve recognition accuracy. Static environmental data were collected across multiple sessions, averaged pixel-wise to create a generalized background model, and subtracted from each action frame to generate motion-salient radar representations. Experiments conducted under identical model configurations reveal an approximate 0.5% increase in classification accuracy compared to the original VIDAR baseline. These findings indicate that environmental noise removal can meaningfully enhance motion features critical for HAR tasks. Future work will investigate the generalizability of the proposed method across diverse settings and neural architectures.

Keywords: mmWave radar, human activity recognition, background subtraction, CNN, SVM, VIDAR, raw radar data.

Introduction

The increasing prominence of millimeter wave (mmWave) radar technology can be attributed to its ability to meet the growing need for contactless, reliable, and non-intrusive sensing solutions across a wide range of application domains. From smart environments and elderly care systems to gesture-based interfaces and indoor security monitoring, mmWave radars offer distinct advantages over conventional methods [1]. Unlike optical sensors or cameras, they function independently of ambient lighting conditions and are largely unaffected by visual obstructions such as smoke, furniture, or low visibility environments [2]. More importantly, as mmWave radars do not generate or store visual imagery,

they inherently uphold user privacy — a concern that continues to gain importance in personal and public monitoring contexts. Additionally, their capability to detect fine-grained motion through non-metallic obstructions makes them especially useful in environments where occlusions are common or where wearing a sensor is impractical or undesired [3, 4]. With their small form factor, low energy consumption, and seamless integration potential with embedded systems, mmWave radars are becoming increasingly relevant in developing intelligent, real-time perception systems. These unique attributes have positioned mmWave radar as a transformative sensing modality, motivating further research into signal enhancement and feature extraction techniques that can fully leverage the richness of the data they produce [5].

Millimeter wave (mmWave) radar technology has gained significant traction in recent years, particularly in the domain of non-contact human activity recognition. Operating within the 30 to 300 GHz frequency range of the electromagnetic spectrum, mmWave radar systems leverage short wavelengths to achieve high spatial and velocity resolution. These systems transmit electromagnetic waves and analyze the returned reflections to estimate range, Doppler velocity, and angle of arrival [6]. The resulting multi-dimensional signal structure provides a rich source of information that is well-suited for characterizing and distinguishing between complex human motions. Furthermore, unlike vision-based systems, mmWave radars are robust to poor lighting, occlusions, and privacy concerns, making them ideal for indoor human monitoring applications [7].

A mmWave radar system operates based on the emission and subsequent reflection of high-frequency electromagnetic signals. These signals are transmitted toward a target and reflected back to the radar receiver. By analyzing the time delay, frequency shift, and phase difference between the transmitted and received signals, the radar system estimates the target's range, velocity, and angle of arrival. This tri-dimensional measurement capability forms the core of advanced radar sensing and enables rich information extraction for applications such as motion recognition [8].

Most modern mmWave radars utilize Frequency Modulated Continuous Wave (FMCW) modulation schemes. In FMCW systems, the radar continuously transmits a signal whose frequency linearly varies over time (known as chirping). The frequency difference between the transmitted signals provides an accurate measure of the range. The Doppler shift caused by the relative motion of a target enables velocity estimation, while the phase information across antenna arrays is used for angle estimation through direction-of-arrival algorithms [9].

The radar architecture typically incorporates multiple transmit () and receive () antennas to support MIMO (Multiple Input Multiple Output) configurations. For instance, a 3 × 4 setup can generate up to 12 virtual antennas, enhancing angular resolution and target separation. The antenna configuration allows for the generation of high-resolution radar cubes that capture spatial, angular, and temporal characteristics of targets [10].

To extract meaningful information, the analog echoes are digitized through high-speed Analog-to-Digital Converters (ADCs). The digital signals are then transformed into the frequency domain using Fast Fourier Transforms (FFT). A range-FFT reveals distance

information, followed by Doppler-FFT for velocity estimation [11]. To ensure efficient and lossless data acquisition, a computer with at least an Intel i7 processor and 16 GB of RAM is required. Under typical indoor laboratory conditions, the system can effectively capture motion information within a 6–8 meter range.

Despite these advantages, mmWave radar systems are subject to several operational limitations that must be addressed for reliable deployment in real-world environments. One of the primary challenges is the presence of environmental reflections, particularly from metallic or reflective surfaces, which give rise to multipath interference [12]. These unintended reflections can distort the original signal and degrade the quality of extracted motion features. Moreover, static background elements often introduce consistent clutter into the radar returns. This background clutter may appear as dominant, non-varying signal components that persist across multiple frames, thereby obscuring subtle motion cues and reducing the effectiveness of activity classification models, especially in fine-grained recognition tasks [13]. To address these issues, it becomes essential to implement robust signal preprocessing techniques. In this study, raw radar data-collected and stored using the previously described hardware setup-undergo a background subtraction process designed to suppress environmental artifacts and enhance motion-focused representations. The details of this preprocessing methodology are discussed in Section 2.

Method

Dataset and Preprocessing

In this study, we used a custom-built dataset designed to support human activity recognition tasks using raw Frequency-Modulated Continuous Wave (FMCW) mmWave radar data. The radar platform employed for data acquisition was the Texas Instruments IWR1843BOOST radar evaluation module in conjunction with the DCA1000EVM capture card. All data were collected and stored using TI's mmWave Studio software, which allows low-level access to the radar's intermediate frequency (IF) signal streams.

Radar Configuration and Data Collection

The radar configuration was kept fixed throughout the entire acquisition process. The radar system configuration used throughout this study is presented in Table 1.

Table 1: Radar Configuration Parameters

Parameter	Value
Number of RX antennas	4
Number of TX antennas	3
mmWave Frequency	77- 81 Ghz
ADC Sample Recording Rate	1 Gbps
Number of Samples Per Class	500
Total Samples in Dataset	2000
Number of Action Classes	4

Each radar capture consisted of complex-valued I/Q signals, producing a 4-channel raw data array per sample. The system operated in a static indoor environment, and each motion was recorded at approximately the same location and orientation relative to the

radar. A dedicated recording session was also conducted to capture "background-only" data, without any motion in the scene. The dataset used in this study was derived from the research of Nalci [14].

Raw Data Processing with GNU Octave

The raw binary .bin files generated by mmWave Studio were decoded using a custom script written in GNU Octave [15]. This script parsed the I/Q samples and reshaped them into a 4D matrix with the following Equation 1:

Shape: $N_{Rx} \times N_{frames} \times N_{chirps} \times N_{ADC} = [4 \times 1 \times 129 \times 256]$ (1)

For each action instance, only the real channels were used without separating virtual antenna data. These complex matrices were then converted to their magnitude representation using the formula in Equation 2:

$$Magnitude = \sqrt{(I^2 + Q^2)} \quad (2)$$

This conversion preserves signal energy while eliminating phase information, which was not utilized in this study.

Generation of VIDAR Images

To visualize and structure the raw radar data for input into machine learning models, we employed a custom grayscale transformation method we term VIDAR. The magnitude values from the four antennas were concatenated sequentially and reshaped into a square 2D image matrix. If the number of values did not yield a perfect square, the excess elements were discarded. For example; 4 antennas x 129 chirps x 256 ADC samples = 132.096 values; Largest square dimension $\sqrt{132096} \cong 363$.

363×363, with 108 leftover values discarded. This approach not only standardized the input size but also facilitated the use of 2D convolutional neural networks (CNNs), which are optimized for grid-like data representations.

Dual-Range Normalization

A unique challenge in radar data is that magnitude values derived from I/Q signals can retain information about positive and negative Doppler shifts, reflecting the direction of motion. To preserve this property in the grayscale domain, we applied a dual-range normalization scheme. Negative-valued signals were scaled into the range [0–127]. Positive-valued signals were scaled into the range [128–255]. This normalization preserves polarity information within the image intensity space while ensuring compatibility with standard 8-bit image representations used in CNN input pipelines.

Background Subtraction Method

One of the key limitations of radar-based human activity recognition systems is the presence of residual environmental noise caused by static objects and multipath reflections. These background elements often persist across frames and can be misinterpreted by classification models as part of the motion signature. To address this issue, we imple-

mented a background subtraction technique that explicitly removes the static component of the radar signal before classification.

Generalized Background Acquisition

To construct a reliable estimate of the background environment, we conducted multiple radar recordings in the exact same location where the activity data was collected. During these sessions, the environment remained static—no human motion occurred. Each background instance was recorded with the same radar parameters as the action recordings and processed through the VIDAR pipeline described in Section 2.2.

At least 10 separate background captures were collected to ensure statistical robustness and temporal stability. Each capture was converted into a grayscale square image representing the magnitude of radar reflections in that frame.

Pixel-Wise Averaging

After collecting the background-only VIDAR images, a generalized background image was computed by taking the pixel-wise arithmetic mean across all background instances:

$$Background_{avg}(i,j) = \frac{1}{N}\sum_{k=1}^{N} I_k(i,j) \quad (3)$$

Where ; $I_k(i,j) I_k(i,j)$ denotes the pixel value at location $(i,j)(i,j)$ in the k-th background image, and N is the total number of background samples as shown in Equation 3.

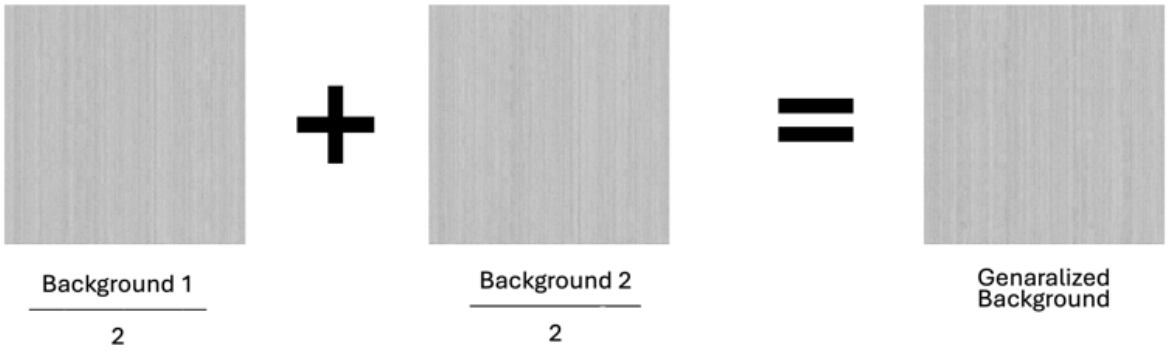


Figure 1: The generalized background image is generated by computing the average signal value at each pixel location across multiple frames and writing it to the corresponding index in the output image.

As illustrated in Figure 1, a generalized background image is generated by computing the pixel-wise average across multiple radar frames. For each pixel, the average signal value is assigned to the corresponding position in the output image. This process captures the static characteristics of the environment and is later used for background subtraction in subsequent motion frames.

This averaging process mitigates minor sensor noise or transient reflections while preserving persistent background patterns. The resulting generalized image represents a stable estimate of static environmental reflections.

Subtraction from Action Images

As demonstrated in Figure 2, a background-subtracted image is obtained by performing pixel-wise subtraction between each motion frame and the pre-computed background reference. The resulting subtracted representation emphasizes dynamic motion features while suppressing static environmental noise, enabling clearer distinction of human activities such as the "backhand" gesture.

For each action instance in the dataset, the generalized background image was subtracted pixel-wise from the corresponding VIDAR image, using the formula in Equation 4:

$$Cleaned(i, j)=Action_{raw}(i, j)-Background_{avg}(i, j) \quad (4)$$

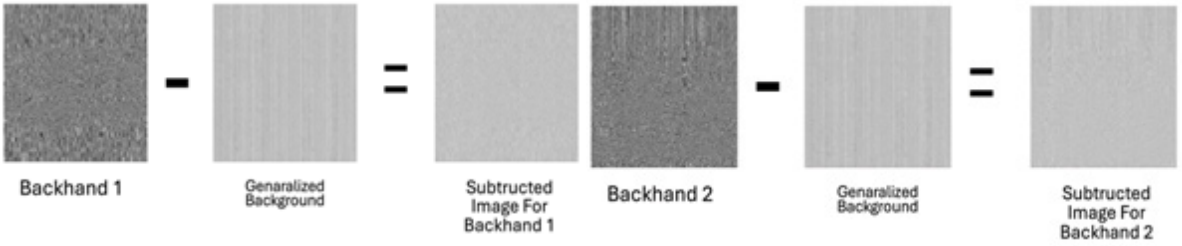


Figure 2: The generalized background image is generated by computing the average signal value at each pixel location across multiple frames and writing it to the corresponding index in the output image. A background-subtracted image is generated by applying pixel

The subtraction was performed in float32 precision before re-normalizing the result to the grayscale range [0–255]. Negative values resulting from the subtraction were clipped to 0. This process effectively removed static environmental features and retained only the dynamic signal components corresponding to human movement.

Impact on Image Quality and Model Performance

The cleaned images exhibited visibly reduced background clutter and more clearly defined motion contours. This was especially apparent in classes involving finer movements such as forehand and backhand swings. The background subtraction also appeared to suppress inter-class ambiguity caused by spatial noise patterns.

As discussed in Section 4, classification experiments showed a consistent improvement of approximately 0.5% in overall accuracy when using background-subtracted images across both CNN and SVM models. This confirms that reducing non-motion signal components helps improve model discriminability.

Model Architecture

Support Vector Machine (SVM)

The SVM classifier was used as a non-parametric, kernel-based machine learning model to establish a baseline performance level [16]. For each sample, the square grayscale VIDAR image was flattened into a one-dimensional feature vector. No handcrafted features or dimensionality reduction techniques were applied; the model was trained directly on the full pixel intensity vector. Configuration of SVM model is shown in Table 2.

Although lacking the feature learning capabilities of deep architectures, SVM remains effective in low-data scenarios and provides interpretable baseline performance, particularly for well-separated classes such as background and serve.

Table 2: SVM Configuration

Parameter	Value
Input	Flattened grayscale VIDAR image (e.g., 363×363 131.769 features)
2	7
Kernel	Radial Basis Function (RBF)
Regularization	'scale' (1 / number of features)
Gamma	1

VGG16 Transfer Learning

To harness the representational power of deep convolutional features, we employed the VGG16 architecture, pre-trained on the ImageNet dataset [17]. Given the difference between photographic and radar imagery, we froze all convolutional layers and fine-tuned only the final dense layers. This allowed the model to retain generic low-level features (e.g., edge detectors) while adapting its high-level decision boundaries to the radar domain. Configuration of CNN model is Shown in Table 3.

Model configuration:

Table 3: CNN Network Configuration Details

Parameter	Value
Input	224×224 pixels VIDAR image
Epoch	5
Pre-trained weights	ImageNet
Batch Size	128
Frozen layers	All convolutional blocks
Trainable parameters	1,608,076 of 16,322,764
Optimizer	Adam
Learning rate	0.00001
Loss function	Categorical cross-entropy
Output layer	4 units (Softmax)
Training Platform	Google Colaboratory
Training Time	10 Hours
Validation strategy	4-fold cross-validation
Loop Over K-Fold	77

Despite the domain gap, VGG16 demonstrated promising performance, especially when trained on background-subtracted data, suggesting that even domain-general features can be effective once environmental noise is reduced. To evaluate the effectiveness of the proposed background subtraction method, two classification models were employed: a classical Support Vector Machine (SVM) and a transfer learning model based on VGG16. Both models were trained on grayscale radar images in VIDAR format, which were generated from complex-valued mmWave radar signals as described in Section II.B-3.

The use of these two complementary approaches enables the comparative analysis of both traditional and deep learning paradigms under identical preprocessing and data conditions.

Experimental Results

In order to assess the effectiveness of the proposed background subtraction method, both the SVM and VGG16 models were trained and evaluated under two distinct scenarios:

Baseline (without background subtraction): VIDAR images generated directly from the original radar signals.

Enhanced (with background subtraction): VIDAR images processed by subtracting the generalized background image, as described in Section III.B.

Each model was trained and validated using 4-fold cross-validation, and the average classification accuracy across all folds was reported. The performance was evaluated for four motion categories: background, forehand, backhand, and serve. Following the application of background subtraction, the VGG16 model achieved a test accuracy of 92.7%, a clear improvement over the baseline (92.1%). The corresponding test confusion matrix as shown in Equation 5:

$$\begin{bmatrix} 21 & 0 & 0 & 0 \\ 0 & 20 & 0 & 1 \\ 0 & 2 & 18 & 1 \\ 0 & 1 & 1 & 18 \end{bmatrix} \quad (5)$$

The training accuracy reached approximately 94%, with its confusion matrix as shown in Equation 6:

$$\begin{bmatrix} 119 & 0 & 0 & 0 \\ 0 & 116 & 4 & 0 \\ 0 & 18 & 102 & 0 \\ 0 & 3 & 1 & 118 \end{bmatrix} \quad (6)$$

These results show that misclassifications were consistent across training and testing, particularly between forehand and backhand classes—indicating that model limitations stem from inherent radar signal ambiguities, rather than overfitting. Additionally, the SVM model, while less expressive, also exhibited a minor improvement from 86.48% to 86.59% after background subtraction.

Figure 3 illustrates the evolution of performance metrics over 1400 training epochs. Accuracy, precision, recall, and F1-score all demonstrate a consistent upward trend, while Mean Squared Error (MSE) continuously decreases. The test accuracy curve remains aligned with training accuracy throughout, reflecting a stable and generalizable training process.

These findings support the hypothesis that background subtraction enables models to better isolate motion-relevant signal patterns. However, it is important to note that the

method does not remove the physical background itself. Instead, it attenuates static signal components by subtracting an average of background radar responses. Since radar signal strength varies independently of absolute background presence, this process primarily cleans non-motion (idle) periods, rather than directly amplifying movement signals.

This clarification has important implications: the observed performance gain stems from eliminating clutter in non-motion segments rather than boosting action-specific signal features. This insight establishes a basis for more detailed temporal segmentation in future studies—particularly in recordings where multiple actions occur with idle intervals in between. The consistent and stable performance shown across models also suggests that the proposed method improves both signal clarity and training dynamics in radar-based activity recognition tasks.

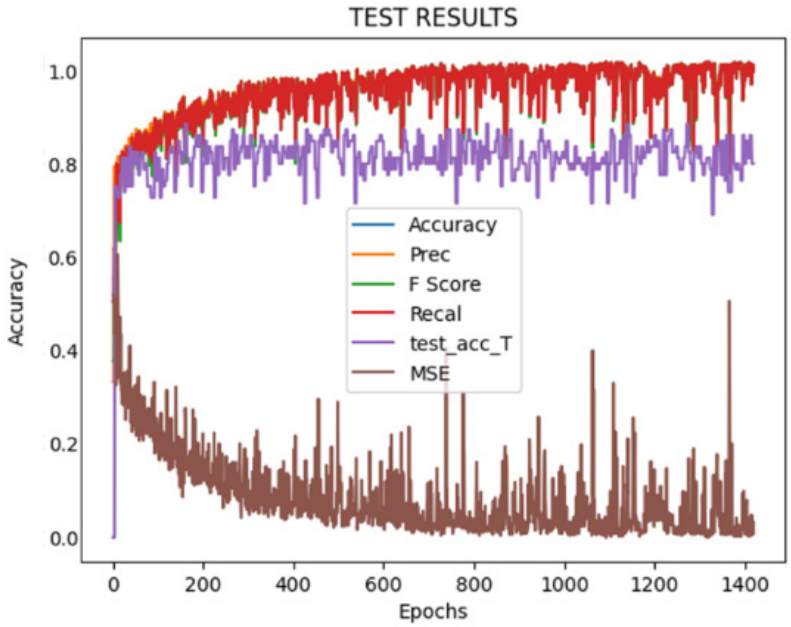


Figure 3: Evolution of performance metrics during training. Accuracy-related scores rise steadily, while MSE declines, suggesting strong generalization with low overfitting risk.

Discussion

The experimental results demonstrate that applying background subtraction to VIDAR images can lead to consistent improvements in human activity classification accuracy across different model architectures. This supports the idea that suppressing static signal artifacts—often present in indoor radar recordings—helps reduce class ambiguity, especially between dynamic motion classes like forehand and backhand.

However, a crucial clarification is warranted: the proposed background subtraction method does not perform semantic background removal in the sense of eliminating physical scene elements. Rather, it statistically suppresses radar energy reflections that remain stable over time. Since mmWave radar signal strength varies due to object geometry, reflectivity, and multipath propagation, the generalized background image acts as a coarse approximation of static environmental influence. When subtracted, it primarily reduces residual energy in non-motion intervals.

This leads to a nuanced interpretation of our results: the accuracy improvement may not result from stronger identification of the motion itself, but from cleaner segmentation between active and idle phases of each action. In other words, the model becomes less likely to misclassify background clutter during periods of inactivity that occur between motion bursts. This distinction is especially relevant for continuous activity recognition tasks, where transitions between motion and stillness must be clearly captured.

In addition to suppressing static environmental components, the proposed background subtraction method also mitigates the influence of *quasi-static* or recurring motion patterns that are irrelevant to the classification task. For instance, continuously rotating fan blades, swaying tree leaves in the wind, or the periodic movement of analog clock hands (e.g., minute and hour hands) may be persistently detected by the radar system as motion. However, such background dynamics do not carry semantic meaning related to human activity and instead contribute *distracting signal noise* to the classifier. In this context, the generalized background model effectively attenuates these non-informative motion sources, thereby enhancing the model's ability to detect *symbolic human actions*. Thus, the proposed approach improves the reliability of radar-based HAR systems not only by eliminating static clutter but also by filtering out behaviorally irrelevant motion artifacts from the classification pipeline.

Future Work

Future studies will aim to explore the temporal structure of motion sequences by analyzing radar signal changes across distinct phases of a single action (e.g., pre-movement, peak-movement, post-movement). Segmenting actions into temporally meaningful slices may help evaluate how background subtraction affects classification accuracy at specific moments. Additionally, the method will be extended to longer recordings where different actions occur sequentially, separated by idle periods, allowing the removal of noise from non-movement intervals. To test robustness and generalizability, future datasets will be collected in multiple environments. Finally, applying this preprocessing technique to transformer-based and other advanced deep learning architectures—such as ResNet, EfficientNet, or attention mechanisms—will allow comparative performance analysis and may yield further improvements in radar-based activity recognition.

Conclusion

This study presented a background subtraction-based preprocessing method for improving human activity recognition using raw mmWave radar data. By generating generalized background representations from multiple static captures and subtracting them from action samples, the proposed method effectively reduced noise in non-motion segments. Experiments with both SVM and VGG16 models showed consistent performance improvements, with VGG16 achieving a test accuracy of 92.7%. The method also contributed to a more stable and interpretable training process, as evidenced by parallel improvements in training and validation metrics. These findings suggest that background subtraction is a lightweight yet impactful enhancement for radar-based classification pipelines and provides a promising direction for future research on temporal segmentation and advanced deep learning architectures.

These findings demonstrate that background subtraction is a viable, lightweight enhancement strategy for radar-based activity recognition systems. Moreover, the results motivate further investigation into the temporal dynamics of radar signals and provide a foundation for evaluating the method's generalizability across different environments and model architectures.

REFERENCES

[1] A. Singh, S. Sandha, and L Garcia, "Radhar: Human activity recognition from point clouds generated through a millimeter-wave radar," dl.acm.orgAD Singh, SS Sandha, L Garcia, M SrivastavaProceedings of the 3rd ACM Workshop on Millimeter-wave Networks and Sensing, 2019•dl.acm.org, vol. 19, pp. 51–56, Oct. 2019, doi: 10.1145/3349624.3356768.

[2] Y. Jang, I. Jeong, M. Y. Heravi, S. Sarkar, H. Shin, and Y. Ahn, "Multi-Camera-Based human activity recognition for Human–Robot collaboration in construction," Sensors, vol. 23, no. 15, p. 6997, Aug. 2023, doi: 10.3390/s23156997.

[3] F. Zeng, M. Guo, L. Tan, F. Guo, and X. Liu, "Wearable Sensor-Based residual Multifeature fusion shrinkage networks for human activity recognition," Sensors, vol. 24, no. 3, p. 758, Jan. 2024, doi: 10.3390/s24030758.

[4] Z. Wang, S. Huan, L. Wu, Q. Wang, J. Liu, and Z. Hu, "Attention-based Vision Transformer for Human Activity Classification using mmWave Radar," ACM International Conference Proceeding Series, pp. 128–134, Nov. 2022, doi: 10.1145/3577164.3577184.

[5] A. K. Alhazmi et al., "Intelligent Millimeter-Wave System for Human Activity Monitoring for telemedicine," Sensors, vol. 24, no. 1, p. 268, Jan. 2024, doi: 10.3390/s24010268.

[6] Z. Bilici, I. Aytutuldu, Y. Genc, and Y. S. Akgul, "mmWave Frequency Modulated Continuous Wave Radar-based Human Action Recognition," 32nd IEEE Conference on Signal Processing and Communications Applications, SIU 2024 - Proceedings, 2024, doi: 10.1109/SIU61531.2024.10600787.

[7] I. Aytutuldu, Z. Bilici, Y. Genc, and Y. S. Akgul, "Indoor Human Counting Based on Range-Doppler Spectrum in the Presence of Material Penetration and Occlusion," 2024 6th International Conference on Communications, Signal Processing, and their Applications, ICCSPA 2024, 2024, doi: 10.1109/ICCSPA61559.2024.10794279.

[8] D. Nalci and Y. S. Akgul, "Human Action Recognition with Raw Millimeter Wave Radar Data," HORA 2022 - 4th International Congress on Human-Computer Interaction, Optimization and Robotic Applications, Proceedings, 2022, doi: 10.1109/HORA55278.2022.9800009.

[9] V. C. Chen, F. Li, S. S. Ho, and H. Wechsler, "Micro-doppler effect in radar: Phenomenon, model, and simulation study," IEEE Trans Aerosp Electron Syst, vol. 42, no. 1, pp. 2–21, Jan. 2006, doi: 10.1109/TAES.2006.1603402.

[10] M. Shaban, "Design and Modeling of a Reconfigurable Multiple Input, Multiple Output Antenna for 24 GHz Radar Sensors," Modelling 2025, Vol. 6, Page 2, vol. 6, no. 1, p. 2, Jan. 2025, doi: 10.3390/MODELLING6010002.

[11] J. Bhatia et al., "Classification of Targets Using Statistical Features from Range FFT of mmWave FMCW Radars," Electronics 2021, Vol. 10, Page 1965, vol. 10, no. 16, p. 1965, Aug. 2021, doi: 10.3390/ELECTRONICS10161965.

[12] K. Qian, L. Yao, X. Zhang, and T. N. Ng, "MilliMirror: 3D Printed reflecting surface for millimeter-wave coverage expansion," Proceedings of the Annual International Conference on Mobile Computing and Networking, MOBICOM, pp. 15–28, Oct. 2022, doi: 10.1145/3495243.3517024.

[13] B. Yan, P. Wang, L. Du, X. Chen, Z. Fang, and Y. Wu, "mmGesture: Semi-supervised gesture recognition system using mmWave radar," Expert Syst Appl, vol. 213, p. 119042, Mar. 2023, doi: 10.1016/J.ESWA.2022.119042.

[14] Doğa Nalcı, "Action recognition using microwave radar sensors," Gebze Technical University, Kocaeli, 2023.

[15] "GNU Octave." https://octave.org/, Accessed: Apr. 20, 2025. [Online].

[16] C. Cortes, V. Vapnik, and L. Saitta, "Support-vector networks," Machine Learning 1995 20:3, vol. 20, no. 3, pp. 273–297, Sep. 1995, doi: 10.1007/BF00994018.

[17] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," 3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings, Sep. 2014, Accessed: Apr. 17, 2025. [Online]. Available: https://arxiv.org/abs/1409.1556v6

AUTOMATED RETINAL DISEASE CLASSIFICATION FOR EARLY DETECTION USING MACHINE LEARNING MODELS

Abdul Samad*
Electrical-Electronics Engineering, Faculty of Engineering, Sakarya University, Serdivan, Sakarya, Türkiye

Muhammed Kürsad Uçar
Electrical-Electronics Engineering, Faculty of Engineering, Sakarya University, Serdivan, Sakarya, Türkiye
(Corresponding author) abdul.--@ogr.sakarya.edu.tr

ABSTRACT

Retinal disorders, including Diabetic Retinopathy (DR), Glaucoma, and Cataract, are major contributors to preventable blindness worldwide, with early stages often being asymptomatic. Timely detection using retinal fundus images can significantly reduce vision loss, but manual diagnosis remains resource-intensive and not scalable. This study presents an automated binary classification framework to categorize retinal images as either Normal or Diseased, without differentiating between specific pathologies. By employing feature selection techniques and various machine learning algorithms, the study aims to streamline the diagnostic process and improve its scalability for primary care and teleophthalmology. The proposed method integrates image preprocessing, including noise removal and contrast enhancement using CLAHE, followed by feature extraction based on statistical measures. Spearman correlation was applied for feature selection to identify highly correlated variables. Machine learning models, such as Decision Trees, Random Forest, K-Nearest Neighbors, and a Hybrid model, were tested for classification performance. The results indicate that the Ensemble model achieves the highest accuracy and robustness, with an accuracy rate of 90.44%, sensitivity of 94.88%, and AUC of 0.904. This approach offers an efficient and scalable solution for initial retinal disease screening, making it feasible for widespread use in healthcare settings.

Keywords: Eye disease, glaucoma, cataract, early diagnosis, machine learning.

Introduction

Retinal disorders contribute significantly to preventable blindness worldwide [1, 2]. Among them, Diabetic Retinopathy (DR), Glaucoma, and Cataract are common and often asymptomatic in preliminary stages [3]. Timely detection using retinal fundus images can help reduce vision loss, but manual examination is resource-intensive and not scalable.

Machine learning offers a promising approach to automate the diagnostic process. This study focuses on a binary classification framework, categorizing retinal images into

Normal or Diseased, regardless of the specific pathology. This simplifies the classification problem and creates a scalable model for initial screening in primary care environments.

In this study, a classical machine learning approach has been applied to the images, utilizing feature selection and extraction methods. This approach aims to efficiently distinguish between normal and diseased conditions, ensuring faster detection that is more feasible for use in initial care environments.

Related Works

Recent advancements in artificial intelligence, particularly deep learning models, have accelerated progress in detecting eye diseases. Convolutional neural networks (CNNs) and transfer learning architectures have been extensively explored to enhance accuracy and clinical use.

In recent years, significant research efforts have been devoted to enhancing the accuracy and reliability of artificial intelligence models for eye disease detection, with a particular focus on Convolutional Neural Networks and transfer learning techniques. In research [4], presents a comprehensive review of commonly adopted CNN architectures such as VGG16, ResNet, and Inception for retinal disease classification. While these models yielded promising outcomes, the review also identified critical challenges including limited availability of annotated datasets, overfitting, and poor generalization to diverse populations, all of which hinder optimal performance in real-world clinical settings.

In another study [5], reported an artificial intelligence model that achieved 98% accuracy on a dataset comprising over 65,000 retinal images, outperforming well-established models such as ResNet, GoogLeNet, and InceptionV3. Similarly, research [6], which proposed an enhanced version of SqueezeNet, attained 98.9% and 98.1% accuracy through cross-validation datasets, respectively.

In another study [7], a model demonstrated 92.37% accuracy, a kappa score of 0.878, and an AUC of 0.987. The fusion of decisions from multiple models in this approach helped mitigate classification bias, resulting in enhanced reliability for multi-disease recognition. Study [8] introduced the DPLA-Net model, which achieved an average accuracy of 94.3% and AUC values exceeding 0.98 for conditions such as retinal detachment and intraocular tumors. Moreover, this model significantly improved the diagnostic performance of junior ophthalmologists, elevating their accuracy from 69.6% to 91.9%.

Research in [9] applied CNNs to classify Optical Coherence Tomography images into four categories: Normal, Diabetic Macular Edema (DME), Choroidal Neovascularization (CNV), and Age-related Macular Degeneration (AMD). Despite the challenges posed by limited datasets, the model exhibited strong classification capabilities. Another study [10] demonstrated that even with a small dataset of 372 grayscale images, a system could achieve 99.99% accuracy for glaucoma and 98.99% for diabetic retinopathy, underscoring the potential of lightweight and efficient architectures in data-constrained scenarios.

Finally, a comprehensive review in [11] examined the integration of CNNs and Vision Transformers (ViTs) for detecting eye diseases such as diabetic retinopathy, AMD, and

glaucoma. The review highlighted ongoing challenges and future directions, emphasizing the importance of improving model interpretability, ensemble learning strategies, and robustness to facilitate clinical adoption.

In this study, we improved image quality using noise removal and CLAHE. We then used 24 methods to extract features from the images. To fix class imbalance, we balanced the dataset. After that, we used Spearman correlation to select important features and reduce processing time. Finally, we used a mix of machine learning models such as Ensemble, KNN, and Decision Tree and hybrid, to make predictions. This method gave better results and detected diseases faster than deep learning models due to feature selection.

3. Material and Method

This study was carried out following the steps presented in Figure 1. Initially, a noise removal filter was applied, followed by contrast enhancement using the CLAHE technique. After preprocessing, feature extraction and selection were performed based on highly correlated features using the spearman feature selection algorithm. In the final step, machine learning algorithms were implemented to perform disease diagnosis.

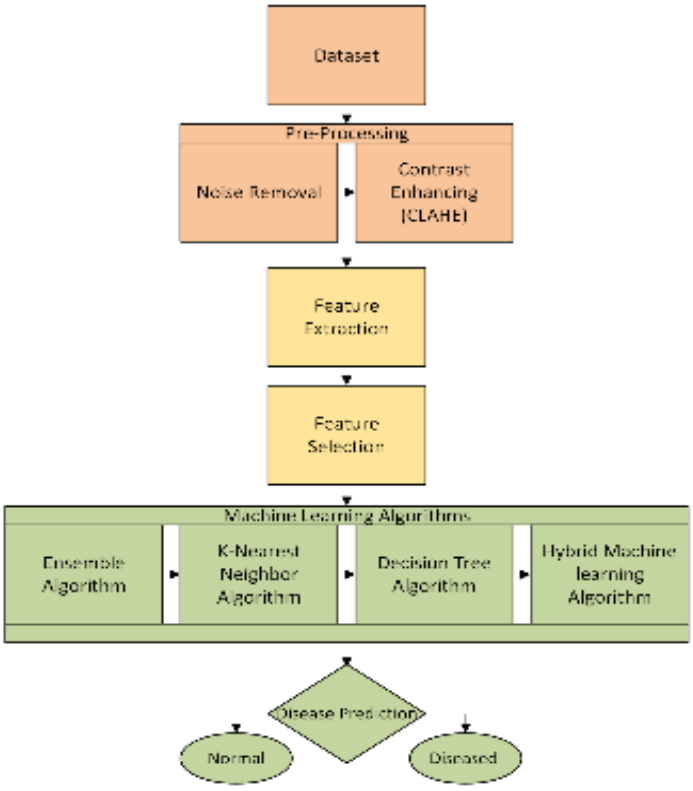


Figure 1: Flow Diagram

Dataset

This study utilizes retinal fundus images obtained from an open source [12] platform. Each dataset consists of high-resolution images accompanied by corresponding diagnostic labels. For classification purposes, all disease types, Diabetic Retinopathy, Glaucoma, and Cataract, were grouped into a single category labeled Diseased, while healthy images were labeled as Normal. The distribution of the data is summarized in Table 1.

Table 1: Data Distribution

Normal	Diseased	Total
1074	3143	4217

Image Processing

In preparation for the classification of retinal fundus images, a structured sequence of preprocessing operations was conducted, as outlined in Figure 3. These procedures play a pivotal role in enhancing image quality, suppressing noise, and highlighting diagnostically significant features [13]. The preprocessing pipeline was designed to address common issues inherent in raw fundus images, such as uneven illumination, low contrast, and background artifacts. By refining image clarity and ensuring consistency across the dataset, these steps facilitate the extraction of meaningful patterns and structures such as vascular architecture, and retinal lesions that are critical for accurate diagnosis. Moreover, this preprocessing stage serves to standardize the input data, thereby improving the robustness and generalizability of the classification model. Overall, the application of these techniques is fundamental to optimizing the downstream performance of automated diagnostic systems in ophthalmic imaging.

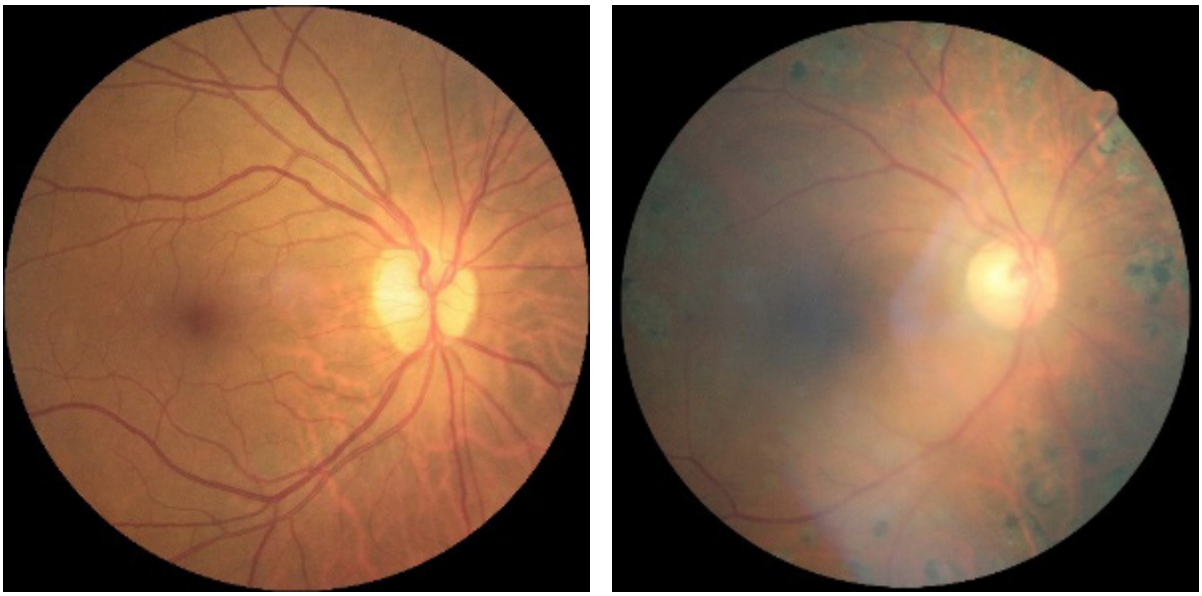


Figure 2: Normal retinal and diseased ımage

Noise Removal

To improve image quality and support precise feature extraction, all color images were initially transformed into their grayscale equivalents, thereby reducing computational complexity and eliminating color-related variance [14, 15]. Subsequently, a two-dimensional Gaussian blur filter, defined by a 5 x 5 convolution kernel and a standard deviation σ, was convolved with each image to attenuate high-frequency noise components and achieve background smoothing. The choice of a 5 x 5 kernel represents a balance between sufficient noise suppression and retention of essential structural details.

$$G(x,y) = \frac{1}{2\pi\sigma^2} exp\left(-\frac{x^2+y^2}{2\sigma^2}\right)$$

Equation 1: Gaussian smoothing kernel equation

In Equation 1, x and y denote spatial offsets relative to the center of the kernel, and σ regulates the degree of blurring. Following convolution, the filtered image exhibits markedly reduced high-frequency variations, which predominantly correspond to noise and small-scale artifacts [16].

This preprocessing stage is critical for downstream tasks: by homogenizing the image background and suppressing spurious pixel-level fluctuations, the Gaussian blur filter enhances the signal-to-noise ratio of relevant anatomical features. Consequently, subsequent operations such as edge detection, morphological analysis, and quantitative feature extraction can be performed with greater reliability and accuracy, thereby improving the overall performance of the machine learning based diagnostic pipeline.

Contrast Enhancement Using CLAHE

Following noise suppression, Contrast Limited Adaptive Histogram Equalization (CLAHE) was applied to augment local contrast and thereby enhance the visibility of diagnostically relevant anatomical structures. CLAHE operates by partitioning the image into a lattice of non-overlapping contextual regions (tiles), within which histogram equalization is performed independently [17]. The resultant contrast adjustment within each tile accentuates subtle features such as retinal microvasculature and optic disc delineations that may be inconspicuous in the original image.

A distinguishing attribute of CLAHE is its clip-limit parameter, which constrains the maximal amplification of the local histogram [18, 19]. By imposing this threshold, the algorithm mitigates the risk of over-enhancement attributable to dominant intensity values and suppresses the amplification of residual noise. Consequently, CLAHE yields enhanced contrast without introducing artefactual intensity spikes that could compromise subsequent analysis [20].

Upon completion of tile-based equalization, bilinear interpolation is employed to merge adjacent tiles seamlessly, thereby obviating boundary artefacts and ensuring spatial continuity across the image. Collectively, these operations produce a uniformly enhanced image that optimizes the delineation of relevant morphological features while preserving overall image integrity facilitating more robust feature extraction and downstream machine learning based diagnostic modeling [21].

Figure 3: Preprocessing steps

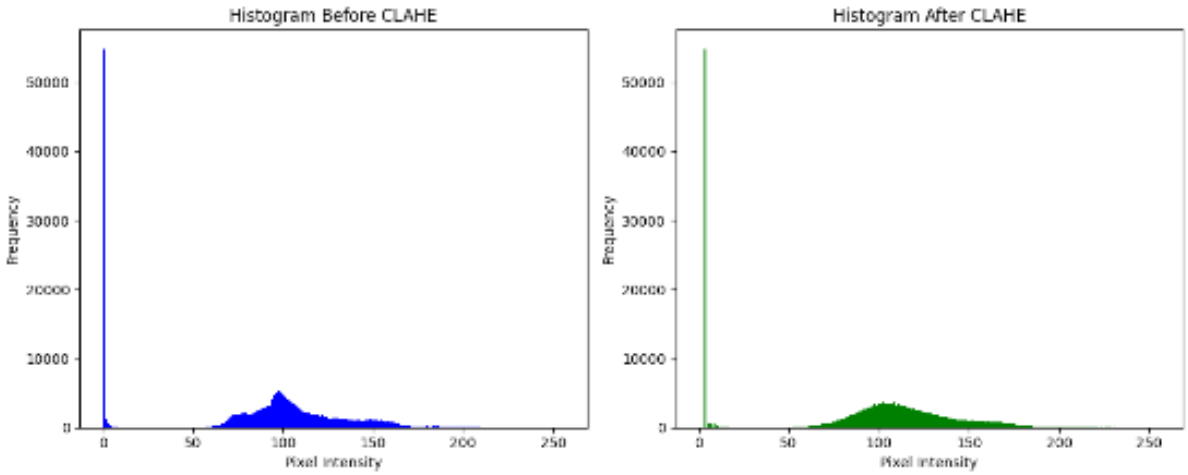


Figure 4: Intensity histogram comparison before and after CLAHE

As shown in Figure 3, CLAHE significantly improves the contrast of retinal images. Furthermore, Figure 4, displays the intensity histogram comparison before and after CLAHE, demonstrating how pixel intensities are more evenly distributed after enhancement.

Feature Extraction

In this study, 24 time-domain features were extracted from the preprocessed retinal images. These features were selected based on their relevance to disease classification and their ability to capture essential characteristics of the image. The extraction process involved calculating various statistical measures, such as mean, standard deviation, skewness, and kurtosis, among others. The specific rules and formulas used to extract these features are summarized in Table 2.

Table 2: Formulas and rules for extracting time-domain features

No	Features	Formulas
1	Mean	$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} = \frac{1}{n}(x_1 + \cdots + x_n)$
2	Standard Deviation	$S = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{x})^2}$

3	Range	$Range = \max(x) - \min(x)$
4	Hjorth Parameter: Activity	$A = S^2$
5	Maximum	$x_{max} = max(x_i)$
6	Minimum	$x_{min} = min(x_i)$
7	Median	$\tilde{x} = \begin{cases} x_{\frac{n+1}{2}} & : x \text{ tek} \\ 1/2\,(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}) & : x \text{ çift} \end{cases}$
8	Geometric Mean	$G = \sqrt[n]{x_1 \times \cdots \times x_n}$
9	Harmonic Mean	$H = n / \left(\frac{1}{x_1} + \cdots + \frac{1}{x_n}\right)$
10	Root Mean Square (RMS)	$X_{rms} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} \lvert x_i \rvert^2}$
11	Skewness	$x_{ske} = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^3}{(n-1)S^3}$
12	Kurtosis	$x_{kur} = \frac{\sum_{i=1}^{n} (x(i) - \overline{x})^4}{(n-1)S^4}$
13	Singular Value Decomposition (SVD)	$SVD = svd(x)$
14	25% Trimmed Mean	$T25 = trimmean(x, 25)$
15	50% Trimmed Mean	$T50 = trimmean(x, 50)$
16	Shannon Entropy	$H = -\sum_{i=1}^{n} p_i \log_2(p_i)$
17	Average Energy	$E = \frac{1}{n}\sum_{i=1}^{n} x_i^2$
18	Hjorth Parameter: Mobility	$M = S_1^2 / S^2$
19	Hjorth Parameter: Complexity	$C = \sqrt{(S_2^2 / S_1^2)^2 - (S_1^2 / S^2)^2}$
20	Interquartile Range (IQR)	$IQR = iqr(x)$
21	Mean or Median Absolute Deviation	$MAD = mad(x)$
22	Total Intensity	$\text{Sum} = \sum_{i=1}^{n} x_i$
23	Average Teager Energy	$TE = \frac{1}{n}\sum_{i=3}^{n} (x_{i-1}^2 - x_i x_{i-2})$
24	Coefficient of Variation (CV)	$DK = (S / \overline{x})100$

Feature Selection

In this study, Spearman correlation was employed as a feature selection technique to identify variables that exhibit significant monotonic relationships with the target outcome [22]. As a non-parametric method, it effectively captures both linear and non-linear associations, making it suitable for datasets that may not meet normality assumptions. Features with high correlation to the target variable and low inter-feature redundancy were retained to improve model generalizability and reduce complexity. This approach not only enhanced the interpretability of the model but also contributed to better predictive performance by eliminating irrelevant or less informative attributes.

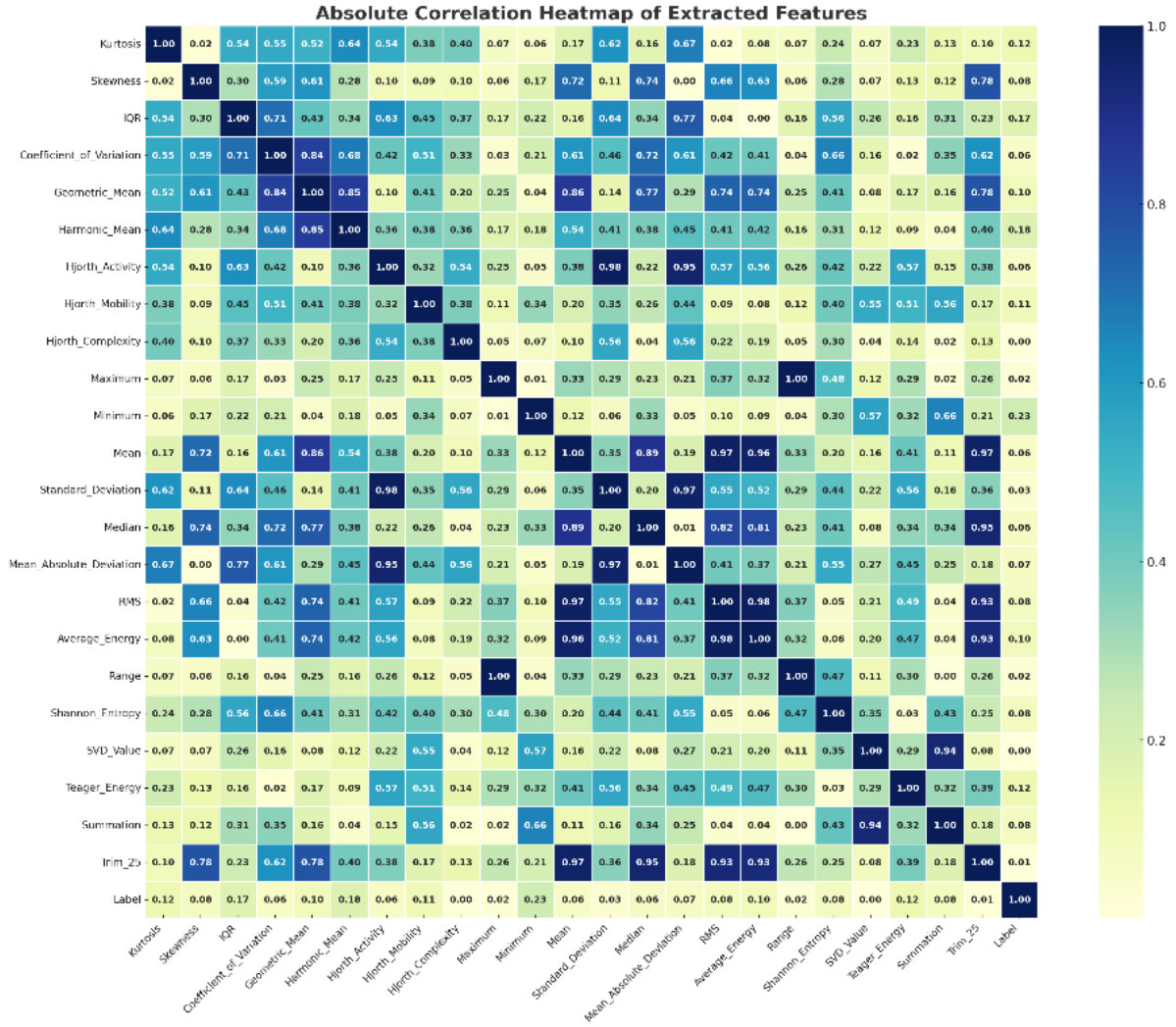


Figure 5: Absolute correlation heatmap of extracted features.

To further illustrate the degree of interdependence among the extracted features, an absolute correlation heatmap was generated Figure 5. Features showing high correlation were considered candidates for removal to improve model generalization and reduce computational complexity.

Machine Learning Algorithm

Machine learning plays a vital role in automated medical image analysis by enabling systems to learn diagnostic patterns from data. In this work, Decision Tree (DT) and k-Nearest Neighbors (k-NN) algorithms were applied for classifying retinal fundus images. To enhance performance, an ensemble approach was used, and a hybrid model was developed to combine the strengths of individual classifiers, resulting in improved accuracy and robustness.

Decision Tree:

Decision Tree algorithm is a non-parametric supervised learning method used for both classification and regression tasks [23, 24]. It models decisions in a tree-like structure by recursively splitting the dataset based on feature values that maximize information gain [25]. Decision Trees are interpretable and computationally efficient but can be prone to overfitting, especially with complex datasets.

Ensemble Methods:

Ensemble learning combines predictions from multiple base models to produce a more accurate and robust outcome [26]. Techniques such as Random Forest (bagging) and Gradient Boosting (boosting) are widely used. Random Forest builds multiple Decision Trees on bootstrapped subsets and averages their results, enhancing generalization and reducing variance [27-29]. Gradient Boosting builds models sequentially, with each model correcting the errors of its predecessor, thus improving performance. These methods are known for their high predictive accuracy and ability to handle complex, nonlinear relationships.

K-Nearest Neighbors (KNN):

KNN is a non-parametric, instance-based learning algorithm used for classification and regression. It predicts the output class of a sample by analyzing the majority class among its "k" closest neighbors in the feature space, typically using distance metrics such as Euclidean or Manhattan distance [30-32]. KNN is simple and effective for low-dimensional data but can be computationally intensive and sensitive to feature scaling and noise in high-dimensional spaces [33-36].

Hybrid Algorithm:

A hybrid machine learning algorithm combines multiple models or techniques to leverage the strengths of each, improving overall performance and robustness [37, 38]. By integrating different approaches, the hybrid model can address limitations of individual algorithms, such as overfitting or computational inefficiency [39], while enhancing predictive accuracy and generalization. Common hybrid strategies include combining ensemble methods with base classifiers, such as Decision Trees, or using different types of learners like K-Nearest Neighbors (KNN) with boosting techniques [40-42]. These hybrid models are particularly effective in handling complex datasets and non-linear relationships, offering improved results in classification and regression tasks [43].

Performance Evaluation Criteria

To assess the performance of the proposed models, a comprehensive set of evaluation metrics was utilized. These included the accuracy rate (Acc), sensitivity (Se), specificity (Sp), F-measure (F_O), Cohen's Kappa coefficient, and the area under the Receiver Operating Characteristic (ROC) curve (AUC). Each of these metrics provides valuable insights into the model's diagnostic capabilities.

Table 3. Performance Evaluation Results

Features		Model	Acc	Sen	Spe	F_O	Kapp	AUC
7		**EN**	**89.97**	**0.94**	**0.85**	**0.89**	**0.79**	**0.89**
DT		85.78	0.89	0.82	0.85	0.71	0.85	
kNN		83.44	0.96	0.70	0.81	0.66	0.83	
HML		88.11	0.97	0.78	0.87	0.76	0.88	
14		EN	88.81	0.94	0.83	0.88	0.77	0.88
DT		87.64	0.93	0.81	0.87	0.75	0.87	
kNN		86.24	0.94	0.78	0.85	0.72	0.86	
HML		88.57	0.98	0.78	0.87	0.77	0.88	
22		**EN**	**90.44**	**0.94**	**0.85**	**0.90**	**0.80**	**0.90**
DT		88.57	0.93	0.83	0.88	0.77	0.88	
kNN		84.61	0.86	0.82	0.84	0.69	0.84	
HML		86.94	0.96	0.77	0.85	0.73	0.86	
All	23	EN	89.04	0.95	0.82	0.88	0.78	0.89
		DT	87.17	0.93	0.81	0.86	0.74	0.87
		kNN	85.54	0.90	0.80	0.85	0.71	0.85
		HML	89.04	0.97	0.80	0.88	0.78	0.89

Result

The primary objective of this study was to assess the effectiveness of various machine learning algorithms in classifying retinal images into "Normal" and "Diseased" categories. The dataset comprised 1074 normal and 3143 diseased retinal images, from which a total of 24 distinct features were extracted. Spearman correlation was utilized for feature reduction, leading to the development of four distinct classification models, each utilizing different subsets of features: 7, 14, 22, and all features.

Classification performance was evaluated using Decision Trees (DT), K-Nearest Neighbors (KNN), Ensemble methods (EN), and a Hybrid Machine Learning (HML) model. The detailed results of these analyses are summarized in Table 3, revealing that all models attained accuracies exceeding 80%. Notably, the Ensemble model demonstrated superior performance with an accuracy of 90.44%, sensitivity of 94.88%, specificity of 85.98%, and an Area Under the Curve (AUC) of 0.9043, indicating robust classification capability.

The Decision Tree (DT) model achieved an accuracy of 88.57% and a sensitivity of 93.95%, whereas the K-Nearest Neighbors (KNN) model exhibited slightly lower accuracy 86.95% but demonstrated high sensitivity 96.74%. Additionally, the Hybrid Machine Learning model, which integrates multiple classifiers, delivered a balanced performance with an

accuracy of 89.04%, sensitivity of 97.67%, and AUC of 0.8902, suggesting good stability and reliable predictive capability.

The results highlight the positive correlation between the number of selected statistical features and improved classification performance. This indicates the feasibility and effectiveness of feature selection combined with machine learning models provide a scalable solution for the initial screening and accurate diagnosis of retinal diseases.

Discussion

This study demonstrates that traditional machine learning models, when combined with effective feature extraction and selection, can classify retinal images as normal or diseased with high accuracy. The best performance was achieved by the Ensemble model using 22 selected features, reaching 90.44% accuracy and 94.88% sensitivity. This shows that well-chosen statistical features, through effective feature extraction and selection, can be very useful for classifying medical images by reducing computational cost, prediction time, and memory load, while achieving accuracy above 85%.

Compared to deep learning models in previous studies, such as [5] and [7], which reported over 98% accuracy using CNNs with large datasets, our model offers a simpler and faster alternative. Deep models often need extensive annotated data, high computing power, and face issues like overfitting. In contrast, our approach requires fewer resources and works well even with limited data. This study highlights that using just 22 well-selected features, classical machine learning can achieve results close to deep learning models, making it a practical solution for early disease screening, especially in primary care or remote areas.

Conclusion

This study presents an efficient and scalable machine learning framework for the classification of retinal diseases using fundus images. By integrating feature selection techniques and various machine learning algorithms, including Decision Trees, K-Nearest Neighbors, and Hybrid models, the proposed approach achieved robust results. The Ensemble model demonstrated the highest performance, with an accuracy of 90.44%, sensitivity of 94.88%, and AUC of 0.904, making it a strong candidate for early retinal disease detection. This method provides a practical solution for primary care and teleophthalmology, offering accurate diagnostics with fewer resources compared to deep learning-based approaches. The results emphasize the potential of classical machine learning models in medical image classification, demonstrating their effectiveness in real-world healthcare applications.

Declaration of Interest: The authors declare that there is no conflict of interest.

Acknowledgements: The authors would like to acknowledge that no external support or contributions were received during this work.

REFERENCES

[1] A. Ben-Arzi, R. Ehrlich, and R. Neumann, "Retinal Diseases: The Next Frontier in Pharmacodelivery," Pharmaceutics 2022, Vol. 14, Page 904, vol. 14, no. 5, p. 904, Apr. 2022, doi: 10.3390/PHARMACEUTICS14050904.

[2] S. Muchuchuti and S. Viriri, "Retinal Disease Detection Using Deep Learning Techniques: A Comprehensive Review," Journal of Imaging 2023, Vol. 9, Page 84, vol. 9, no. 4, p. 84, Apr. 2023, doi: 10.3390/JIMAGING9040084.

[3] D. Wang, J. Lian, and W. Jiao, "Multi-label classification of retinal disease via a novel vision transformer model," Front Neurosci, vol. 17, p. 1290803, Jan. 2023, doi: 10.3389/FNINS.2023.1290803/BIBTEX.

[4] G. Ramanathan, D. Chakrabarti, A. Patil, S. Rishipathak, and S. Kharche, "Eye Disease Detection Using Machine Learning," in 2021 2nd Global Conference for Advancement in Technology, GCAT 2021, Institute of Electrical and Electronics Engineers Inc., Oct. 2021. doi: 10.1109/GCAT52182.2021.9587740.

[5] M. Khalid, M. Z. Sajid, A. Youssef, N. A. Khan, M. F. Hamid, and F. Abbas, "CAD-EYE: An Automated System for Multi-Eye Disease Classification Using Feature Fusion with Deep Learning Models and Fluorescence Imaging for Enhanced Interpretability," Diagnostics, vol. 14, no. 23, Dec. 2024, doi: 10.3390/diagnostics14232679.

[6] A. Zia, R. Mahum, N. Ahmad, M. Awais, and A. M. Alshamrani, "Eye diseases detection using deep learning with BAM attention module," Multimed Tools Appl, vol. 83, no. 20, pp. 59061–59084, Jun. 2024, doi: 10.1007/s11042-023-17839-9.

[7] F. Du et al., "Recognition of eye diseases based on deep neural networks for transfer learning and improved D-S evidence theory," BMC Med Imaging, vol. 24, no. 1, Dec. 2024, doi: 10.1186/s12880-023-01176-2.

[8] X. Ye et al., "Ocular Disease Detection with Deep Learning (Fine-Grained Image Categorization) Applied to Ocular B-Scan Ultrasound Images," Ophthalmol Ther, Oct. 2024, doi: 10.1007/s40123-024-01009-7.

[9] M. Elkholy and M. A. Marzouk, "Deep learning-based classification of eye diseases using Convolutional Neural Network for OCT images," Front Comput Sci, vol. 5, 2023, doi: 10.3389/FCOMP.2023.1252295.

[10] V. Pol, V. Waghmare, R. Shinde, and Prof. R. R. Suryavanshi, "Eye Diseases Detection using ML," Int J Res Appl Sci Eng Technol, vol. 12, no. 5, pp. 3602–3605, May 2024, doi: 10.22214/ijraset.2024.62403.

[11] S. Muchuchuti and S. Viriri, "Retinal Disease Detection Using Deep Learning Techniques: A Comprehensive Review," Apr. 01, 2023, MDPI. doi: 10.3390/jimaging9040084.

[12] "eye_diseases_classification." Accessed: Apr. 22, 2025. [Online]. Available: https://www.kaggle.com/datasets/gunavenkatdoddi/eye-diseases-classification

[13] M. Sachin Sonawane, M. Awasthy, and N. Choubey, "A Literature Review on Image Processing and Classification Techniques for Agriculture Produce and Modeling of Quality Assessment system for Soybean industry Sample," Print) www.arcjournals.org International Journal of Innovative Research in Electronics and Communications, vol. 6, no. 2, pp. 8–16, 2019, doi: 10.20431/2349-4050.0602002.

[14] M. P. Deisenroth and H. Ohlsson, "A general perspective on Gaussian filtering and smoothing: Explaining current and deriving new algorithms," Proceedings of the American Control Conference, pp. 1807–1812, 2011, doi: 10.1109/ACC.2011.5990871.

[15] J. P. F. D'Haeyer, "Gaussian filtering of images: A regularization approach," Signal Processing, vol. 18, no. 2, pp. 169–181, Oct. 1989, doi: 10.1016/0165-1684(89)90048-0.

[16] A. K. Naik, G. Kumar, P. K. Upadhyay, P. Date, and A. K. Singh, "Gaussian Filtering for Simultaneously Occurring Delayed and Missing Measurements," IEEE Access, vol. 10, pp. 100746–100762, 2022, doi: 10.1109/ACCESS.2022.3208119.

[17] olcay palta, "IMAGE ENHANCEMENT IN WHITE BLOOD CELLS USING HYBRID (CHALE+GAUSS) FILTER AND FLASK-BASED WEB SERVICE INTEGRATION," 7. INTERNATIONAL MEDITERRANEAN CONGRESS, Jan. 2025, Accessed: Apr. 26, 2025. [Online]. Available: https://www.academia.edu/127302951/IMAGE_ENHANCEMENT_IN_WHITE_BLOOD_CELLS_USING_HYBRID_CHALE_GAUSS_FILTER_AND_FLASK_BASED_WEB_SERVICE_INTEGRATION

[18] P. Musa, F. Al Rafi, and M. Lamsani, "A review: Contrast-limited adaptive histogram equalization (CLAHE) methods to help the application of face recognition," Proceedings of the 3rd International Conference on Informatics and Computing, ICIC 2018, Oct. 2018, doi: 10.1109/IAC.2018.8780492.

[19] B. Soni and P. Mathur, "An improved image dehazing technique using CLAHE and guided filter," 2020 7th International Conference on Signal Processing and Integrated Networks, SPIN 2020, pp. 902–907, Feb. 2020, doi: 10.1109/SPIN48934.2020.9071296.

[20] R. Sharma and A. Kamra, "A Review on CLAHE Based Enhancement Techniques," Proceedings of International Conference on Contemporary Computing and Informatics, IC3I 2023, pp. 321–325, 2023, doi: 10.1109/IC3I59117.2023.10397722.

[21] B. Hatipoğlu, Prof. Dr. İ. Karagöz, and M. İnal, "Threshold Based Image Enhancement Method for Low Contrast X-Ray Images Using CLAHE," International Journal of Engineering Research and Development, vol. 14, no. 3, pp. 343–350, Dec. 2022, doi: 10.29137/UMAGD.1203617.

[22] M. K. Uçar, K. Uçar, Z. Uçar, and M. R. Bozkurt, "Determination gender-based hybrid artificial intelligence of body muscle percentage by photoplethysmography signal," Comput Methods Programs Biomed, vol. 224, p. 107010, Sep. 2022, doi: 10.1016/J.CMPB.2022.107010.

[23] P. P. Harsh H. Patel, "Study and Analysis of Decision Tree Based Classification Algorithms," vol. 6, no. 10, pp. 74–78, 2018.

[24] A. Navada, A. N. Ansari, S. Patil, and B. A. Sonkamble, "Overview of use of decision tree algorithms in machine learning," Proceedings - 2011 IEEE Control and System Graduate Research Colloquium, ICSGRC 2011, pp. 37–42, 2011, doi: 10.1109/ICSGRC.2011.5991826.

[25] Y. Lu, T. Ye, and J. Zheng, "Decision Tree Algorithm in Machine Learning," 2022 IEEE International Conference on Advances in Electrical Engineering and Computer Applications, AEECA 2022, pp. 1014–1017, 2022, doi: 10.1109/AEECA55500.2022.9918857.

[26] U. Sarmah, P. Borah, and D. K. Bhattacharyya, "Ensemble Learning Methods: An Empirical Study," SN Comput Sci, vol. 5, no. 7, pp. 1–32, Oct. 2024, doi: 10.1007/S42979-024-03252-Y/TABLES/15.

[27] H. Şevgin, "A comparative study of ensemble methods in the field of education: Bagging and Boosting algorithms," International Journal of Assessment Tools in Education, vol. 2023, no. 3, pp. 544–562, 2023, doi: 10.21449/ijate.1167705.

[28] D. Opitz and R. Maclin, "Popular Ensemble Methods: An Empirical Study," Journal of Artificial Intelligence Research, vol. 11, pp. 169–198, 1999.

[29] I. D. Mienye and Y. Sun, "A Survey of Ensemble Learning: Concepts, Algorithms, Applications, and Prospects," IEEE Access, vol. 10, pp. 99129–99149, 2022, doi: 10.1109/ACCESS.2022.3207287.

[30] Z. Zhang, "Introduction to machine learning: k-nearest neighbors," Ann Transl Med, vol. 4, no. 11, p. 218, Jun. 2016, doi: 10.21037/ATM.2016.03.37.

[31] M. Bansal, A. Goyal, and A. Choudhary, "A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning," Decision Analytics Journal, vol. 3, p. 100071, Jun. 2022, doi: 10.1016/J.DAJOUR.2022.100071.

[32] R. S. A. Daulay, S. Efendi, and Suherman, "Review of Literature on Improving the KNN Algorithm," Transactions on Engineering and Computing Sciences, vol. 11, no. 3, pp. 63–72, Jun. 2023, doi: 10.14738/TECS.113.14768.

[33] K. Taunk, S. De, S. Verma, and A. Swetapadma, "A brief review of nearest neighbor algorithm for learning and classification," 2019 International Conference on Intelligent Computing and Control Systems, ICCS 2019, pp. 1255–1260, May 2019, doi: 10.1109/ICCS45141.2019.9065747.

[34] M. Suyal and P. Goyal, "A Review on Analysis of K-Nearest Neighbor Classification Machine Learning Algorithms based on Supervised Learning," International Journal of Engineering Trends and Technology, vol. 70, pp. 43–48, 2022, doi: 10.14445/22315381/IJETT-V70I7P205.

[35] R. K. Halder, M. N. Uddin, M. A. Uddin, S. Aryal, and A. Khraisat, "Enhancing K-nearest neighbor algorithm: a comprehensive review and performance analysis of modifications," Journal of Big Data 2024 11:1, vol. 11, no. 1, pp. 1–55, Aug. 2024, doi: 10.1186/S40537-024-00973-Y.

[36] R. D. Wicaksono and G. Fajar Shidiq, "Comparison of KNN and CNN Algorithms for Gender Classification Based on Eye Images," Scientific Journal of Informatics, vol. 11, no. 4, pp. 915–924, Dec. 2024, doi: 10.15294/SJI.V11I4.13529.

[37] M. Abdelrahim, C. Merlosy, and T. Wang, "Hybrid Machine Learning Approaches: A Method to Improve Expected Output of Semi-structured Sequential Data," Proceedings - 2016 IEEE 10th International Conference on Semantic Computing, ICSC 2016, pp. 342–345, Mar. 2016, doi: 10.1109/ICSC.2016.72.

[38] F. Zhou, H. Fan, Y. Liu, H. Zhang, and R. Ji, "Hybrid Model of Machine Learning Method and Empirical Method for Rate of Penetration Prediction Based on Data Similarity," Applied Sciences 2023, Vol. 13, Page 5870, vol. 13, no. 10, p. 5870, May 2023, doi: 10.3390/APP13105870.

[39] B. F. Azevedo, A. M. A. C. Rocha, and A. I. Pereira, "Hybrid approaches to optimization and machine learning methods: a systematic literature review," Mach Learn, vol. 113, no. 7, pp. 4055–4097, Jul. 2024, doi: 10.1007/S10994-023-06467-X/TABLES/3.

[40] D. C. Asogwa, S. O. Anigbogu, I. E. Onyenwe, and F. A. Sani, "Text Classification Using Hybrid Machine Learning Algorithms on Big Data," International Journal of Trend in Research and Development, vol. 6, no. 5, pp. 2394–9333, Mar. 2021, Accessed: Apr. 23, 2025. [Online]. Available: https://arxiv.org/abs/2103.16624v1

[41] R. Jahan and S. Alam, "Improving Classification Accuracy Using Hybrid Machine Learning Algorithms on Malaria Dataset," Engineering Proceedings 2023, Vol. 56, Page 232, vol. 56, no. 1, p. 232, Nov. 2023, doi: 10.3390/ASEC2023-15924.

[42]M. R. Machado and S. Karray, "Applying hybrid machine learning algorithms to assess customer risk-adjusted revenue in the financial industry," Electron Commer Res Appl, vol. 56, p. 101202, Nov. 2022, doi: 10.1016/J.ELERAP.2022.101202.

[43] P. Sokkhey and T. Okazaki, "Hybrid machine learning algorithms for predicting academic performance," International Journal of Advanced Computer Science and Applications, vol. 11, no. 1, pp. 32–41, 2020, doi: 10.14569/IJACSA.2020.0110104.

BENLİĞİN DİJİTAL DÖNÜŞÜMÜ, YAPAY ZEKÂ VE YALNIZLIK

Nilgün Özcan*
Gençlik ve Spor Bakanlığı, Türkiye
**(İletişimden sorumlu yazar) nilgunguney@gmail.com*

ÖZET

İnsan aklını aşan bir hızla ilerleyen teknoloji, pek çok şey gibi insanın benlik algısını da dönüştürmektedir. Bu dönüşüm, dijital teknolojilerin kimlik, kişisel algı ve öz yönetimi yeniden şekillendirmesi olarak ifade edilebilir. Geleneksel benlik tanımları, öz farkındalık ve sosyal etkileşimler üzerine kuruluyken, dijital çağda benlik; sosyal medya profilleri, yapay zekâ etkileşimleri ve sanal kimlikler ile şekillenmektedir. Peki, bu dönüşüm insanı yalnızlaştırıyor mu? Varoluşundan uzaklaştırıyor mu? Yoksa yeni bir sosyalleşme biçimi mi yaratıyor? Literatürde dijitalleşmenin olumsuz etkilerine ilişkin araştırmaların özellikle yalnızlık, sosyal izolasyon ve benlik algısı çerçevesinde geliştiği görülmektedir. Yapay zekânın kurucu ontolojik varsayımı, zekânın bireye özgü bir kavram olduğu yönündedir [1]. Bu çalışma, benlik ve yalnızlık ilişkisi, benliğin dijital dönüşümü, toplumsal bağlar, geleneksel ve modern benlik anlayışları gibi alt başlıklarla yapay zekâ ve yalnızlık arasındaki ilişkiyi güncel akademik çalışmalar ışığında incelemeyi amaçlamaktadır. Yapay zekâ programları ile idealize edilmiş bir kurguda, yaşadığı toplumla bağ kuramayan, benlik bütünlüğünden yoksun bireyler ait olma duygusundan uzaklaşarak yalnızlığa sürüklenmektedir. Teknolojik çözümler yalnızlık problemi ile başa çıkmak için kullanılabilmektedir ancak yüzeysel bir çözüm imkânı sunmaktadır.

Anahtar kelimeler: Benlik, dijital dönüşüm, kimlik yapay zekâ ve benlik, varoluşçuluk.

Giriş

Yapay zekâ felsefesi literatürde yeni olmakla birlikte hızla gelişen bir alandır. Yapay zekâ; desenleri tanıma, karar alma, sorunları çözme veya dili anlama gibi genellikle insan zekâsı gerektiren görevleri gerçekleştirmek için bilgisayar sistemlerinin ve algoritmaların kullanılması olarak tanımlanmaktadır [2,3]. İnsan benzeri zekâ ile çalışabilen ve tepki verebilen makineler yaratmayı amaçlayan bilgisayar bilimi alanda yüz tanıma ve oyun oynama gibi faaliyetleri içerir [4]. İnsan benzeri zekâ ile çalışabilen ve tepki verebilen makineler yaratmayı amaçlayan bilgisayar bilimi gelecekte ortaya çıkabilecek problemleri öngörecek şekilde düzenlenmelidir. Sohbet robotları, sanal uygulamalar gibi programlar aracılığıyla oluşan yapay arkadaşlıklar tüm yaş gruplarında yalnızlıkla mücadele dayanıklılığı sağlamaktadır [5]. Zekâyı bilgi işleme olarak gören geleneksel anlayış, günümüzde zayıflamıştır ve teknolojik varsayımlardan ziyade içsel yeniden yapılanma, zekâ için daha iyi bir ölçüt olarak görülmektedir. Zekânın ölçütü, bilgi işlemekten ziyade içsel yeniden yapılanma yoluyla ontolojik genişlemeye ulaşma kapasitesidir. Bu açıdan yaklaşıldığında insan zihninin ürünü olan yapay zekâların, insan varoluşu üzerindeki etkisi daha önemli hale gelmektedir [6].

Yapay zekalara, insan varlığının evrim sürecinin bir parçası olarak yaklaşıldığında; yaşamın zirvesi, varoluşun şifresini çözen akıl sahibi bir varlık olarak insan iken, günümüzde insanın yerini, ekolojik krizden korkmayan, evrendeki diğer gezegenleri ve belki de diğer evrenleri keşfedebilen, onun yarattığı robotlar almaya adaydır. Bu çerçevede dünyada baskın türün robotlar olarak değiştiği bir dönem yaşanmaktadır [7].

Benlik Kavramı ve Yalnızlık

Benlik, bireyin kendisine dair inanç, değer ve tutumlarının bütününü ifade eder. Rogers (1951) ve Maslow (1968) gibi hümanist psikologlar, benlik saygısı (self-esteem) ve benlik uyumu (self-congruence) kavramlarını merkeze alarak, bireyin kendini gerçekleştirmesinin psikolojik esenlik için kritik olduğunu vurgulamıştır. Benlik algısı düşük olan bireyler, sosyal ilişkilerde yetersizlik hissi yaşayabilir ve bu durum yalnızlık duygusunu tetikleyebilir [8]. Öte yandan benlik saygısı ile yalnızlık arasında negatif bir korelasyon olduğunu gösteren birçok çalışma bulunmaktadır [9]. Özellikle kendini değersiz hisseden bireylerin, sosyal etkileşimlerden kaçındığı ya da reddedilme korkusuyla derin ilişkiler kuramadığı görülmektedir. Bireyin ben olma bilincini kazanması, sosyal çevresiyle uyumlu bir şekilde benlik algısı geliştirmesini gerektirmektedir. İçinde yaşadığı toplumla bağ kuramayan benlik bütünlüğünden yoksun bireyler, aidiyet hissi eksikliği yaşayarak yalnızlığa daha yatkın hale gelebilmektedir. Yalnızlık, benlik algısını derinden etkileyen bir deneyimdir. Sürekli yalnız kalan bireyler, bu durumu kendi kişisel başarısızlıkları ile ilişkilendirirlerse bu durum benlik saygısının zedelenmesine yol açabilmektedir [10]. Yalnız bireyler, sosyal medya veya çevrelerindeki insanlarla kendilerini kıyaslayarak "eksik" hissetme eğilimindedir [11]. Günümüzde dijital iletişimin artması, benlik ve yalnızlık ilişkisini daha karmaşık hale getirmiştir. Yapay zekânın insan yaşamına girdikçe ortaya çıkan sorunlardan biri de yapay zekâ geliştikçe onun tarafından üretilenler ile insan ürünü şeyler arasındaki farkın idrakinin zorlaşmasıdır. Sosyal medya kanalları, bireylerin "kabul edilme" ihtiyacını karşılarken, buralarda kurulan yüzeysel etkileşim başkaları ile derin bağlar kurulmasını engelleyebilmektedir [12]. Ayrıca, bireyci toplumlarda yalnızlık daha yaygınken, kolektivist kültürlerde benlik algısının grup aidiyetiyle daha güçlü şekillendiği görülmektedir [13].

Benliğin Dijital Dönüşümü

Dijital teknoloji çerçevesinde günümüzde insanların kendilerini algılama biçimleri ve ifade etme şekillerinin kökten bir değişim içerisinde olduğu görülmektedir [13]. *Reclaiming Conversation: The Power of Talk in a Digital Age* başlıklı çalışmasında Turkle (2017), dijital iletişimin insanları yüz yüze etkileşimden uzaklaştırdığını ve benlik algısını zayıflattığını ifade etmektedir. Turkle'a göre, sosyal medya platformları, insanların kendilerini "idealize edilmiş" versiyonlarıyla sunmalarına neden olmakta ve bu durum, bireyin özgün benlik algısını zedelemektedir [12]. Bu durum benliğin nesneleştirilmesi ve nicelleştirilmesi olarak da ifade edilebilir. Davranışları ve kimlikleri dijital mecralarda nicelleştirilen ve analiz edilen bireyler günümüzde veri olarak görülmeye başlanmıştır. Bu nicelleştirme süreci bireylerin aktif, refleksif uygulamalarını ve algoritmalar tarafından yönlendirilen sistemik süreçleri içermektedir.

Felsefe tarihinde Sokrates ve pek çok filozof tarafından vurgulandığını gördüğümüz, insan için en zor deneyimlerden biri olarak bilinen "kendini tanımak", insan benliğini daha karmaşık hale getiren yapay zekâ destekli algoritmalar ile çok daha zor hale gelmiştir,

bireyin özünden uzaklaşması olarak yorumlanabilmektedir. Pariser (2011) The Filter Bubble: What the Internet is Hiding From You başlıklı kitabında, kişiselleştirilmiş içeriklerin insanları kendi "filtre balonlarına" hapsettiğini ve bu durumun benliğin çeşitliliğini sınırladığını öne sürmektedir [13]. Bu durum, insanların kendi inanç ve değerlerini sorgulamalarını engelleyerek, benliğin dijital dönüşümünü olumsuz yönde etkilemektedir.

Yeni bir varoluş biçimi olarak bakıldığında insanın kendini inşa etme yollarından biri olarak görülebilecek dijital iletişim, bireylerin kimliklerini keşfetmeleri ve ifade etmeleri için yeni yöntemler de sunabilmektedir. Sanal evrenlerde inşa edilen ve dijital platformlar ve sosyoteknik sistemler tarafından dışarıdan etkilenen yeni benlik kavramlarının ve kimliklerin yaratılmasına olanak sağlamaktadır [14]. Ayrıca bu imkânlarla dijital hümanizmin desteklendiği de söylenebilir.

Geleneksel ve Modern Benlik Anlayışları

Geleneksel toplumlarda benlik, bireyin aile, din, kültür ve yerel topluluklarla kurduğu bağlar üzerinden tanımlanmaktadır. Freud'un bilinçaltı teorileri, Jung'un kolektif bilinçdışı kavramı ve sosyolojideki "sosyal rol" teorileri, benliği daha çok içsel ve toplumsal süreçlerle açıklamıştır. Benlik, sabit ve derinlikli bir kavram olarak görülmüştür. Öte yandan modernizm ile birlikte benlik kavramı, dijital platformlarda sergilenen bir "performans" haline gelmektedir Bireyler sosyal medya platformlarında farklı benlik versiyonları ile var olmaktadır. Ayrıca sosyal medya uygulamalarındaki algoritmalar bireylerin uygulamada geçmiş kullanım kayıtlarına göre uygun içerikleri göstererek benlik algılarını doğrudan şekillendirmektedir. Turkle (2017) bu durumu; bir yandan sürekli bağlantıda olduğumuz diğer yandan derin ilişkiler kurmakta zorlandığımız birliktelikleri ifade eden "yalnız beraberlik" çağı olarak ifade etmektedir [12].

Dijital dünyanın sunduğu avantajlar birikimli kültür kavramını güçlendirmiştir. Bu kültür, sosyal olarak öğrenilen davranışlar ve insanın bilişsel yeteneklerini geliştiren karmaşık teknolojilerle karakterize edilmektedir. Dijital ortamlar, insan zekâsının ve yaratıcılığının evrimi için gerekli olan kültürel öğrenmeyi ve sosyal referanslamayı desteklemektedir [17]. Teknolojik gelişmelerin kültürel evrimi güçlendirdiği ve buna paralel olarak günlük yaşamda bilgi teknolojilerine olan bağımlılığımızı artırdığı görülmektedir. Öte yandan bu tartışmalarda bilinç de son derece önemlidir. Bilinç olmadan insanların diğerleri ile anlamlı şekilde yaşamalarını sağlayan öznellik söz konusu olamaz. İnsan bilinci, başkalarının bilinciyle gönüllü olarak temas kurmak için bir özveri eylemi olarak uzanan bir bilinçtir. Özneler arasılık karşılıklı duygudaşlık, şefkat, inanç ve arzulara ilişkin çıkarımdan daha fazlasıdır.

Dijital Mecraların Değişimdeki Rolü

Yapay zekâ teknolojilerinin yakın dost ihtiyacını karşılayarak insanların sosyalleşmesini sağlayabildiği gibi yalnızlık duygularını da artırabilmektedir. Yapay zekâ programları ile yaratılan sanal asistanların ve chatbotlar, insanların duygusal ihtiyaçlarını karşılamada yetersiz kalabilmektedir. Bu hali ile teknoloji; insanların gerçek ilişkilerden uzaklaşmasına ve sosyal izolasyon yaşamasına neden olmaktadır. Aynı zamanda, yapay zekaların sunduğu kişiselleştirilmiş öneriler, insanları sürekli bir "optimizasyon" ve "mükemmeliyet" arayışına itmektedir. Bu durum, benlik algısını olumsuz etkileyerek, insanların kendilerini yetersiz hissetmelerine

yol açmaktadır *Digital Sociology* başlıklı çalışmasında Lupton, dijital teknolojilerin insanların kendilerini sürekli olarak başkalarıyla karşılaştırmalarına neden olduğunu ve bu durumun yalnızlık duygularını artırdığını ifade etmektedir [18]. Bireylerin, Instagram'da ideal yaşamlar, LinkedIn'de profesyonel kimlikler, anonim forumlarda ise farklı kişilikler sergilediklerinden hareketle çoklu kimlik oluşturulması ve sonuçlarına ilişkin araştırmalar yapılması ve sonuçlarının öngörülebilir olması gerekmektedir. Örneğin, sanal mecralarda kurulan diyalogların, bireylerin benlik algısına etkileri önemli görülmektedir. Yalnız yaşayan bireylerin bir yapay zekâ asistanıyla sürdürdüğü iletişimin kendisini ifade edebilmesini sağlayarak iletişim ihtiyacını karşıladığı düşünülebilir ancak bu iletişimin gerçek bir insan etkileşiminin yerini tutabildiğini söylemek güçtür. Dijital mecralarda yayınlanan bir düşünce, fotoğraf ya da reels, eğer beğeniliyorsa dijital onay mekanizması olarak benlik saygısını doğrudan etkilemektedir. Bu çerçevede sosyal medyanın, dış onay mekanizmalarına bağımlı bir benlik yapısı oluşturduğu ifade edilebilecektir. Sosyal medya kullanımı yüzeysel, zayıf ilişkilere, benlik odaklı yaşama eğiliminde artışa ve ruhsal sağlık sorunlarına yol açabilmektedir [19].

Benlik ve yalnızlık kavramları, bireyin içsel ve sosyal deneyimlerini anlamada merkezi role sahip olgulardır. Benlik, bireyin kendini tanımlama, değerlendirme ve anlamlandırma süreçlerini kapsarken; yalnızlık, arzu edilen sosyal ilişkiler ile mevcut ilişkiler arasındaki uyumsuzluktan kaynaklanan subjektif bir deneyimdir.

Yalnızlık tezi bağlamında yaklaşıldığında; konuya sanal alemdeki yüzeysel ilişkilerden başlamak gerekmektedir. Dijital etkileşimlerin, çoğu zaman derinlikten yoksun olduğu bilinmektedir. Dolayısıyla sosyal medya kullanımı arttıkça, insanlar yüz yüze iletişimde daha fazla yabancılaşma hissetmektedir [20]. Ayrıca sosyal medyada parlatılmış kurgusal hayatları gören bireylerin kendilerini yetersiz hissetmesi kaçınılmazdır. Bu deneyimler depresyon ve sosyal izolasyonu tetikleyebilmektedir. Global sosyal ağlar ve çeşitli dijital platformlar, coğrafi sınırları aşarak benzer ilgi alanlarına sahip insanları bir araya getirebilmektedir. Özellikle yalnız yaşayan bireyler için yapay zekâ asistanları, "dijital arkadaş" olabilmekte, yalnızlıktan uzaklaştırarak sosyalleşmelerine olumlu katkı sağlayabilmektedir.

SONUÇ

Benliğin dijital dönüşümü, yapay zekâ ve yalnızlık arasındaki ilişki, günümüz toplumlarında en önemli inceleme alanlarından biri durumundadır. Zira dijitalleşme; benliği daha akışkan ve çoğul hale getirirken, yalnızlık ve aidiyet ikilemini de beraberinde getirmektedir. Önemli olan, teknolojiyi insani bağları güçlendiren bir araç olarak kullanabilmektir. Gelecekte, yapay zekâ ve sosyal medya, yalnızlığı azaltacak yeni toplumsal modeller sunabileceği düşünülmektedir. Güncel bir sorun olan varoluşsal krizleri aşabilmek ve dijital evrende kaybolmamak için dijital molalar verebilmeli ve derin insani bağlar kurarak insan ilişkilerinin güçlendirilmesi gerekmektedir. Teknolojiden bir süreliğine uzaklaşmak, insanın içsel dünyasında çeşitli değişimlere yol açabilir. Dijital molaların amacı, insanın dijital dünyadan uzaklaşarak özüne dönmesi, kendisi ile bağ kurması ve gerçek dünyadaki ilişkileri daha fazla önemsemesini sağlamaktır. Bu çaba kısa vadede bireylerde bir boşluk hissi doğursa da uzun vadede özgürleştirici deneyimlere kapı açabilecektir. Öte yandan yapay zekâ uygulamaları insanlar tarafından kullanıldığı sürece, eğitilmektedir. Yapay zekânın dikkatli kullanılması son derece önemlidir. Yapay zekânın insanlığın yararına gelişmesi bu şekilde

sağlanabilecektir. Bir yapay zekâ uygulamasının adilliği veya taraflılığı büyük ölçüde eğitildiği veriler, algoritmaların tasarımı ve nasıl uygulandığı gibi çeşitli faktörlere bağlıdır. Dijital teknolojiler, insanların benlik algısını dönüştürürken, aynı zamanda yalnızlık ve içe kapanma gibi olumsuz sonuçlar doğurmaktadır. Bu sorunların çözümü, yapay zekâ programlarının etik bir yaklaşımla tasarlanması ve insanlar arası bağların güçlendirilmesiyle mümkün olabilir. Yapay zekâ programlarının ve bu alandaki yeniliklerin, insan varoluşu üzerindeki etkisi; insan türünün yüzyıllardır süren varoluşsal evriminin bir aşaması olarak kabul edildiğinde bu konudaki endişelerin azalabileceği düşünülmektedir.

REFERANSLAR

[1] Halpin, H. (2025). Artificial intelligence versus collective intelligence. AI & Soc. https://doi.org/10.1007/s00146-025-02240-x

[2] Morandín-Ahuerma, F. (2022). What is Artificial Intelligence? *International Journal of Research Publication and Reviews*, https://doi.org/10.55248/gengpi.2022.31261

[3] Panesar, A. (2020). What Is Artificial Intelligence? Machine Learning and AI for Healthcare, 1 - 18. https://doi.org/10.1007/978-1-4842-6537-6_1

[4] Coleman, J. (2020). AI and Our Understanding of Intelligence. K. Arai, S. Kapoor, & R. Bhatia (Ed.), Advances in Intelligent Systems and Computing Intelligent Systems and Applications. 1250, p. 183-190. IntelliSys 2020. https://doi.org/10.1007/978-3-030-55180-3_

[5] Kachhi, R. (2024). Combating Lonaliness Using AI. *International Journal of Innovative Research in Computer and Communication Engineering.* https://doi.org/10.15680/ijircce.2024.1203512

[6] Denizhan, Y. (2023). Intelligence as a Border Activity Between the Modelled and the Unmodelled. *Angelaki*, 28(3), 25–37. https://doi.org/10.1080/0969725X.2023.2216542

[7] Maltsev, Y. (2024). The Phenomenon of Artificial Intelligence in the Context of Self-Development and Self-Realization of Being. Philosophical Thought, 7, 16-32. https://doi.org/10.25136/2409-8728.2024.7.70674 EDN: PQTGPW

[8] Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachment as a fundamental human motivation. *Psychological Bulletin*, 117(3), 497–529. https://doi.org/10.1037/0033-2909.117.3.497

[9] Rosenberg, M. (1965). Society and the Adolescent Self-Image. Princeton, NJ: Princeton University Press.

[10] Peplau, L., & Perlman, D. (1982). Perspectives on loneliness. Loneliness: A sourcebook of current theory, research and therapy (p. 1-18). New York: John Wiley & Sons.

[11] Suls, J. M, & Wills, T. (1991). Social comparison: contemporary theory and research. Hillsdale, N.J. : L. Erlbaum Associates.

[12] Turkle, S. (2017). Reclaiming Conversation: The Power of Talk in a Digital Age (V:27). B. Books.

[13] Hofstede, G. (1980). Culture's consequences: International differences in work-related values. Beverly Hills, CA: Sage.

[14] Martín Lucas, J., Muñoz Rodríguez, J., & Pessoa, T. (2020). The OnLife self: CONECT-ID. Construction of Youth Identity in a Hyperconnected World. In Eighth International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'20). (s. 907-908). USA, Association for Computing Machinery. https://doi.org/10.1145/3434780.343670

[15] Pariser, E. (2011). The Filter Bubble: What The Internet Is Hiding From You. UK: Penguin Book.

[16] Brubaker, R. (2020). Digital hyperconnectivity and the self. Theory and Society, 49, 771 - 801. https://doi.org/10.1007/s11186-020-09405-1

[17] Seetharaman, P., Mathew, S., Sein, M., & Tallamraju, R. (2020). Being (more) Human in a Digitized World. Information Systems Frontiers. 1-4. https://doi.org/10.1007/s10796-020-10020-9

[18] Lupton, D. (2018). Digital Sociology: Critical Perspectives. Routledge.

[19] Twenge, j. (2020). Why Increased Social Media Use May Be Linked to Depression. Journal of College and Character, 14(1), 11-20.

[20] Qureshi-Hurst, E. (2022). Anxiety, alienation, and estrangement in the context of social media. Religious Studies, 58(3), 522–533. https://doi.org/10.1017/S0034412521000093

BLOCKCHAIN-ENABLED SMART CONTRACTS FOR SUSTAINABLE MEDICAL WASTE MANAGEMENT: ENHANCING TRACEABILITY, COMPLIANCE, AND EFFICIENCY

Arif Soyler*
İzmir Bakırçay University, Department of Health Management, İzmir, Turkiye

Esra Dundar Aravacık
İzmir Kâtip Çelebi University, Department of Health Management, İzmir, Turkiye

Serhat Burmaoglu
İzmir Kâtip Çelebi University, Department of Health Management, İzmir, Turkiye
(Corresponding author) arifsoyler@gmail.com

ABSTRACT

Medical waste management represents a critical healthcare challenge requiring strict tracking, regulatory compliance, and effective stakeholder coordination to mitigate environmental and public health risks. Traditional systems suffer from transparency gaps, fragmented tracking, and compliance issues. This paper explores the potential of blockchain-based smart contracts to transform medical waste management through automated tracking, compliance assurance, and environmental monitoring. Through a systematic literature review of 87 publications on smart contracts in healthcare, waste, and sustainability domains, we synthesize existing knowledge and identify implementation frameworks for medical waste. Our findings show that smart contracts, integrated with IoT devices, enable immutable audit trails from generation to disposal, automate compliance checks, and support secure stakeholder communication while protecting sensitive data. Blockchain's decentralized structure enhances data integrity and traceability, ensuring adherence to environmental and health regulations. However, challenges such as scalability, security risks, regulatory ambiguity, and integration with existing healthcare systems remain. Based on our analysis, we propose a comprehensive implementation framework (Figure 1) addressing system architecture, stakeholder roles, technical aspects, and governance mechanisms for blockchain-based medical waste solutions. This research contributes a conceptual model for smart contract use in medical waste management that balances technical feasibility with regulatory demands, proposing a socio-technical approach that considers technology, processes, policies, and societal impacts.

Keywords: Blockchain, smart contracts, medical waste management, healthcare, sustainability, automated tracking.

Introduction

Medical waste management represents a critical challenge for healthcare institutions globally, with significant implications for public health, environmental sustainability, and regulatory compliance. The improper handling of medical waste—including infectious, pathological, pharmaceutical, and radioactive materials—poses substantial risks to healthcare workers, waste handlers, communities, and ecosystems [1]. Despite stringent regulations and established protocols, traditional medical waste management systems frequently suffer from transparency deficiencies, fragmented tracking, and compliance uncertainties [2], [3].

The healthcare sector generates approximately 0.5 kg of hazardous waste per hospital bed daily in developed countries and 0.2 kg in developing countries, totaling millions of tons annually worldwide [4]. Conventional waste management approaches rely heavily on manual documentation, intermediary verification, and limited traceability, creating vulnerabilities in the waste management chain. These limitations contribute to inefficiencies, environmental contamination, and potential public health hazards when medical waste is improperly handled or disposed of [5].

Blockchain technology—a distributed, immutable ledger system—and its associated smart contracts have emerged as transformative innovations with the potential to address complex coordination challenges across various industries. Smart contracts, self-executing agreements with terms directly encoded into computer protocols, offer promising capabilities for enhancing transparency, security, and efficiency in transactional processes [6]. Their implementation across supply chain management, finance, energy, and construction sectors has demonstrated significant potential for automation, disintermediation, and enhanced trust in multi-stakeholder environments [7].

The application of blockchain and smart contracts in healthcare contexts has garnered increasing attention, with implementations ranging from electronic health record management to clinical trial data integrity and remote patient monitoring [8]. Similarly, waste management systems have begun exploring blockchain-based solutions to improve tracking, incentivize proper disposal, and enhance sustainability metrics [9]. However, the specific intersection of smart contracts and medical waste management remains largely unexplored in the academic literature, representing a significant research gap despite its considerable potential for addressing sector-specific challenges.

This research aims to bridge this knowledge gap by systematically analyzing how blockchain-based smart contracts can be effectively implemented to address the unique challenges of medical waste management. This paper proposes a comprehensive framework (Figure 1) for implementing blockchain-based smart contracts in medical waste management systems to address these challenges. Specifically, this study investigates the potential for smart contracts to enhance medical waste tracking from point of generation to final disposal, automate compliance verification processes, facilitate secure coordination among diverse stakeholders, and monitor environmental impacts throughout the waste management lifecycle.

The primary research questions guiding this study are:

1. What are the specific challenges in medical waste management that can be effectively addressed through blockchain-based smart contracts?

2.How can smart contracts be designed and implemented to enhance tracking, compliance verification, and stakeholder coordination in medical waste management systems?

3.What technical, regulatory, and operational considerations must be addressed to ensure successful implementation of smart contract solutions in medical waste management contexts?

4.What sustainability impacts—economic, environmental, and social—might result from implementing smart contract solutions in medical waste management?

The remainder of this paper is structured as follows: Section II presents a comprehensive literature review covering smart contract technology, its applications across industries, healthcare implementations, waste management applications, medical waste management considerations, and sustainability aspects. Section III outlines the research methodology. Section IV presents the results of our analysis. Section V discusses the implications and challenges of implementing smart contracts in medical waste management. Finally, Section VI concludes with a summary of key contributions and directions for future research.

Literature Review

Smart Contract Technology

Smart contracts represent a critical component of blockchain technology, enabling automated execution of agreements without intermediaries. They function as self-executing computer programs designed to facilitate, verify, and enforce negotiation and implementation of digital contracts [10], [11]. These digital agreements are designed to be self-enforcing, with embedded terms automatically executed when predetermined conditions are met [12].

At their core, smart contracts are event-driven and stateful, employing protocols and user interfaces to execute all steps of the contracting process [13]. Operating on blockchain networks, they verify and enforce contracts digitally, functioning essentially as programmable transactions [14]. The primary goal of smart contracts is to provide solutions for trust, security, and transparency across various domains [15], [16].

Despite their potential, smart contracts face several implementation challenges. Security vulnerabilities exist due to code complexity, particularly in languages like Solidity, leading to errors and loopholes with potential for significant financial losses [15]. Common security incidents include re-entrancy attacks, demonstrated in high-profile cases like The DAO, where substantial cryptocurrency theft occurred [14].

Privacy concerns arise from blockchain's inherent transparency, as all transaction data is publicly available on the distributed ledger [11]. This fundamental tension between blockchain's transparency and privacy requirements poses challenges for sensitive information management [13]. Additionally, scalability issues emerge from blockchain's decentralized

nature, which inherently leads to slower transaction processing compared to centralized systems [17].

Technical barriers to smart contract adoption include complexity in development, testing, and deployment, creating accessibility challenges for non-technical users [13]. The lack of standardization and regulations leads to interoperability and implementation issues [7]. Smart contracts' immutability, while a key feature, complicates bug fixes or logical updates after deployment [14].

Applications Across Industries

Smart contracts have found diverse applications across multiple industries, each with unique implementation approaches and success metrics. These applications leverage blockchain's features such as decentralization, transparency, and immutability to address specific challenges.

In supply chain management, smart contracts enhance transparency and traceability of goods throughout their journey [7]. They improve visibility by recording each step on the blockchain, ensure authenticity by tracking products from origin to delivery, streamline logistics through inventory management and shipping coordination, and facilitate efficient procurement processes [18]. Implementation approaches in this sector often integrate blockchain with IoT devices such as GPS and RFID for tracking goods [19].

Financial services extensively utilize smart contracts to automate transactions like payments and settlements, enable faster and more secure international payments, manage digital assets, facilitate decentralized finance (DeFi) applications, simplify insurance claims processing, and support crowdfunding initiatives [20]. Implementation approaches include using public blockchains for DeFi and open financial systems, employing smart contracts to enforce financial agreements, creating digital tokens to represent assets and rights, and integrating with existing financial infrastructure through APIs [21].

In the energy sector, smart contracts enable peer-to-peer (P2P) energy trading within microgrids, automate energy billing and payments, facilitate renewable energy integration into the grid, support smart grids and demand response programs, and track carbon emissions to promote environmental sustainability [22]. Implementation typically involves private or permissioned blockchains to manage data access and transactions, smart contracts to automatically execute energy trades and settlements, integration with IoT devices like smart meters for real-time data, and tailored consensus mechanisms for specific use cases [23].

Construction applications include managing contracts and payments through smart contracts, tracking materials and equipment to ensure authenticity and reduce waste, improving stakeholder collaboration, enhancing project management transparency, and automating compliance with building codes [24]. Implementation approaches include using smart contracts to automate milestone-based payments, integrating blockchain with Building Information Modeling (BIM) for data management, and employing IoT devices to collect construction process data [7].

Smart Contracts in Healthcare

Smart contracts have diverse applications in healthcare, primarily aimed at enhancing data management, security, and patient empowerment. These applications leverage blockchain's decentralization, transparency, and immutability to address existing healthcare system challenges.

For health information management, smart contracts facilitate secure and efficient electronic health record (EHR) management, ensuring data privacy, integrity, and availability while addressing fragmented medical records and interoperability challenges [25]. A notable example is MedRec, a decentralized EHR management system using Ethereum smart contracts to manage authentication, confidentiality, accountability, and data sharing, enabling patients to access records across multiple treatment sites [26].

Smart contracts enable patients to control medical data access through multi-signature approvals between patients and providers, ensuring only authorized users can access or append records [27]. Ancile, a framework using Ethereum smart contracts, provides secure, interoperable, and efficient access control to medical records, transferring ownership and control to patients [28].

Remote patient monitoring applications use smart contracts to automate the process by connecting sensors to smart devices that execute contracts transferring records to immutable ledgers [29]. One system using private Ethereum blockchain with a master-slave model for medical device deployment connects sensors to devices interacting with smart contracts to evaluate data and trigger alerts to patients and providers [30].

In clinical trials and research, smart contracts enhance data transparency and enforce regulatory requirements [31]. Regulator contracts and trial contracts act as trusted administrators in Ethereum-based systems, with regulator contracts holding data structures for clinical trial authorization, while trial contracts leverage functions within authorization contracts [32].

Despite these promising applications, healthcare smart contracts face significant challenges regarding compliance, integration, and regulation. A major challenge involves ensuring blockchain-based healthcare systems comply with regulations like HIPAA, which mandates patient data protection through stringent privacy, security, and access control measures [28]. Blockchain's transparent and immutable nature can conflict with these requirements, particularly regarding the right to be forgotten and personal data access control [33].

Waste Management Applications

The literature describes various tracking methods, digital solutions, and operational challenges in waste management systems, encompassing both traditional and modern approaches. Traditional waste management methods often lack operational transparency, traceability, auditability, security, and trusted data provenance [34]. These manual processes lead to several challenges: coordination issues requiring full stakeholder coordination, governance, technical support, and monitoring [35]; transparency deficiencies

in waste collection, transfer, and disposal, complicating inefficiency identification [36]; inefficient manual monitoring of waste bins and disposal sites without real-time data or effective analysis of waste bin placement or disposal locations [37].

Modern waste management approaches have proposed various technical solutions to address these limitations. IoT-based approaches focus primarily on monitoring waste management processes, including bin fill levels and collection [38]. Smart bins use IoT devices to identify locations, track vehicles, and monitor conditions, with sensors identifying bin levels, volume, distance, and weight [39]. Blockchain technology provides authenticity, integrity, and security to waste management solutions, enabling monitoring of waste management processes, including tracking, tracing, and authentication at every step while identifying misleading activities [40].

Smart contracts offer innovative solutions for waste management tracking, compliance monitoring, stakeholder coordination, and documentation automation. Blockchain-based tracking systems monitor waste throughout the management process, providing transparency and identifying misleading activities [41]. These solutions often combine sensors and IoT devices to monitor waste type, quality, location, and shipment, as demonstrated by studies using blockchain and sensors to manage and segregate plastic waste [42].

Smart contracts can enforce waste management regulations through automated collection, transportation, and disposal processes ensuring adherence to standards [43]. Some systems monitor waste management using penalties for non-compliance and rewards for effective handling [44]. Blockchain enables decentralized systems improving coordination among waste management stakeholders [45]. Smart contracts ensure secure communication channels for coordinating home appliances, minimizing bills and reducing carbon footprints [46].

Medical Waste Management Specific Considerations

Medical waste management presents unique challenges compared to general waste management, particularly in tracking, compliance, stakeholder coordination, and environmental concerns. Medical waste requires meticulous tracking due to potential biohazard risks [47]. Unlike general waste, monitoring medical waste from origin to final disposal is crucial, ensuring proper handling, storage, and transportation according to strict guidelines [24]. Blockchain technology has been suggested to improve medical waste traceability [48].

Healthcare operates under strict regulatory frameworks like HIPAA, mandating data privacy and integrity [31]. These mandates extend to medical waste handling and disposal, requiring adherence to protocols often more stringent than general waste standards [49]. Medical waste management involves various stakeholders including hospitals, clinics, waste disposal companies, and regulatory bodies [12]. Coordinating these stakeholders' activities and ensuring smooth communication is essential for effective waste management [50].

Medical waste often contains hazardous materials including sharps, infectious substances, and pharmaceuticals, posing higher environmental risks than general waste [18]. Improper handling can lead to soil and water contamination and public health risks, necessitating proper disposal methods and strict guideline adherence to mitigate these risks [51].

Smart contract solutions address the unique challenges of medical waste management, particularly in automated tracking, compliance verification, and environmental impact monitoring. Smart contracts, often combined with IoT devices, enable real-time medical waste tracking, monitoring movement from origin to disposal [29]. Sensors monitor waste bins with data recorded on blockchains, offering transparency and accountability [52]. Smart contracts automatically enforce regulatory compliance by verifying condition fulfillment before transaction execution [25]. For example, contracts can verify waste disposal company permits before payment, preventing improper medical waste handling [53].

Sustainability Aspects

Several papers analyze economic aspects of smart contract implementation, focusing on cost-benefit analyses, resource optimization, and long-term viability, employing various metrics and evaluation methods to assess economic impact across industries.

Smart contracts reduce costs by automating processes and eliminating intermediaries [54]. In construction, they automate payments, reducing transaction fees and administration requirements while avoiding payment delays, contractual disputes, and insolvency [55]. In insurance, smart contracts reduce operational costs and processing time [56], improve data reliability, and offer privacy preservation [14]. They also reduce costs by streamlining processes and preventing corruption [57].

Smart contracts facilitate efficient resource management by optimizing business processes [7]. In energy, they enable automated trading, leading to more efficient building energy control and use [22]. They support improved time and cost efficiencies through real-time project process monitoring and recording [24]. Smart contracts optimize supply chains through automated procurement, reducing transaction costs, improving inventory management, and promptly identifying parties responsible for quality defects or wastage [58].

Smart contract implementation in waste management has several environmental and social implications spanning waste handling, treatment, monitoring, and proper disposal incentivization. Smart contracts optimize resource utilization through advanced analytics, minimizing waste and enhancing operational efficiency [16]. They enable carbon credit and emissions allowance possession and exchange monitoring, preventing fraudulent activities and ensuring single-transaction credits [59].

Combined with technologies like IoT, smart contracts support automated, decentralized water and air quality pollution monitoring [60]. They help monitor waste handlers and reduce waste [43]. Smart contracts develop incentive programs rewarding efficient waste management and reduction, potentially including “pay-as-you-throw” schemes significantly impacting municipal solid waste management [42].

Methods

Research Design

This study employed a systematic literature review methodology to comprehensively investigate the potential of blockchain-based smart contracts in addressing medical waste management challenges. The systematic approach enables a rigorous, transparent, and reproducible process for identifying, evaluating, and synthesizing relevant literature [61]. This method was particularly appropriate for this research as it allows for the integration of findings from diverse domains including blockchain technology, healthcare applications, waste management, and sustainability studies.

The review was guided by the following research questions:

1. What specific challenges in medical waste management can be effectively addressed through blockchain-based smart contracts?

2.How can smart contracts be designed and implemented to enhance tracking, compliance verification, and stakeholder coordination in medical waste management systems?

3.What technical, regulatory, and operational considerations must be addressed for successful implementation?

4.What are the potential sustainability impacts of implementing smart contract solutions in medical waste management?

The PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework was utilized to structure the review process, ensuring methodological rigor and comprehensive reporting [62]. This framework includes a four-phase flow diagram that guides the identification, screening, eligibility assessment, and inclusion of studies.

Search Strategy

A comprehensive search strategy was developed to identify relevant literature across multiple electronic databases. The primary databases searched included Scopus, Web of Science, PubMed. These databases were selected to ensure comprehensive coverage across computer science, healthcare, environmental sciences, and sustainability domains.

The search terms were organized into three main categories and combined using Boolean operators:

1. Blockchain technology: “blockchain” OR “distributed ledger” OR “smart contract”

2.Waste management: “waste management” OR “medical waste” OR “healthcare waste” OR “hospital waste” OR “biomedical waste”

3.Applications: “track” OR “monitor” OR “compliance” OR “regulation” OR “sustainable” OR “environment”

The initial search was conducted in January 2025, covering publications from 2015 (when Ethereum introduced advanced smart contract capabilities) to December 2024. The

search was limited to peer-reviewed journal articles, conference proceedings, and book chapters published in English.

In addition to database searches, a supplementary search strategy was employed to identify relevant gray literature, including technical reports, white papers, and policy documents from reputable organizations. Reference lists of included studies were also manually searched to identify additional relevant publications.

Inclusion and Exclusion Criteria

Studies were selected based on the following inclusion criteria:

1. Focus on blockchain technology and/or smart contracts

2. Direct application or potential application to waste management

3. Discussion of healthcare or medical contexts, or transferable insights

4. Addressing at least one aspect of tracking, compliance, coordination, or environmental monitoring

5. Published between 2015 and 2024

6. Written in English

Exclusion criteria were:

1. Studies focusing solely on cryptocurrency aspects without smart contract applications

2.Publications addressing general blockchain concepts without waste management relevance

3.Opinion pieces, editorials, or news articles without substantial technical or implementation content

4.Duplicate publications or multiple reporting of the same study

5.Studies with no full-text availability

Data Collection and Analysis

The review process began with the initial database search, which identified 423 potentially relevant publications. After removing duplicates, 312 unique records remained for screening. Two independent reviewers screened titles and ABSTRACTs against the inclusion and exclusion criteria, resulting in 143 publications for full-text assessment. Following full-text review, 87 studies were included in the final analysis.

A standardized data extraction form was developed to systematically extract relevant information from each included study. The form captured:

1. Publication details (authors, year, publication type)

2. Study characteristics (objectives, methodology, context)

3. Smart contract application details (blockchain platform, implementation approach)

4. Waste management context (general waste, medical waste, specific aspects addressed)

5. Key findings and outcomes

6. Implementation challenges and proposed solutions

7. Sustainability considerations (economic, environmental, social)

Thematic analysis was employed to synthesize the extracted data [63]. This involved an iterative process of open coding, where preliminary codes were assigned to text segments; axial coding, where relationships between codes were identified; and selective coding, where core categories were refined and integrated. This approach enabled the identification of key themes, patterns, and insights across the included studies.

A conceptual framework was developed through iterative refinement, integrating findings from the thematic analysis. The framework aimed to capture the essential components, relationships, and considerations for implementing smart contracts in medical waste management contexts.

Results

Overview of Literature

The systematic review identified 87 publications that met the inclusion criteria. The temporal distribution of publications showed a significant increase in research interest over time, with 5% of publications from 2015-2017, 27% from 2018-2020, and 68% from 2021-2024. This trend reflects the growing recognition of blockchain technology's potential in waste management applications, particularly in healthcare contexts.

In terms of publication types, the review included 53 journal articles (61%), 27 conference papers (31%), and 7 book chapters (8%). The geographical distribution of studies revealed that research was predominantly conducted in developed countries, with the largest contributions from the United States (18%), China (15%), United Kingdom (12%), Germany (9%), and Australia (8%). This geographical concentration highlights a potential research gap regarding blockchain applications in developing regions where medical waste management challenges are often more acute.

Regarding disciplinary focus, publications spanned multiple domains, including computer science and information technology (43%), healthcare and medicine (21%), environmental science and waste management (18%), and interdisciplinary studies combining multiple domains (18%). This diverse disciplinary representation demonstrates the inherently multifaceted nature of applying blockchain technology to medical waste management.

The analysis of blockchain platforms discussed in the literature revealed that Ethereum was the most referenced platform (62%), followed by Hyperledger Fabric (18%), and other platforms including Corda, EOS, and IOTA (collectively 20%). This distribution reflects Ethereum's prominence in supporting complex smart contract applications despite its sustainability concerns related to energy consumption.

Smart Contract Applications in Medical Waste Management

The analysis identified three primary categories of smart contract applications in medical waste management: automated tracking systems, compliance verification mechanisms, and environmental impact monitoring.

Automated Tracking Systems

Smart contracts, often integrated with IoT devices, enable comprehensive tracking of medical waste throughout its lifecycle. The review identified several key components of automated tracking systems:

***1. Origin Verification and Tagging*:** Smart contracts register the generation of medical waste, capturing essential data including waste type, quantity, generation time, and responsible personnel [47]. This initial registration creates a verifiable starting point for the waste tracking chain.

***2. Real-time Monitoring of Waste Movement*:** IoT sensors embedded in waste containers transmit data to blockchain networks, with smart contracts automatically updating waste status and location as it moves through the management chain [52]. This capability addresses a critical gap in traditional systems where waste tracking often relies on manual documentation and is subject to errors and fraud.

***3. Secure Chain of Custody Documentation*:** Smart contracts create immutable records of every stakeholder interaction with medical waste, documenting each transfer of responsibility [64]. This automated documentation can significantly reduce administrative burden while enhancing accountability.

***4. End-of-Life Verification*:** Smart contracts verify proper disposal or treatment of medical waste, creating closure in the tracking process [48]. This verification can include confirmation of incineration, sterilization, or other appropriate treatment methods depending on waste type.

The integration of these tracking components creates a transparent, tamper-resistant record of medical waste from generation to disposal, addressing critical challenges in current waste management systems related to fragmented tracking and limited transparency.

Compliance Verification Mechanisms

Smart contracts provide automated mechanisms for verifying and enforcing regulatory compliance in medical waste management:

***1. Automated Regulatory Adherence*:** Smart contracts encode regulatory requirements as executable conditions, automatically verifying compliance before allowing waste processing to proceed [25]. This encodes regulations as algorithmic rules, ensuring consistent application.

***2. Digital Permitting and Certification*:** Smart contracts verify and manage digital permits for waste handlers, transporters, and disposal facilities, allowing only authorized entities to participate in the waste management process [53]. This mechanism can signifi-

cantly reduce the risk of improper handling by unqualified parties.

***3. Audit Trail Generation*:** The immutable nature of blockchain enables comprehensive audit trails for regulatory inspection [28]. Smart contracts automatically generate compliance reports, documenting adherence to regulations throughout the waste management process.

***4. Non-compliance Detection and Reporting*:** Smart contracts can identify regulatory violations and automatically trigger notification systems to alert relevant stakeholders [27]. This early warning system enables prompt corrective action.

These compliance mechanisms address significant challenges in current medical waste management systems, where regulatory adherence often depends on periodic inspections and manual documentation, leaving substantial gaps in continuous compliance monitoring.

Environmental Impact Monitoring

Smart contracts facilitate the monitoring and management of environmental impacts associated with medical waste:

***1. Carbon Footprint Tracking*:** Smart contracts calculate and record the carbon emissions associated with waste transportation and treatment [59]. This capability enables more accurate environmental impact assessment and supports carbon reduction initiatives.

***2. Resource Optimization*:** Smart contracts analyze waste data to identify opportunities for waste reduction and resource recovery [44]. This can include identifying patterns in waste generation that could be addressed through improved procurement or operational practices.

***3. Incentive Mechanisms for Sustainable Practices*:** Smart contracts implement reward systems for environmentally responsible waste management practices [42]. These incentive structures can promote waste reduction, proper segregation, and sustainable disposal methods.

***4. Environmental Compliance Verification*:** Smart contracts verify adherence to environmental regulations, automatically documenting compliance with emissions standards, proper disposal methods, and contamination prevention measures [60].

These environmental monitoring applications demonstrate the potential for smart contracts to enhance the sustainability of medical waste management beyond basic tracking and compliance functions.

Implementation Framework

Based on the synthesis of literature findings, a comprehensive framework for implementing smart contracts in medical waste management was developed. Figure 1 illustrates this framework, which consists of four key components: system architecture, stakeholder roles, technical considerations, and governance mechanisms.

As shown in Figure 1, the system architecture component includes five primary layers: blockchain, smart contract, IoT integration, user interface, and data storage. Each layer serves a distinct function in the overall implementation.

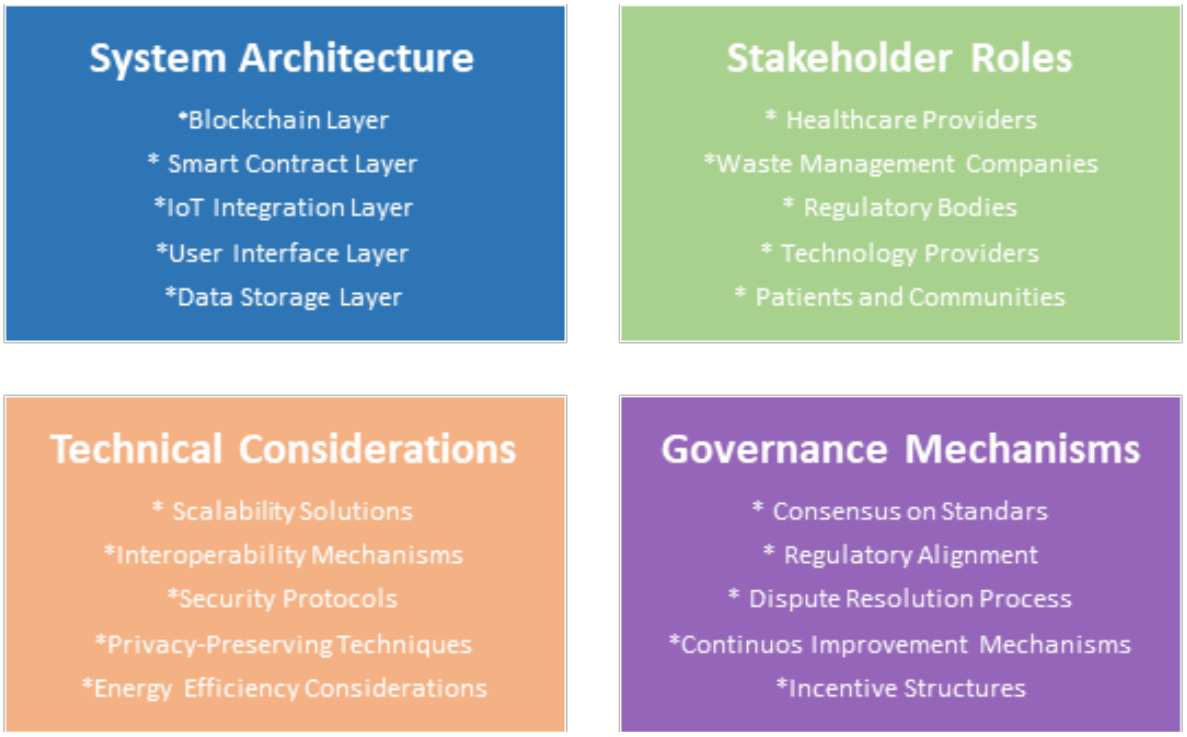


Figure 1: Smart Contract Framework for Medical Waste Management

System Architecture

The proposed system architecture includes:

***1. Blockchain Layer*:** Typically, a permissioned blockchain network (e.g., Hyperledger Fabric) that balances transparency with privacy considerations essential in healthcare contexts [28]. This layer provides the foundational distributed ledger infrastructure.

***2. Smart Contract Layer*:** Includes contracts for waste registration, tracking, compliance verification, stakeholder authentication, and incentive mechanisms [47]. These contracts encode the business logic for medical waste management processes.

***3. IoT Integration Layer*:** Connects physical waste management processes with the digital blockchain system through sensors, RFID tags, and monitoring devices [52]. This layer enables automated data collection without human intervention.

***4. User Interface Layer*:** Provides accessible interfaces for different stakeholders to interact with the system based on their specific roles and permissions [25]. This layer translates complex blockchain interactions into user-friendly applications.

***5. Data Storage Layer*:** Combines on-chain storage for critical transaction data with off-chain storage for larger datasets, balancing transparency with efficiency and privacy [53]. This hybrid approach addresses blockchain's inherent storage limitations.

Stakeholder Roles

The framework identifies key stakeholders and their roles within the system:

***1. Healthcare Providers*:** Generate medical waste, initiate the tracking process, and maintain responsibility for proper segregation and initial handling [47]. Smart contracts verify their compliance with waste segregation and packaging requirements.

***2. Waste Management Companies*:** Transport and process medical waste in accordance with regulations [52]. Smart contracts verify their credentials, track their activities, and automate payment based on service completion.

***3. Regulatory Bodies*:** Establish and enforce compliance standards [25]. Smart contracts provide real-time compliance monitoring and automated reporting capabilities.

***4. Technology Providers*:** Develop and maintain the blockchain infrastructure and associated applications [28]. Smart contracts manage system access and ensure technological component integration.

***5. Patients and Communities*:** Though not directly interacting with the system, they benefit from improved safety and environmental protection [27]. Smart contracts support transparency and accountability to these indirect stakeholders.

Technical Considerations

The framework addresses several critical technical considerations:

***1. Scalability Solutions*:** Implementing layer-two scaling or sharding techniques to handle the high transaction volume associated with comprehensive waste tracking [14]. These approaches address blockchain's inherent scalability limitations.

***2. Interoperability Mechanisms*:** Creating standards and protocols for data exchange between different blockchain systems and existing healthcare information systems [7]. This enables integration with existing waste management infrastructure.

***3. Security Protocols*:** Employing formal verification, code auditing, and security-focused development practices to prevent vulnerabilities in smart contract code [15]. These measures mitigate the risks associated with immutable code deployment.

***4. Privacy-Preserving Techniques*:** Implementing zero-knowledge proofs, secure multi-party computation, and data minimization practices to protect sensitive information while maintaining verification capabilities [11]. These techniques address the tension between blockchain transparency and healthcare privacy requirements.

***5. Energy Efficiency Considerations*:** Selecting consensus mechanisms with lower energy requirements, particularly for applications in healthcare settings concerned with environmental impact [15]. This addresses sustainability concerns associated with blockchain technology.

Governance Mechanisms

The framework includes governance mechanisms essential for effective implementation:

***1. Consensus on Standards*:** Establishing agreed-upon standards for data formats, communication protocols, and compliance requirements [24]. This enables consistent implementation across different healthcare contexts.

***2. Regulatory Alignment*:** Ensuring smart contract implementations comply with relevant healthcare regulations including HIPAA, GDPR, and waste-specific regulations [28]. This addresses the critical need for legal compliance.

***3. Dispute Resolution Processes*:** Defining procedures for addressing discrepancies, system failures, or disagreements between stakeholders [25]. These processes acknowledge that automated systems may still require human intervention in exceptional circumstances.

***4. Continuous Improvement Mechanisms*:** Creating processes for updating smart contracts and system components in response to regulatory changes, technological advancements, and stakeholder feedback [7]. These mechanisms balance blockchain's immutability with the need for system evolution.

***5. Incentive Structures*:** Designing economic and non-economic incentives to promote stakeholder participation and sustainable waste management practices [42]. These structures align individual stakeholder interests with broader system goals.

This comprehensive framework addresses the technical, operational, and governance dimensions of implementing smart contracts in medical waste management, providing a foundation for practical application development and implementation.

Discussion

Key Insights and Implications

The systematic review reveals several significant insights regarding the potential of blockchain-based smart contracts to transform medical waste management practices. These insights have important implications for healthcare providers, waste management companies, regulators, and technology developers.

First, smart contracts offer a paradigm shift in medical waste tracking by creating continuous, automated, and tamper-resistant documentation throughout the waste lifecycle. Unlike traditional systems that rely on periodic reporting and manual documentation, blockchain-enabled tracking provides real-time visibility and accountability [47]. This enhanced traceability can significantly reduce the risk of improper disposal, which is particularly critical for hazardous medical waste [48]. Healthcare facilities can benefit from reduced administrative burden associated with waste documentation while simultaneously improving their compliance posture and environmental performance.

Second, the automation of compliance verification through smart contracts represents a transformative approach to regulatory adherence. By encoding regulatory requirements directly into executable contract terms, smart contracts create a "compliance by design"

system where adherence is verified at each step of the waste management process [25]. This approach shifts compliance from a retrospective documentation exercise to a proactive, continuous verification process. Regulatory bodies can benefit from more reliable compliance data, enabling them to focus inspection resources on higher-risk facilities or activities rather than routine documentation checks.

Third, the integration of blockchain with IoT devices creates a powerful combination for medical waste management. IoT sensors provide automated data inputs about waste quantity, location, temperature, and other relevant parameters, while blockchain ensures the integrity and availability of this data [52]. This integration enables more sophisticated waste management approaches, including optimization of collection routes, real-time monitoring of waste treatment conditions, and verification of proper disposal methods. Waste management companies can leverage these capabilities to improve operational efficiency and service quality.

Fourth, smart contracts enable more transparent and efficient financial transactions among waste management stakeholders. Automated payment execution based on verified service completion can reduce payment delays, minimize disputes, and lower administrative costs [53]. This financial efficiency can benefit both healthcare facilities and waste management service providers, potentially leading to cost reductions in medical waste management services.

Finally, the framework developed in this study provides a comprehensive approach for implementing smart contracts in medical waste management contexts. By addressing system architecture, stakeholder roles, technical considerations, and governance mechanisms, the framework offers a practical guide for organizations seeking to develop blockchain-based waste management solutions [24]. This holistic approach acknowledges the socio-technical nature of medical waste management, recognizing that successful implementation requires attention to both technological and organizational dimensions. The framework presented in Figure 1 addresses both technical and governance aspects of smart contract implementation, providing a balanced approach that considers all relevant stakeholders.

These insights directly address our initial research questions by: (1) identifying specific medical waste management challenges that can be effectively addressed through blockchain-based smart contracts, particularly in tracking, compliance verification, and stakeholder coordination; (2) outlining how smart contracts can be designed and implemented through our proposed framework incorporating system architecture, stakeholder roles, technical considerations, and governance mechanisms; (3) detailing the technical, regulatory, and operational considerations that must be addressed for successful implementation, particularly regarding scalability, security, interoperability, and regulatory alignment; and (4) assessing potential sustainability impacts across economic, environmental, and social dimensions. This comprehensive analysis provides both theoretical insights and practical guidance for implementing blockchain-based smart contracts in medical waste management contexts.

Implementation Challenges

Despite the significant potential benefits, implementing blockchain-based smart contracts in medical waste management faces several substantial challenges that must be addressed for successful adoption.

Technical challenges include scalability limitations, as blockchain networks may struggle to process the high transaction volumes required for comprehensive waste tracking in large healthcare facilities [14]. Security vulnerabilities in smart contract code pose another significant concern, as flaws in contract logic could potentially lead to system exploitation with serious consequences for waste management operations [15]. Additionally, interoperability between blockchain systems and existing healthcare information systems remains problematic, with limited standardization creating barriers to seamless data exchange [7].

Regulatory challenges stem from the evolving nature of both blockchain regulations and waste management requirements. Smart contract implementations must navigate complex and sometimes contradictory regulatory frameworks, including healthcare-specific regulations (HIPAA, GDPR) and environmental regulations governing waste management [28]. The immutable nature of blockchain records may conflict with "right to be forgotten" provisions in privacy regulations, requiring careful system design to ensure compliance [11]. Furthermore, the legal status of smart contracts varies across jurisdictions, creating uncertainty regarding their enforceability in medical waste management contexts [25].

Organizational challenges include stakeholder adoption barriers stemming from limited understanding of blockchain technology among healthcare and waste management professionals [27]. The complexity of blockchain systems may create resistance to adoption, particularly in healthcare settings already experiencing technology fatigue. Integration with existing workflows and processes presents another significant challenge, as blockchain implementation requires substantial changes to established procedures and potentially threatens existing roles and responsibilities [24]. Additionally, the initial implementation costs, including technology development, system integration, and staff training, may deter organizations from adopting blockchain solutions despite potential long-term benefits [19].

Sustainability challenges arise from the energy consumption associated with certain blockchain consensus mechanisms; particularly proof-of-work systems used in public blockchains like Ethereum [15]. This energy consumption could potentially undermine the environmental benefits of improved waste management. Additionally, the economic sustainability of blockchain-based waste management systems remains uncertain, with questions about how implementation and operational costs will be distributed among stakeholders [7].

Sustainability Considerations

The implementation of blockchain-based smart contracts in medical waste management has significant implications for sustainability across economic, environmental, and social dimensions.

From an economic perspective, smart contracts offer potential cost reductions through process automation, reduced intermediation, and improved efficiency [54]. Automated

tracking and verification can minimize administrative overhead, while smart contract-executed payments can reduce transaction costs and payment delays [53]. However, these benefits must be weighed against the substantial implementation costs, including technology development, integration, and maintenance [19]. The economic viability of blockchain-based systems will depend on equitable distribution of both costs and benefits among stakeholders, requiring careful business model design and consideration of incentive structures [65]. The long-term economic sustainability of these systems will depend on demonstrating clear return on investment through reduced compliance risks, improved operational efficiency, and potential new value creation opportunities [66].

From an environmental perspective, blockchain-enabled smart contracts can contribute to more sustainable waste management practices through several mechanisms. Improved tracking and compliance verification can reduce instances of improper disposal, minimizing environmental contamination risks [47]. Enhanced data availability can support waste reduction initiatives by identifying patterns and opportunities for waste minimization at source [44]. Additionally, blockchain systems can enable more sophisticated recycling and resource recovery programs for appropriate medical waste streams [42].

However, these environmental benefits must be considered alongside the environmental impact of the blockchain technology itself. The energy consumption associated with certain consensus mechanisms, particularly proof-of-work, has raised significant sustainability concerns [67]. This has led to increasing interest in more energy-efficient alternatives like proof-of-stake and permissioned blockchain implementations that offer significantly lower energy requirements [68]. Medical waste management applications should prioritize these more sustainable blockchain architectures to ensure that technological environmental impacts do not undermine waste management environmental benefits [69].

From a social perspective, blockchain-based medical waste management systems can enhance transparency and accountability, providing greater assurance to communities regarding proper waste handling [70]. This transparency can build trust among stakeholders and address concerns about environmental justice related to waste disposal facility siting and operations [71]. Additionally, the automation of routine tasks through smart contracts can potentially improve working conditions for waste management personnel by reducing exposure to hazardous materials through better tracking and handling procedures [72].

However, the social implications of blockchain implementation extend beyond these benefits to include potential workforce impacts as automation changes job requirements and responsibilities [73]. The digital divide and varying levels of technological literacy among stakeholders may also create inequitable access to system benefits if not adequately addressed through inclusive design and implementation approaches [74]. Successful implementation will require attention to these social dimensions, including stakeholder engagement, education, and equitable benefit distribution [75].

Conclusion

This research has examined the potential of blockchain-based smart contracts to transform medical waste management practices through enhanced tracking, automated

compliance verification, and improved stakeholder coordination. The systematic analysis of literature spanning smart contract applications, healthcare implementations, waste management systems, and sustainability considerations has revealed both significant opportunities and persistent challenges in applying this emerging technology to the distinctive requirements of medical waste management.

The findings demonstrate that smart contracts, particularly when integrated with Internet of Things (IoT) devices and existing healthcare information systems, can address several critical challenges in medical waste management. By creating immutable audit trails that track medical waste from generation to disposal, smart contracts enable transparent verification of proper handling procedures, reducing the risk of improper disposal and its associated environmental and public health consequences [47, 48]. The automation of compliance verification processes through predefined contract terms can significantly reduce administrative burden while enhancing adherence to regulatory requirements [25, 28].

The conceptual framework developed in this research provides a comprehensive approach for implementing smart contracts in medical waste management, addressing system architecture, stakeholder roles, technical considerations, and governance mechanisms. This holistic framework acknowledges the socio-technical nature of medical waste management and provides a foundation for practical application development [24], 47]. However, implementation faces substantial challenges including scalability limitations [14], security vulnerabilities [15], regulatory uncertainties [11, 28], and adoption barriers [27, 74].

This research contributes to both scholarly understanding and practical application by identifying specific use cases where smart contracts offer substantial benefits, synthesizing insights from diverse disciplines, and providing a balanced assessment of both opportunities and challenges. Future research should focus on empirical validation through pilot implementations, quantitative economic analysis, and comparative studies of different blockchain platforms for specific medical waste management contexts. As blockchain technology matures, collaborative efforts across healthcare, technology, regulatory, and waste management sectors will be essential to harness its transformative potential for more sustainable waste management practices.

Declaration of Interest: The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgements: The authors would like to acknowledge the support provided by Izmir Katip Çelebi University and Izmir Bakırcay University during the research process. We also thank the anonymous reviewers for their valuable comments and suggestions that helped improve the quality of this paper.

Author Contributions: Arif Soyler: Conceptualization, methodology, data collection, formal analysis, writing - original draft preparation. Esra Dundar Aravacık: Supervision, validation, writing - review and editing. Serhat Burmaoglu: Supervision, methodology, data collection, validation, formal analysis, writing - review and editing.

REFERENCES

[1] M. Kassou, H. Badir, and N. Sbihi, "Blockchain-based medical and water waste management conception," *E3S Web Conf.*, vol. 234, p. 00070, 2021.

[2] A. Gehlot, R. Gupta, and S. R. Das, "Animal welfare monitoring," *Information*, vol. 13, no. 2, p. 67, 2022.

[3] A. Chidepatil et al., "From trash to cash: How blockchain and multi-sensor-driven artificial intelligence can transform circular economy of plastic waste?," *Admin. Sci.*, vol. 10, no. 1, p. 23, 2020.

[4] World Health Organization, "Health-care waste," WHO, Feb. 2018. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/health-care-waste

[5] P. Sharma, S. Singh, and J. H. Park, "Blockchain-based waste management in smart cities: A comprehensive survey," *Sensors*, vol. 23, no. 1, p. 56, 2023.

[6] S. N. Khan, F. Loukil, C. Ghedira-Guegan, E. Benkhelifa, and A. Bani-Hani, "Blockchain smart contracts: Applications, challenges, and future trends," *Peer-to-Peer Netw. Appl.*, vol. 14, no. 5, pp. 2901-2925, 2021.

[7] Y. Xu, Y. Huang, and Y. Bian, "A review of smart contracts applications in various industries: A procurement perspective," *Adv. Civil Eng.*, vol. 2021, pp. 1-17, 2021.

[8] A. Khatoon, "A blockchain-based smart contract system for healthcare management," *Electronics*, vol. 9, no. 1, p. 94, 2020.

[9] G. Ongena, K. Smit, J. Boksebeld, G. Adams, Y. Roelofs, and P. Ravesteyn, "Blockchain-based smart contracts in waste management: A silver bullet?," *Bled eConf.*, vol. 19, pp. 1-19, 2018.

[10] M. Alharby and A. van Moorsel, "Blockchain-based smart contracts: A systematic mapping study," *arXiv preprint arXiv:1710.06372*, 2017.

[11] S.-Y. Lin, L. Zhang, J. Li, L.-L. Ji, and Y. Sun, "A survey of application research based on blockchain smart contract," *Cluster Comput.*, vol. 25, no. 5, pp. 4013-4029, 2022.

[12] D. Macrinici, C. Cartofeanu, and S. Gao, "Smart contract applications within blockchain technology: A systematic mapping study," *Telematics Inform.*, vol. 35, no. 8, pp. 2337-2354, 2018.

[13] M. Di Angelo and G. Salzer, "A survey of tools for analyzing Ethereum smart contracts," in *Proc. IEEE Int. Conf. Decentralized Appl. Infrastruct.*, 2019, pp. 69-78.

[14] S. Rouhani and R. Deters, "Security, performance, and applications of smart contracts: A systematic survey," *IEEE Access*, vol. 7, pp. 50759-50779, 2019.

[15] G. Wu, W. Jiang, K. Li, Y. Li, and X. Zhang, "Smart contract security analysis: A systematic review," *J. Network Comput. Appl.*, vol. 226, p. 103882, 2024.

[16] P. Sharma, R. Jindal, and M. D. Borah, "A review of smart contract-based platforms, applications, and challenges," *Cluster Comput.*, vol. 26, no. 1, pp. 395-421, 2023.

[17] S. Wang et al., "Blockchain-enabled smart contracts: Architecture, applications, and future trends," *IEEE Trans. Syst., Man, Cybern. Syst.*, vol. 49, no. 11, pp. 2266-2277, 2019.

[18] V. Paliwal, S. Chandra, and S. Sharma, "Blockchain technology for sustainable supply chain management: A systematic literature review and a classification framework," *Sustainability*, vol. 12, no. 18, p. 7638, 2020.

[19] S. Aggarwal, N. Kumar, and S. Sharma, "Blockchain applications in smart community context: A comprehensive analysis," *J. Ambient Intell. Humanized Comput.*, vol. 10, no. 12, pp. 4707-4731, 2019.

[20] P. M. Hartmann and O. Hohn, "Crowdfunding and blockchain: What to expect?," *Eur. J. Innov. Manage.*, vol. 22, no. 3, pp. 518-535, 2019.

[21] W. Meng and Z. Qian, "Blockchain for business: An overview of functions, applications and open issues," *J. Phys.: Conf. Ser.*, vol. 1069, p. 012006, 2018.

[22] D. Kirli, T. Yigitcanlar, and U. Madanayake, "Smart contracts for energy: A systematic review of applications and challenges," *Renewable Sustainable Energy Rev.*, vol. 158, p. 112013, 2022.

[23] A. El and A. Bacha, "Review on the applications of blockchain-based smart contracts in the smart grid," *Energies*, vol. 14, no. 11, p. 3063, 2021.

[24] J. Li and M. Kassem, "A critical review of the application of distributed ledger technologies and smart contracts in the construction sector," *Autom. Construction*, vol. 132, p. 103955, 2021.

[25] A. Khatoon, "A blockchain-based smart contract system for healthcare management," *Electronics*, vol. 9, no. 1, p. 94, 2020.

[26] A. Azaria, A. Ekblaw, T. Vieira, and A. Lippman, "MedRec: Using blockchain for medical data access and permission management," in *Proc. 2nd Int. Conf. Open Big Data*, 2016, pp. 25-30.

[27] Z. O. Candereli, S. Burmaoglu, L. B. Kidak, and D. O. Gungor, "Applying blockchain technologies in healthcare: A scientometric analysis," in *Multidimensional Perspectives and Global Analysis of Universal Health Coverage*, IGI Global, 2020, pp. 69-92.

[28] G. G. Dagher, J. Mohler, M. Milojkovic, and P. B. Marella, "Ancile: Privacy-preserving framework for access control and interoperability of electronic health records using blockchain technology," *Sustainable Cities Society*, vol. 39, pp. 283-297, 2018.

[29] K. N. Griggs, O. Ossipova, C. P. Kohlios, A. N. Baccarini, E. A. Howson, and T. Hayajneh, "Healthcare blockchain system using smart contracts for secure automated remote patient monitoring," *J. Med. Syst.*, vol. 42, no. 7, pp. 1-7, 2018.

[30] M. A. Salahuddin, A. Al-Fuqaha, and I. Khan, "Agile softwarized infrastructure for IoT with secure and privacy preserving deployment in smart healthcare services," in *Proc. IEEE Int. Conf. Commun.*, 2018, pp. 1-6.

[31] T. Hewa, M. Ylianttila, and M. Liyanage, "Survey on blockchain based smart contracts: Applications, challenges and future directions," *J. Network Comput. Appl.*, vol. 177, p. 102857, 2021.

[32] T. Nugent, D. Upton, and M. Cimpoesu, "Improving data transparency in clinical trials using blockchain smart contracts," *F1000Research*, vol. 5, p. 2541, 2016.

[33] T.-T. Kuo, H.-E. Kim, and L. Ohno-Machado, "Blockchain distributed ledger technologies for biomedical and health care applications," *J. Am. Med. Inform. Assoc.*, vol. 24, no. 6, pp. 1211-1220, 2017.

[34] O. Akter, *Blockchain leveraged incentive providing waste management system*. Technical University of Munich, 2017.

[35] R. Borchard, R. Zeiss, and J. Recker, "Digitalization of waste management: Insights from German private and public waste management firms," *Waste Manage. Res.*, vol. 39, no. 9, pp. 1355-1365, 2021.

[36] K. Chary and K. Kumar, "GSM and GPS based solid waste management for smart cities," *Indian J. Public Health Res. Dev.*, vol. 9, no. 68, pp. 68-72, 2018.

[37] D. Dutta and S. Goel, "Applications of remote sensing and GIS in solid waste management: A review," in *Advances in Solid and Hazardous Waste Management*, Springer, 2017, pp. 133-151.

[38] M. Hannan et al., "An automated solid waste bin level detection system using Gabor wavelet filters and multi-layer perception," *Resour., Conserv. Recycl.*, vol. 72, pp. 33-42, 2013.

[39] M. Hannan, M. Al Mamun, A. Hussain, H. Basri, and R. Begum, "A review on technologies and their usage in solid waste monitoring and management systems: Issues and challenges," *Waste Manage.*, vol. 43, pp. 509-523, 2015.

[40] A. Kumar and S. Samadder, "A review on technological options of waste to energy for effective management of municipal solid waste," *Waste Manage.*, vol. 69, pp. 407-422, 2017.

[41] P. Bindra, A. Chidepatil, D. Kulkarni, and M. Kshirsagar, "Smart contracts and multi-sensor driven AI in circular economy of plastic waste," *Admin. Sci.*, vol. 10, no. 1, p. 23, 2019.

[42] A. Chidepatil et al., "From trash to cash: How blockchain and multi-sensor-driven artificial intelligence can transform circular economy of plastic waste?," *Admin. Sci.*, vol. 10, no. 1, p. 23, 2020.

[43] G. Ongena et al., "Blockchain-based smart contracts in waste management: A silver bullet?," *Bled eConf.*, vol. 19, pp. 1-19, 2018.

[44] M. Poongodi et al., "An effective electronic waste management solution based on blockchain smart contract in 5G communities," in *Proc. IEEE 3rd 5G World Forum*, 2020, pp. 1-6.

[45] A. Richardson and J. Xu, "Carbon trading with blockchain," in *Int. Conf. Financial Cryptography Data Security*, Springer, 2020, pp. 105-124.

[46] N. Aung and T. Tantidham, "Policy based smart home system using blockchain technology," in *Proc. 9th Int. Conf. Knowledge Smart Technol.*, 2017, pp. 135-139.

[47] M. Kassou, H. Badir, and N. Sbihi, "Blockchain-based medical and water waste management conception," *E3S Web Conf.*, vol. 234, p. 00070, 2021.

[48] A. Gehlot, R. Gupta, and S. R. Das, "Animal welfare monitoring," *Information*, vol. 13, no. 2, p. 67, 2022.

[49] M. Bugdol, M. Grebosz-Krawczyk, and D. Jelonek, "Systemic support and environmental awareness in a normalised environmental management system consistent with ISO 14001," *Manage. Environ. Quality Int. J.*, vol. 32, no. 5, pp. 949-969, 2021.

[50] A. Manzoor, A. Trivedi, and S. K. Singh, "Securing digital payment system using blockchain technology," in *Proc. 11th Int. Conf. Security Inform. Networks*, 2018, pp. 1-4.

[51] D. Pradeepkumar, B. J. Jansen, and U. Hacioglu, "A model-based approach for automated analysis of regulatory representations," in *Proc. IEEE 17th Int. Conf. Trust, Security Privacy Comput. Commun.*, 2018, pp. 125-131.

[52] M. Sharma et al., "Internet of Things (IoT) adoption barriers of smart cities' waste management: An Indian context," *J. Cleaner Prod.*, vol. 270, p. 122047, 2020.

[53] B. Zhao, P. Fan, and M. Zhang, "Y-DWMS: A digital watermark management system based on smart contracts," *Sensors*, vol. 19, no. 14, p. 3091, 2019.

[54] H. Cardeira, "Smart contracts and their applications in the construction industry," *Proc. New Perspectives Construction Law Conf.*, 2015.

[55] H. Y. Chong and T. Diamantopoulos, "The potential of smart contracts for addressing payment problems in the construction industry," *Construction Manage. Econ.*, vol. 38, no. 11, pp. 1013-1030, 2020.

[56] M. Alharby, A. Aldweesh, and A. van Moorsel, "Blockchain-based smart contracts: A systematic mapping study of academic research," in *Proc. Int. Conf. Cloud Comput., Big Data Blockchain*, 2018, pp. 1-6.

[57] D. Rustiana, D. Ramadhan, L. Wibowo, A. W. Nugroho, and G. Mahardika, "State of the art blockchain enabled smart contract applications in the university," *Blockchain Frontier Technol.*, vol. 2, no. 2, pp. 70-80, 2022.

[58] A. Ahmadisheykhsarmast and R. Sonmez, "Smart contracts in construction industry," in *Proc. 5th Int. Project Construction Manage. Conf.*, 2018, pp. 767-776.

[59] C. Fu, W. Wang, and Y. Zhang, "A blockchain powered environmentally sustainable emission trading framework for the fashion apparel manufacturing industry," in *Proc. IEEE Int. Conf. Software Quality, Reliability Security Companion*, 2018, pp. 393-397.

[60] S. R. Niya, S. S. Jha, T. Bocek, and B. Stiller, "Design and implementation of an automated and decentralized pollution monitoring system with blockchains, smart contracts, and LoRaWAN," in *NOMS 2018-2018 IEEE/IFIP Network Operations Manage. Symp.*, 2018, pp. 1-4.

[61] M. U. Hassan and A. Ahmad, "A systematic mapping study of smart contracts," *Comput. Standards Interfaces*, vol. 62, pp. 49-72, 2019.

[62] F. Alkhudayri, M. A. Hassan, K. Z. Ghafoor, and R. A. Khan, "Smart contracts applications in the university: A systematic literature review," *Blockchain Frontier Technol.*, vol. 2, no. 2, pp. 69-88, 2023.

[63] M. Cheng, L. Zhang, and Z. Qian, "Blockchain-smart contracts for sustainable project performance: A systematic review," *Eng., Construction Architectural Manage.*, vol. 30, no. 3, pp. 1091-1127, 2023.

[64] H. Hasan, E. AlHadhrami, A. AlDhaheri, K. Salah, and R. Jayaraman, "Smart contract-based approach for efficient shipment management," *Comput. Ind. Eng.*, vol. 136, pp. 149-159, 2019.

[65] P. Pagano et al., "Implementation of blockchain technology in insurance contracts against natural hazards: A methodological multi-disciplinary approach," *Environ. Climate Technol.*, vol. 23, no. 3, pp. 211-229, 2019.

[66] P. Hegedűs, "Towards analyzing the complexity landscape of Solidity based Ethereum smart contracts," *Technologies*, vol. 7, no. 1, p. 6, 2019.

[67] P. Rani, S. Kumar, and S. Sharma, "Sustainable blockchain networks: Applications, challenges, and future directions," *J. Environ. Manage.*, vol. 354, p. 120273, 2024.

[68] S. Garg, A. Singh, and N. Kumar, "A survey of blockchain applications in IoT and smart cities," *Multimedia Tools Appl.*, vol. 83, no. 13, pp. 35323-35357, 2024.

[69] A. Di Vaio, F. Boccia, and M. Zito, "Blockchain for a circular economy," *Sustainability*, vol. 15, no. 15, p. 11445, 2023.

[70] A. Bracciali, I. Chatzigiannakis, A. Vitaletti, and M. Zecchini, "Citizens vote to act: Smart contracts for the management of water resources in smart cities," in *Proc. First Int. Conf. Societal Autom.*, 2019, pp. 1-8.

[71] A. Upadhyay, A. Kumar, and A. Tripathi, "A framework for blockchain implementation for promoting sustainability and social responsibility in circular economy," *Environ. Technol. Innov.*, vol. 22, p. 101441, 2021.

[72] D. Zhang, "Application of blockchain technology in incentivizing efficient use of rural wastes: A case study on yitong system," *Energy Procedia*, vol. 158, pp. 6707-6714, 2019.

[73] S. Singh, R. Chhabra, and J. Arora, "A systematic review of waste management solutions using machine learning, Internet of Things and blockchain technologies: State-of-Art, methodologies, and challenges," *Arch. Comput. Methods Eng.*, vol. 31, pp. 1255-1276, 2024.

[74] T. Stroumpoulis, S. Triantafyllou, and K. Moustakas, "Exploring the role of blockchain technology in reducing food waste in the hospitality industry: A proposed framework," *J. Environ. Manage.*, vol. 300, p. 113747, 2021.

[75] N. Kumar and A. K. Singh, "Smart contract-IoT combination: Foundations, current state, and future directions," *Multimedia Tools Appl.*, vol. 83, no. 10, pp. 26609-26648, 2024.

CALLIGRAPHY DRAWING WITH A SIMPLE ROBOTIC ARM

Rıdvan Yayla*
Bilecik Şeyh Edebali University, Computer Engineering Department, Bilecik, Türkiye

Hüseyin Hakan Çetinkaya
Başkent University, Computer Education and Instructional Technology Department, Ankara, Türkiye

Semih Çakır
Zonguldak Bülent Ecevit University, Karadeniz Ereğli Vocational School, Zonguldak, Türkiye
**(Corresponding author) ridvan.yayla@bilecik.edu.tr*

ABSTRACT

Nowadays, autonomous systems are widely used across various technological domains, including the Internet of Things, electric vehicles, defense systems, and robotics. These systems also offer innovative solutions for artistic applications by integrating traditional techniques with modern automation. Calligraphy, as an art form that merges textual communication with aesthetic expression, holds historical and cultural significance. It has been extensively used in manuscripts, religious texts, and formal documents, contributing to the preservation of cultural heritage. Beyond its artistic value, calligraphy enhances creativity and precision, making it a valuable skill in both traditional and contemporary design. In this study, a Raspberry Pi-based drawing system has been assembled using servo motors and simple materials, such as ice cream sticks, to facilitate calligraphic writing. While similar implementations exist, an additional graphical user interface (GUI) has been developed to enhance usability. The interface allows users to input calligraphic text, select from five different font styles, visualize the text on-screen, and initiate the drawing process via a dedicated control button. This work demonstrates how open-source hardware and software can be adapted for artistic purposes, bridging the gap between technology and traditional art forms.

Keywords: Raspberry pi, Calligraphy, Autonomous systems, Artistic Automation.

Introduction

The fusion of technology and art has led to the emergence of novel approaches that redefine traditional artistic expressions. Among these, calligraphy remains a significant art form that has been cherished for centuries due to its aesthetic and cultural value. While traditionally practised by skilled artisans using manual tools, advancements in automation and digital technology have introduced new possibilities for replicating calligraphic techniques. By leveraging modern hardware and software, it is now possible to create systems

that assist in or fully automate the process of artistic writing, making it more accessible and adaptable to contemporary applications.

This study explores the application of open-source hardware in calligraphic writing by developing a Raspberry Pi-based drawing system. The system utilizes servo motors and simple materials, such as ice cream sticks, to mimic human-like calligraphic strokes. Unlike existing implementations, this project enhances user interaction by incorporating a graphical user interface (GUI) that allows users to input text, select from multiple font styles, and visualize the output before execution [1][2]. By combining affordability and ease of use, this approach not only preserves the essence of calligraphy but also demonstrates how automation can contribute to the evolution of artistic craftsmanship.

Literature Review

The robotic arm technology is widely used for many generative fields. Additionally, this technology is also used for creative artwork. In the literature, there are a few featured studies for calligraphy. Hsu et al, created a robot for calligraphic text based on Chinese characters [3]. Wu et al., proposed a robotic writing framework based on a robotic hand-eye coordination method to solve these limitations [4]. Hsu et. al., created a robot design to attempts to use a delta robot's structure and reliable coordinates to develop a self-learning Chinese calligraphy-writing system that requires precise control [5]. Chao et. al, proposed a robotic learning approach that introduces a inverse reinforcement learning algorithm with human pREFERENCES into a robotic writing system [6]. Bai et. al., built a prototype parallel calligraphy robot system, including an execution subsystem, a software-defined subsystem and an acquisition subsystem [7]. Gao et. al., proposed a new robotic learning framework to perform numeral writing tasks by investigating human-machine interactions with human pREFERENCES [8]. Wang et. al., proposed a novel neuron-based hypothesis generation model, called hypothesis generation net, to model human cognition, including how to make decisions and how to do actions [9]. They showed a virtual robot with its cognition system can learn how to write Chinese calligraphy to validate the feasibility of the proposed hypothesis generation model in a simulation environment, where an image-to-action translation via a cognitive framework is proposed to learn the pattern of Chinese characters. Bayrak et. al., studied a project on aimed to recognize and draw objects by using TensorFlow, OpenCV and BrachioGraph libraries together with a drawing machine created with hardware [10].

System Design

Hardware Components and Assembly

The system is designed using a Raspberry Pi 3 Model B, which serves as the central controller for the robotic arm responsible for executing calligraphic strokes. The Raspberry Pi 3 Model B features a 1.2 GHz 64-bit quad-core ARM Cortex-A53 processor and 1 GB of RAM, making it capable of handling real-time motor control and graphical processing tasks [11]. The system follows the fundamental design principles of Brachiograph, a minimalistic pen plotter arm, and adapts it for calligraphic writing applications [12]. The pieces of equipment used in the robotic arm construction are shown in Figure 1.

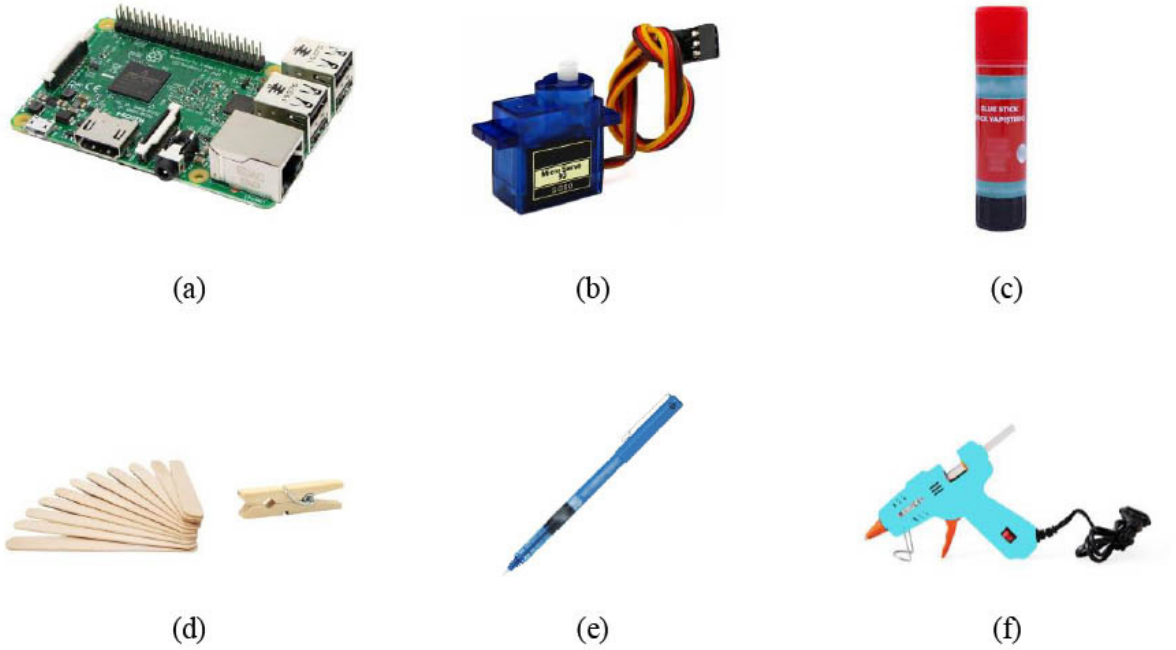


Figure 1: (a) Raspberry Pi 3 Model B (b) Micro Servo Motor SG90 (c) Glue Stick & Bamboo Clothes Peg (d) Popsicle Sticks (e) Pilot Pen (f) Silicone Gun

The Micro Servo Motor SG90 is a small and lightweight servo motor commonly used in robotics, RC vehicles, and embedded systems [13]. It operates on 4.8V to 6V and provides a torque of up to 1.8 kg/cm at 6V. The motor has a rotation angle of approximately 180 degrees and is controlled using PWM (Pulse Width Modulation) signals [13]. With its plastic gears, it offers a balance between affordability and performance, making it ideal for hobbyist projects and low-power applications.

Software and Control Mechanism

The software architecture is based on Python, utilizing RPi.GPIO and pigpio libraries for servo motor control. The drawing system use Biograph libraries for drawing, calibration and server motor rotation as the working principle [14]. The pen movements are achieved through Pulse Width Modulation (PWM) signals, which adjust the angles of the shoulder, elbow, and pen-lift servos. The system converts input text into coordinate-based instructions, which are then mapped to specific angular movements of the servos, ensuring that the pen accurately follows the desired calligraphic path. A Graphical User Interface (GUI) has been implemented using Tkinter as shown in Figure 2, allowing users to input custom text (number 1 text field) and choose from five different calligraphic font styles (number 2 dropdown menu) [15-17].

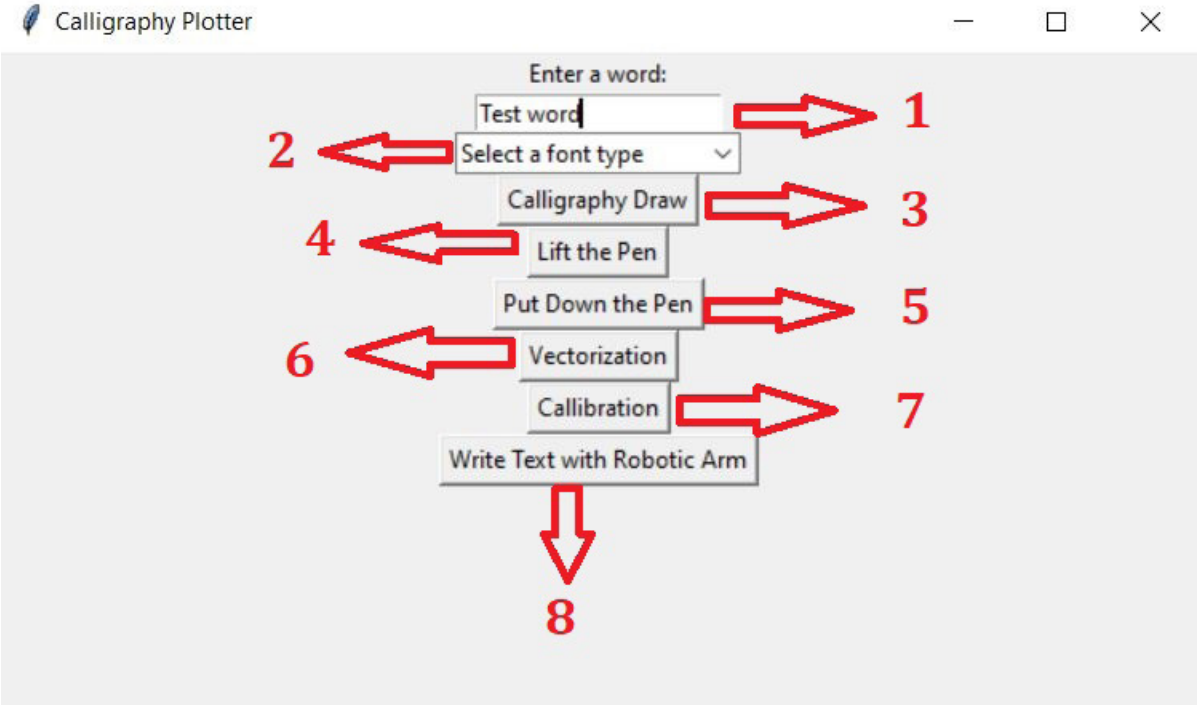


Figure 2: Tkinter GUI for calligraphic plotting

According to Figure 2, once the text and font are selected, the system generates a real-time preview on the screen by the number 3 button. The generated text is converted to vectorized points to recognize the text using the number 6 button. Upon pressing the number 8 button, the Raspberry Pi translates the font strokes into servo movement commands, initiating the drawing process.

To improve the fluidity of the strokes, the system implements motion smoothing algorithms, reducing abrupt changes in servo angles and ensuring a more natural replication of calligraphy. Before the plotting, a calibration should be made to prevent excessive pen shaking with the number 7 button. Additionally, the number 4 and 5 buttons provide to lift and put down to pen, respectively. Since servo motors have inherent positional limitations, the software includes calibration mechanisms, allowing users to fine-tune the range of movement and adjust the starting position for optimal performance.

The system consists of a three-servo motor structure that mimics a robotic arm. The first servo, referred to as the shoulder servo, controls the base rotation, allowing horizontal movements along the X-axis. The second servo, the elbow servo, governs the forearm movement, enabling vertical motion along the Y-axis. The third servo is responsible for lifting and lowering the pen, ensuring that it only makes contact with the drawing surface when necessary. These servos are mounted on a lightweight frame constructed using ice cream sticks and 3D-printed components, ensuring both stability and affordability. The pen holder is designed to securely grip the pen while allowing smooth movement during the drawing process.The robotic arm hardware is shown in Figure 3 and Figure 4.

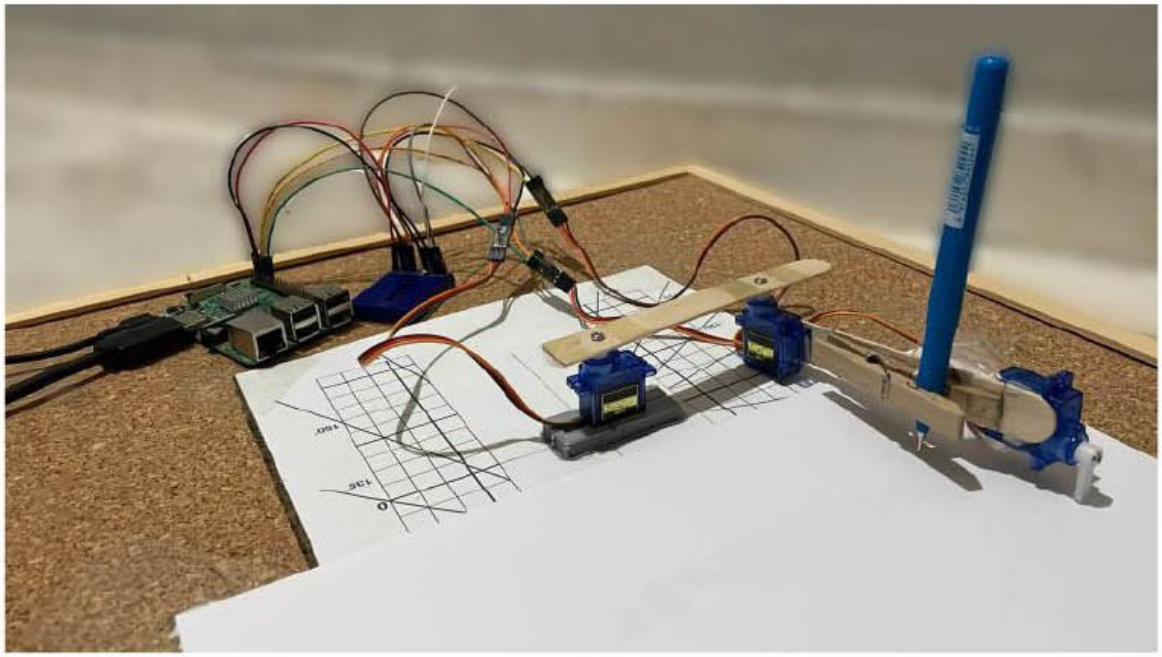


Figure 3: Robotic Arm Hardware (side view)

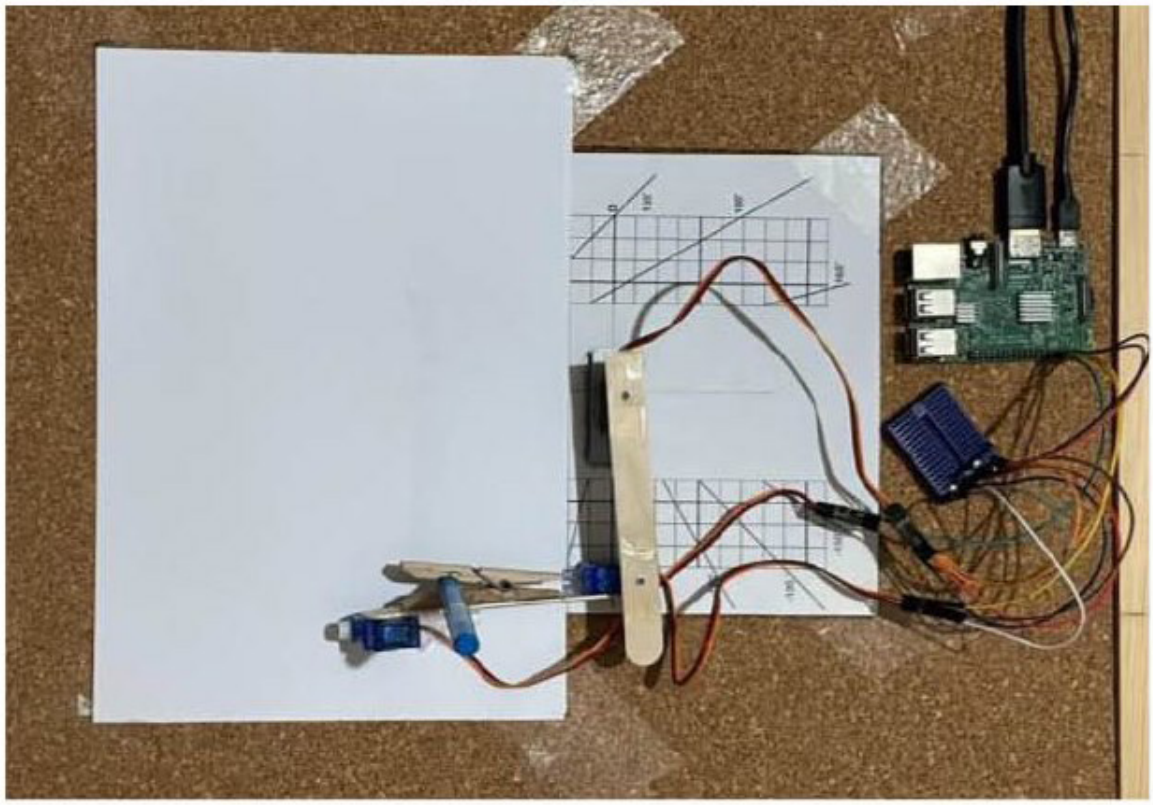


Figure 4: Robotic Arm Hardware (top view)

System Workflow

The system operates in a structured sequence, ensuring precise execution of calligraphic writing. The process begins with user interaction via the GUI, where the user enters the desired text and selects a calligraphic font. The system processes this input by converting the text into a set of vector-based coordinates, defining the movement path required to replicate each stroke.

Once the coordinate mapping is complete, the Raspberry Pi calculates the necessary angular adjustments for each servo. The shoulder servo determines the horizontal displacement, the elbow servo manages vertical positioning, and the pen-lift servo controls contact with the surface. The Raspberry Pi then generates PWM signals to direct the servo motors, ensuring smooth and synchronized movement. As the pen moves along the designated path, the calligraphic text is progressively drawn on the surface.

Throughout the drawing process, real-time adjustments are made to account for mechanical inaccuracies, reducing errors in stroke execution. The integration of feedback mechanisms helps refine motion precision, preventing jerky movements that could distort the calligraphic output.

Advantages and Limitations

This system offers several key advantages. First, it is built using affordable and accessible hardware, making it a cost-effective alternative to professional calligraphy robots. The ability to select from multiple font styles provides flexibility in artistic expression, while the open-source nature of the software enables further enhancements and customization. Additionally, the use of a three-servo robotic arm allows for more fluid and natural strokes, better replicating the nuances of traditional calligraphy compared to simpler two-motor plotters.

However, certain limitations must be considered. Due to the mechanical constraints of servo motors, the system has limited precision, particularly for highly intricate calligraphic details. The speed and smoothness of movements depend on servo response times, and minor calibration errors may lead to distorted letter shapes. Additionally, the pen-lift servo may introduce small delays in stroke transitions, affecting the overall fluidity of the writing process.

Experimental Results

The proposed Raspberry Pi-based calligraphy drawing system was tested using various text inputs and font styles to evaluate its performance in rendering artistic handwriting. Multiple trials were conducted, and a few sample outputs are shown in Figure 4, where the different words were drawn using the system.

Due to the low-cost construction of the mechanical structure, which was assembled using ice cream sticks and servo motors, the system exhibited noticeable mechanical vibrations and instability during operation. These vibrations led to visible distortions in the resulting calligraphic paths, particularly in curved or closely spaced strokes. As shown in the sample drawing, letter shapes are recognizable but suffer from jitter and overlapping

strokes, especially in complex characters.

The vector-based path planning successfully converted input text into draw-ready trajectories; however, the mechanical inaccuracy of the servo motors impacted the precision of execution. Despite these limitations, the system is still capable of producing rudimentary artistic outputs that reflect the intended calligraphic style to a reasonable degree.

These experimental results suggest that while the concept and software components of the system function correctly, improvements in hardware design—such as using higher-precision motors or a more rigid frame—would significantly enhance the fidelity and aesthetic quality of the generated calligraphy. Nevertheless, the project effectively demonstrates the feasibility of combining open-source electronics and software to bridge traditional artistic forms with modern automation.

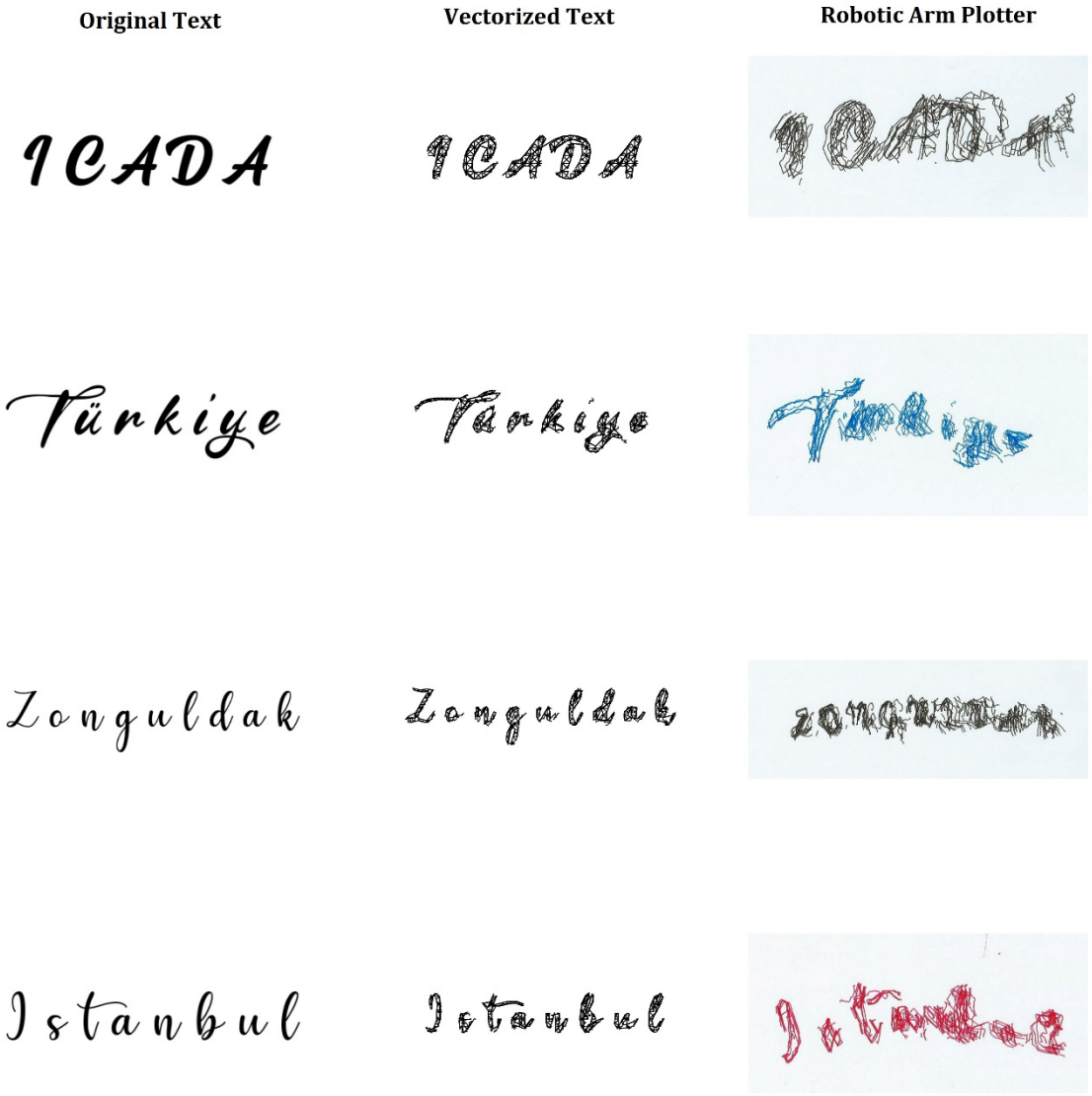


Figure 4: Experimental Results

Conclusion

This study demonstrates how low-cost robotics and autonomous systems can be applied to artistic domains, bridging the gap between traditional calligraphy and modern automation. By adapting the Brachiograph model and integrating a three-servo robotic arm, the system offers an innovative approach to automated calligraphy. Future improvements may focus on enhancing precision, expanding font customization options, and refining servo control algorithms to achieve higher accuracy and artistic quality.

Declaration of Interest: The authors declare that there is no conflict of interest.

Author Contributions: Rıdvan YAYLA prepared the hardware and experimental results, Hüseyin Hakan Çetinkaya contributed to the literature review and Semih Çakır revised to proofread and grammar the article.

REFERENCES

[1] P. C. Sai, K. Karthik, K. B. Prasad, C. V. Sai Pranav and D. K.V, "Real-Time Task Manager: A Python-Based Approach Using Psutil and Tkinter," 2024 8th International Conference on Computational System and Information Technology for Sustainable Solutions (CSITSS), Bengaluru, India, 2024, pp. 1-6, doi: 10.1109/CSITSS64042.2024.10816758.

[2] A. Narendran, A. Hothri, M. Yashaswini, S. Sundaram and K. Aiswariya Milan, "Personalized Travel Itinerary Recommender System using Tkinter and SQLite," 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), Kamand, India, 2024, pp. 1-6, doi: 10.1109/ICCCNT61001.2024.10725832.

[3] L. Gan et al., "Towards a Robotic Chinese Calligraphy Writing Framework," 2018 IEEE International Conference on Robotics and Biomimetics (ROBIO), Kuala Lumpur, Malaysia, 2018, pp. 493-498, doi: 10.1109/ROBIO.2018.8665143.

[4]R. Wu et al., "Internal Model Control Structure Inspired Robotic Calligraphy System," in IEEE Transactions on Industrial Informatics, vol. 20, no. 2, pp. 2600-2610, Feb. 2024, doi: 10.1109/TII.2023.3295415.

[5] M. -J. Hsu, P. -C. Yeh, Y. -H. Chien, C. -K. Lu, W. -Y. Wang and C. -C. J. Hsu, "A Chinese Calligraphy-Writing Robotic System Based on Image-to-Action Translations and a Hypothesis Generation Net," in IEEE Access, vol. 11, pp. 25801-25816, 2023, doi: 10.1109/ACCESS.2023.3252902.

[6] F. Chao et al., "Robotic Chinese Calligraphy with Human Preference," 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), Leicester, UK, 2019, pp. 360-366, doi: 10.1109/SmartWorld-UIC-ATC-SCALCOM-IOP-SCI.2019.00105.

[7]T. Bai, C. Guo, Y. Liu, Y. Lu, X. Dai and F. -Y. Wang, "Parallel Calligraphy Robot: Framework and System Implementation," in IEEE Journal of Radio Frequency Identification, vol. 7, pp. 163-167, 2023, doi: 10.1109/JRFID.2022.3212589.

[8]X. Gao, C. Zhou, F. Chao, L. Yang, C. -M. Lin and C. Shang, "A Robotic Writing Framework–Learning Human Aesthetic PREFERENCES via Human–Machine Interactions," in IEEE Access, vol. 7, pp. 144043-144053, 2019, doi: 10.1109/ACCESS.2019.2944912.

[9]W. -Y. Wang, M. -J. Hsu, L. -A. Yu, Y. -H. Chien and C. -C. Hsu, "Deep Learning-Based Hypothesis Generation Model and Its Application on Virtual Chinese Calligraphy-Writing Robot," in IEEE Access, vol. 8, pp. 87243-87251, 2020, doi: 10.1109/ACCESS.2020.2991767.

[10] T. BAYRAK Et Al. , "Raspberry Pi based Object Detection and Drawing," Global Conference On Engıneerıng Research (GLOBCER'21) , Balıkesir, Turkey, pp.194-201, 2021

[11]Raspberry Pi Foundation, "Raspberry Pi Documentation," [Online]. Available: https://www.raspberrypi.com/documentation/computers/raspberry-pi.html. [Accessed: 04-Feb-2025].

[12]K. McAleer, "BrachioGraph," Kev's Robots, 26 July 2024. [Online]. Available: https://www.kevsrobots.com/blog/brachiograph.html. [Accessed: 8-Jan-2025].

[13]"SG90 Datasheet," [Online]. Available: http://www.ee.ic.ac.uk/pcheung/teaching/DE1_EE/stores/sg90_datasheet.pdf. [Accessed: 13-Feb-2025].

[14]evildmp, "BrachioGraph – The cheapest, simplest pen-plotter," GitHub Repository, [Online]. Available: https://github.com/evildmp/brachiograph. [Accessed: 9-Jan-2025].

[15]Alan D. Moore, Python GUI Programming with Tkinter: Design and build functional and user-friendly GUI applications , Packt Publishing, 2021.

[16]S. B. Kasetty, I. N. Margret, K. S. Madhuri, R. K. Ojha, N. Sudha and K. Rajakumar, "A Hybrid Machine Learning and Deep Learning Approach for Diabetes Prediction With a Tkinter-Based User Interface," 2024 International Conference on Intelligent Computing and Sustainable Innovations in Technology (IC-SIT), Bhubaneswar, India, 2024, pp. 1-5, doi: 10.1109/IC-SIT63503.2024.10862656.

[17]S. Vasavi, P. D. L. Nikhita Sri and P. V. Sai Krishna, "GUI-Enabled Boundary Regularization System for Urban Buildings Using the Tkinter," 2024 2nd International Conference on Device Intelligence, Computing and Communication Technologies (DICCT), Dehradun, India, 2024, pp. 424-429, doi: 10.1109/DICCT61038.2024.10532910.

CHAOTIC SIGNAL-INDUCED FIRING CHARACTERISTIC OF HODGKIN-HUXLEY NEURON IN NEURAL NETWORK TOPOLOGY BASED ON NEWMAN–WATTS APPROACH

Muhammed Elhatip
Zonguldak Bülent Ecevit University, Biomedical Engineering, Zonguldak, Türkiye

Tuğba Palabaş*
Zonguldak Bülent Ecevit University, Biomedical Engineering, Zonguldak, Türkiye

Ziynet Pamuk
Zonguldak Bülent Ecevit University, Biomedical Engineering, Zonguldak, Türkiye
*(Corresponding author) tugba.palabas@gmail.com

ABSTRACT

In this study, the firing characteristics of neurons under constant excitation are examined within a biologically realistic hybrid network composed of Hodgkin-Huxley neurons. Each neuron in the network is subjected to a chaotic input signal, and the network topology is shaped according to the Newman-Watts algorithm. This approach allows the construction of regular or irregular Small-World (SW) networks by introducing new connections between non-adjacent neurons with a certain probability p. For computational analyses, we first observe how low and high levels of chaotic signal modulate neuronal activity patterns. Then, to provide a more general perspective about the effect of chaos on firing characteristics of the neurons, the variation of firing rate as a function of chaotic signal level is investigated. It is shown that the firing rate of the hybrid network using the Newman-Watts method reaches a maximum at the optimal level of chaotic signal. It is also stated that this behavior is independent of the constant excitation level and the network size.

Keywords: Firing characteristics, hybrid network, Newman-Watts.

Introduction

The human nervous system is a very complex system consisting of billions of neurons. This nervous system receives, codes, processes and transmits environmental information to the relevant organs to perform vital functions. All these processes are managed by sending electrochemical signals known as action potentials between neurons via synapses. In electrophysiological studies, it is observed that these neuronal activities recorded exhibit stochastic fluctuations at a certain level when there is or is not any stimulation. Such an activity has put forward the idea that biological neurons can operate in noisy environments. The studies in the literature, many noise sources such as ion channel noise, synaptic

noise and synaptic bombardment noise have been identified [1-4]. The role of noise in regulating neuronal activities has been investigated in many studies. However, it has been accepted that chaos can also organize the irregular behavior of individual neurons or neuronal populations [5-6].

Some systems such as Lorenz, Chen and Chua [6-10], which are thought to mimic the chaotic behavior of the neuronal environment, are frequently used in computational studies. In addition that it is also very important to model the network dynamics of the brain in order to reveal the collective behavior of neuronal systems. In this context, Scale-Free (SF) and Small-World (SW) network topologies, which are thought to support realistic neuronal information processing in terms of the anatomical and functional organization of the brain, are widely used [11-13]. And Watts-Strogatz and Newman-Watts approaches, in which new connections are established between two neurons at random, are suggested to obtain a random network structure containing the SW topology. The Newman-Watts method includes an algorithm structure in which previous connections are preserved while new connections are established [14, 15].

In this study, the Newman-Watts approach is used to investigate the effects of chaotic activity on the collective behavior of Hodgkin-Huxley neuron populations. The results show that in a biologically realistic hybrid network, the firing rate values reach a maximum at an optimal level of the chaotic signal and then they reach to minimum. It can be said that this behavior is partially independent of any external stimulation or network size.

Model and Method

The Hodgkin Huxley model equation used to calculate neuron dynamics in this study is as follows:

$$C_m \frac{dV_i}{dt} = -I_{Na} - I_K - I_L - (I_{ext} + I_{chaos} + I_c + I_e) \quad (1)$$

C_m=1/cm² is the membrane capacitance per unit area, *Vi* is the membrane potential of the i. neuron. *Iext=6. /cm²* is the constant current, Ic and Ie are synaptic currents including excitatory chemical coupling and electrical gap junction. Ionic currents I_{Na}, I_K, ve I_L are defined as follows for sodium, potassium, and other ions, respectively.

$$I_{Na} = g_{Na}^{max} m^3 h(V_i - E_{Na}) \quad (2)$$

$$I_K = g_K^{max} n^4 (V_i - E_K) \quad (3)$$

$$I_L = g_L (V_i - E_L) \quad (4)$$

Here, g_{Na}^{max}=120/cm² represents the maximum channel conductance of sodium and g_K^{max}=36/cm² represents the potassium ions, and g_L=0.3/cm² represents the total channel conductance of leakage ions. E_{Na} = 115mV, E_K = -12mV ve E_L = 10.6mV variables are the equilibrium potentials of sodium, potassium and leakage ions. Initial values of the gate variables m, n, h are determined randomly and the time-dependent behavior of the gate variables is described by the Langevin equation as follows:

$$\frac{dx}{dt} = \alpha_x(V)(1-x) - \beta_x(V)(x); x=m, n, h \quad (5)$$

The rate functions of the gate variables ve which vary with the membrane potential V, are as follows:

$$\alpha_m = 0.1 \frac{(25 - V)}{exp[(25 - V)/10] - 1} \quad (6)$$

$$\beta_m = 4\, exp[-V/10] \quad (7)$$

$$\alpha_n = 0.01 \frac{(10 - V)}{exp[(10 - V)/10] - 1} \quad (8)$$

$$\beta_n = 0.125\, exp[-V/80] \quad (9)$$

$$\alpha_h = 0.07\, exp[-V/20] \quad (10)$$

$$\beta_h = \frac{1}{exp[(30 - V)/10] + 1} \quad (11)$$

I_{chaos}= εx representation is integrated into the model (in Eq.1) by adjusting the intensity with the ε modulation parameter and this chaotic current signal x is calculated as follows:

$$\frac{d_x}{d_t} = \sigma(y - x) \quad (12)$$

$$\frac{d_y}{d_t} == x(p - z) - y \quad (13)$$

$$\frac{d_z}{d_t} == z(x - \gamma) \quad (14)$$

Here, σ, p and γ are constant values determined as 10, 28, 8/3 and the x and z phase portraits, which are two-dimensional cuts of 3-dimensional dynamical model that exhibits chaotic behaviors, are as shown in Fig. 1(a) and Fig. 1(b).

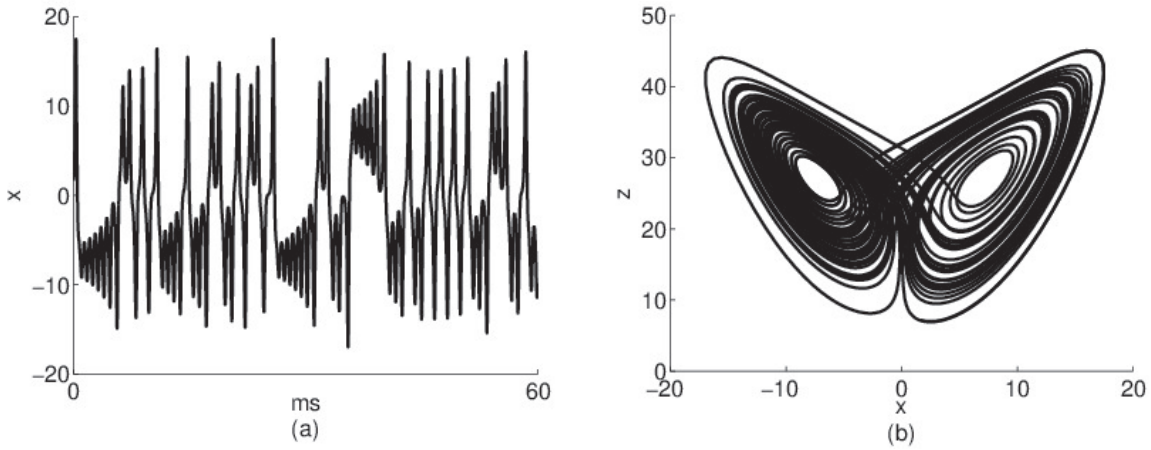


Figure 1. x and z phase plane diagram of a three-dimensional Lorenz system. (a) Time series of x variable of chaotic system. (b) Phase diagram on a two-dimensional x - z plane of chaotic attractor.

The x signal produced by the Lorenz system exhibits chaotic oscillation and the trajectory continues with a random number of turns left and right, as seen in panel b. A negative x in panel a corresponds to the left loop in panel b, and a positive x corresponds to the right loop.

In Equation 1, the synaptic current I_c, which is defined for the chemical excitatory synapse structure, entering the postsynaptic neuron is calculated as in Equation 15, depending on the membrane potential V_i of the i_{th} neuron, the resting potential E_{rev}=70mV of the presynaptic neuron and the synaptic conductance parameter g_c. For the electrical connection, it is calculated as in Equation 16, depending on the membrane voltages of the neurons and the electrical conductance parameter g_e.

$$I_i^{syn} = \sum_{j\in N(i)} g_c sj(E_{rev} - V_i) \quad (15)$$

$$I_i^{syn} = \sum_{i\in N(i)} g_e(V_j - V_i) \quad (16)$$

Here, the conductance change in a postsynaptic neuron is determined with the fraction of open receptor channels s_j for neuron j for the case of the chemical synapse.

In the hybrid network consisting of 200 neurons connected by excitatory chemical synapse and electrical gap junction, neurons have two local connections with each other. Initially, a network with a ring topology with equal local connections is defined. The topology of the neural network is transformed into a random network structure using the Newman-Watts approach. In this method, new connections are created randomly between neurons in the regular network. In order to examine the firing patterns of neurons in the network, membrane potentials and chaotic signals applied to the neuronal network are recorded for 1 s. Then, the average firing rate of each neuron during the T period is calculated and the collective firing rate value for N=200 neurons is obtained as follows:

$$\mathrm{r} = \frac{1}{N}\sum_{i=1}^{N}(fr_{mean}, \mathrm{i}) \quad (17)$$

Results

We begin our analysis by obtaining the firing activities of the neurons in the network at low and high levels of the chaotic signal. As seen in Fig. 2(a), when ε=0.078, the I_{kaos} current oscillates in the range of -1, 1 with respect to time. The neuron under the influence of constant current and chaotic current exhibits random firing activity. In Fig. 2(b), when ε=5 the chaotic current level has increased by 100 times. This level is quite high for real applications. Here, under the influence of the chaotic signal, the neuronal activity frequency increases and almost shows a chaos-like behavior. When we look at the collective firing rate (r) graph of the hybrid network in Fig. 2(c), it is understood that the spiking activity of all neurons in the network increases as ε increases and that r becomes zero after a certain level of ε, i.e., the neurons do not fire spikes.

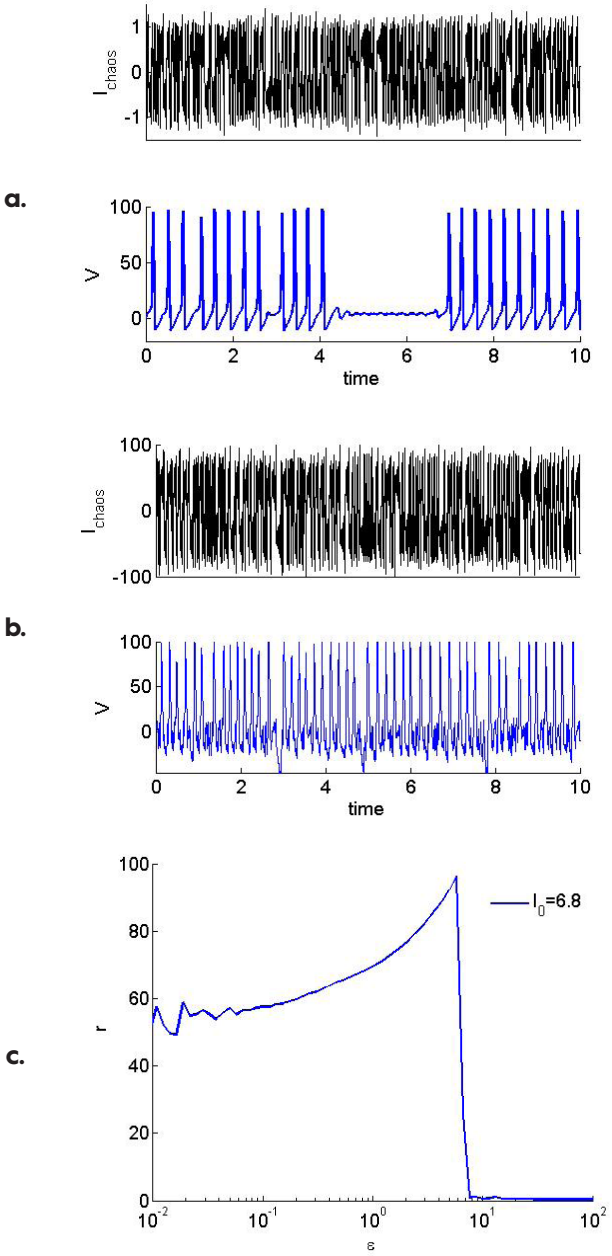


Figure 2: Firing characteristics of neurons depending on the chaotic current level (a) the chaotic current (black) and the firing activity of the 1st neuron when ε=0.078, (b) the chaotic current (black) and the firing activity of the 1st neuron when ε=5, (c) the dependence of collective firing rate r on chaotic current intensity ε for hybrid network.

Fig. 3 is obtained to show how this firing characteristic of the hybrid network changes according to the applied constant current value. Here, as a natural result of the increase in the current level, high r values are obtained even at smaller values of ε. Since a very large amplitude chaotic current is obtained at the maximum ε level, the ε at which r is maximum for all constant current values is almost the same. In addition, larger r is obtained for larger I_0.

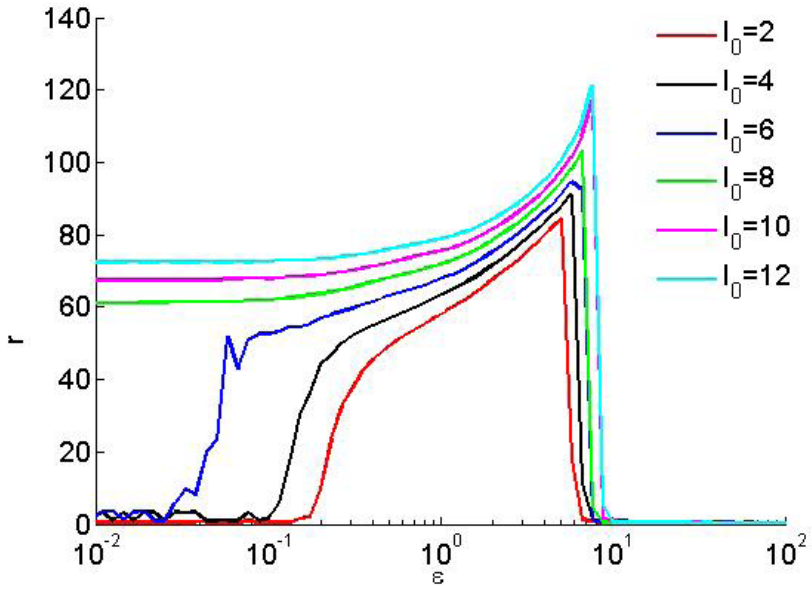


Figure 3: Collective firing rate (r) of neuronal network for 6 different constant external stimulus levels.

The analyses so far have been conducted by considering 200 neurons connected with a small-world network topology. Fig. 4 shows how the network size affects the collective firing rate value. As can be understood from the figure, the number of neurons in the network does not have a significant effect on the collective firing rate. It should be noted that when N = 10, a stable graph could not be obtained due to the small number of neurons at low epsilon levels and the small number of simulation repetitions. Considering that neuronal activity exhibits random firing behavior according to the chaotic environment, this can be overcome with more simulations.

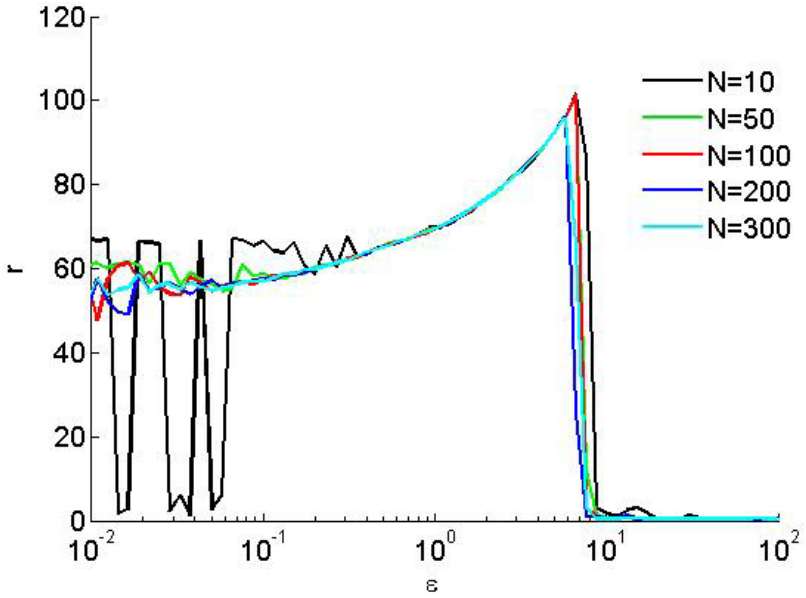


Figure 4: Collective firing activity of hybrid network for different network size.

As can be understood from the obtained results, at a significant level of chaotic activity, the firing rate of individual neurons or networks varies independently of the network size according to the constant current and chaotic signal level. In computational studies, these levels of chaos can be taken into account to investigate the role of chaotic signals in signal transmission.

Declaration of Interest: The authors declare that there is no conflict of interest.

Author Contributions: Muhammed Elhatip: Conceptualization, Methodology, Software, Investigation. Tugba Palabas: Conceptualization, Methodology, Software, Formal analysis, Investigation, Supervision, Validation, Visualization, Writing - review & editing – original draft. Ziynet Pamuk: Supervision, Validation, Visualization, Writing – review & editing

REFERENCES

[1] Guo, D., Perc, M., Liu, T., & Yao, D. (2018). Functional importance of noise in neuronal information processing. Europhysics Letters, 124(5), 50001.

[2]Hille, B. (1992). Ionic channels of excitable membranes. Published by Sinauer Assoc. Inc. Sunderland, MA.

[3]Koch, C. (2004). Biophysics of computation: information processing in single neurons. Oxford University Press.

[4]Hô, N., & Destexhe, A. (2000). Synaptic background activity enhances the responsiveness of neocortical pyramidal neurons. Journal of neurophysiology, 84(3), 1488-1496.

[5]Rabinovich, M. I., & Abarbanel, H. D. I. (1998). The role of chaos in neural systems. Neuroscience, 87(1), 5-14.

[6]Destexhe, A., & Rudolph-Lilith, M. (2012). Neuronal noise (Vol. 8). Springer Science & Business Media.

[7]Baysal, V., & Yılmaz, E. (2021). Chaotic signal induced delay decay in Hodgkin-Huxley Neuron. Applied Mathematics and Computation, 411, 126540.

[8]G. Chen and T. Ueta, "Yet another chaotic attractor", Int. J. Bifurc. Chaos, vol. 9, no. 7, pp. 1465-1466, 1999.

[9]Liang, X., & Qi, G. (2017). Mechanical analysis of Chen chaotic system. Chaos, Solitons & Fractals, 98, 173-177.

[10]Madan, R. N. (Ed.). (1993). Chua's circuit: a paradigm for chaos (Vol. 1). World Scientific.

[11]Albert R, Jeong H and Barabási A L (1999) Internet: Diameter of the World-Wide Web. Nature, 401 (6749): 130-131.

[12]Strogatz S H (2001) Exploring Complex Networks. Nature, 410 (6825): 268-276.

[13]Kaiser M (2011) A Tutorial in Connectome Analysis: Topological and Spatial Features of Brain networks. Neuroimage, 57 (3): 892-907.

[14]Yamakou, M. E., & Inack, E. M. (2023). Coherence resonance and stochastic synchronization in a small-world neural network: an interplay in the presence of spike-timing-dependent plasticity. Nonlinear Dynamics, 111(8), 7789-7805.

[15]Li, T., Wu, Y., Yang, L., Zhan, X., & Jia, Y. (2022). Spike-timing-dependent plasticity enhances chaotic resonance in small-world network. Physica A: Statistical Mechanics and its Applications, 606, 128069.

COLLISION DETECTION FOR THE SAFETY OF AUTONOMOUS SYSTEMS USING MACHINE LEARNING ALGORITHMS AND AUTOML LIBRARIES

Elif Cesur*
Endüstri Mühendisliği, Mühendislik ve Doğa Bilimleri, İstanbul Medeniyet Üniversitesi, Türkiye

Handegül İpek
Endüstri Mühendisliği, Mühendislik ve Doğa Bilimleri, İstanbul Medeniyet Üniversitesi, Türkiye

Hilal Gülmez
Endüstri Mühendisliği, Mühendislik ve Doğa Bilimleri, İstanbul Medeniyet Üniversitesi, Türkiye

Büşra Çetiner
Endüstri Mühendisliği, Mühendislik ve Doğa Bilimleri, İstanbul Medeniyet Üniversitesi, Türkiye
**(Corresponding author) elif.karakaya@medeniyet.edu.tr*

ABSTRACT

The success of autonomous systems depends on the accuracy and reliability of their abilities to perceive the environment, plan movement, and make decisions. Collision scenarios that may arise when encountering other vehicles, pedestrians, animals, and physical obstacles in traffic represent a significant challenge. The ability to prevent traffic accidents and predict potential collisions in advance is a key factor in increasing the reliability of autonomous systems from the perspective of human users.

Artificial intelligence algorithms such as machine learning and deep learning have significantly advanced the field of collision detection. These technologies can achieve highly successful results even with complex datasets, enabling them to predict collision risks in advance. Analyzing sensor, LiDAR, radar, and camera data using machine learning algorithms has proven to be highly effective in evaluating possible scenarios.

In this study, machine learning and AutoML algorithms were used to detect collisions for autonomous vehicles. Using a dataset collected from vehicle sensors, predictions were made with various machine learning algorithms and an AutoML library. Furthermore, hyperparameter optimization was performed in the AutoML applications to improve the results. While machine learning achieved an accuracy rate of 94%, AutoML increased this rate to as high as 98%.

Keywords: Autonomous vehicles, collision detection, machine learning, AutoML

Introduction

Autonomous vehicles have emerged as a revolutionary technology, transforming transportation by integrating artificial intelligence, sensors, and machine learning algorithms [1]. The primary goal of these vehicles is to enhance efficiency, reduce human error, and ultimately increase road safety [2]. However, despite technological advancements, autonomous systems still face challenges, particularly in predicting and preventing collisions [3].

One of the main concerns in autonomous driving is ensuring real-time decision-making capability to avoid crashes. Traffic accidents primarily result from human errors, including distraction, fatigue, and misjudgment [4]. By utilizing advanced computational methods, autonomous systems can reduce reliance on human perception and execute precise maneuvers based on sensor data. Nevertheless, the accuracy of these predictions depends on robust models that can analyze environmental inputs and forecast potential risks [5].

Machine learning plays a crucial role in improving collision detection. Traditional rule-based models struggle to adapt to dynamic driving conditions, whereas machine learning techniques enable vehicles to learn from past scenarios and continuously optimize their decision-making process [6]. Supervised learning algorithms, such as Artificial Neural Networks (ANN), Decision Trees (DT), and Support Vector Machines (SVM), have proven effective in classifying risky driving scenarios [7]. More recently, Automated Machine Learning (AutoML) frameworks, such as PyCaret, have gained attention for their ability to streamline model selection and hyperparameter tuning, improving predictive accuracy with minimal human intervention [8].

This study aims to enhance autonomous driving safety by comparing various machine learning methods in predicting collision risks. By applying supervised learning models and AutoML techniques to autonomous vehicle sensor data, we analyze how different classification algorithms perform in identifying risky interactions. The research also explores the impact of feature selection, data preprocessing, and hyperparameter optimization in improving model accuracy [9].

Ultimately, the findings of this study aim to contribute to autonomous vehicle technology by providing a data-driven approach to collision detection. As autonomous systems become more prevalent in urban mobility, optimizing prediction accuracy will be essential for their widespread adoption and regulatory approval. Through comprehensive analysis, this research seeks to address existing gaps and pave the way for future advancements in autonomous driving safety [10] .

Materials And Method

Machine learning refers to the entirety of learning-based methods that enable classification and prediction through the interpretation of large-scale data using statistical techniques [11]. Learning occurs when the system memorizes patterns within an existing dataset and evaluates future data accordingly. Machine learning is generally categorized into three types: supervised learning, unsupervised learning, and reinforcement learning.

Unsupervised learning involves methods where only input data is present, and there are no labeled outputs. The primary objective in this type of learning is not classification but rather clustering, examining relationships between attributes, or dimensionality reduction [12] .

Reinforcement learning, on the other hand, is a learning style in which the machine explores different actions to reach a defined goal and learns the optimal path through trial and error [13].

Supervised learning, also known as guided learning, involves both input and output data, and during the training process, the relationship between them is examined. The presence of input and output makes the data labeled. Supervised learning is among the most commonly used methods in machine learning [14]. It is typically studied under two main categories: regression and classification. Classification techniques include algorithms such as decision trees, artificial neural networks, discriminant analysis, support vector machines, and k-nearest neighbors.

Artificial neural networks are artificial intelligence models inspired by the working system of neurons in the human brain. This model consists of three main layers: the input layer, hidden (intermediate) layers, and the output layer. These layers come together to form a network similar to neural connections. The data entering through the input layer is processed in the hidden layers using weight values and is then transmitted to the output layer. While the weights are initially assigned randomly, they are optimized during the training process; however, the exact final values of these weights are not known to users [15]. The number of hidden layers is determined by the user according to the desired complexity of the model.

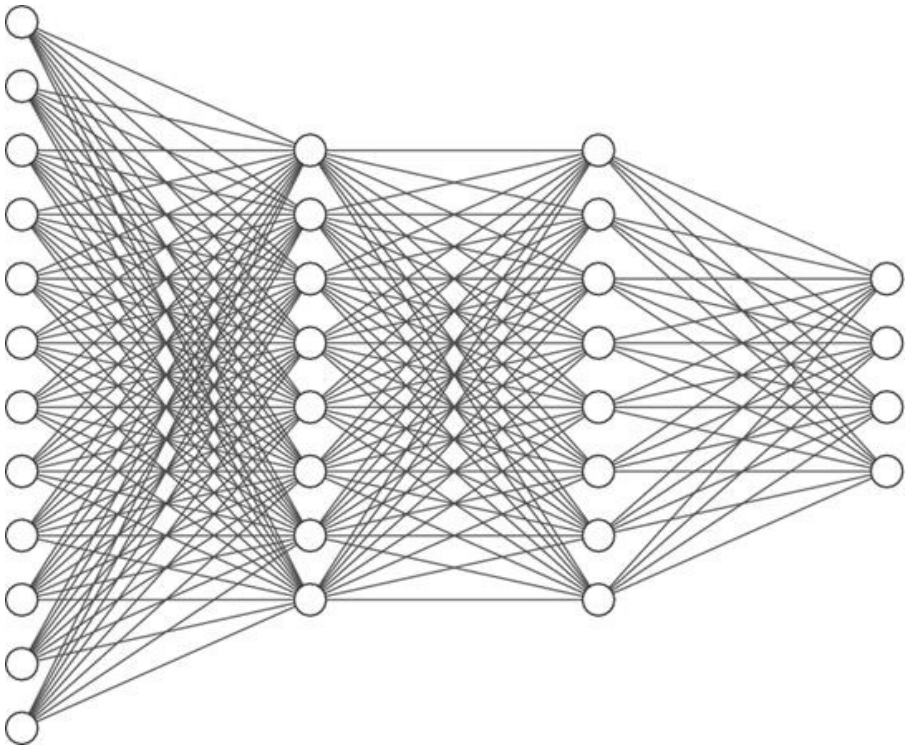


Figure 1: Artificial Neural Networks

Decision trees are a classification method applied using an inductive approach. A decision tree consists of a root node that contains all the data, branches associated with this root node, and leaves connected to these branches [16]. Each node from the root to the leaves includes a decision rule that follows an IF-THEN structure. These rules aim to determine the most appropriate sequence for predicting outcomes based on the decisions made [17].

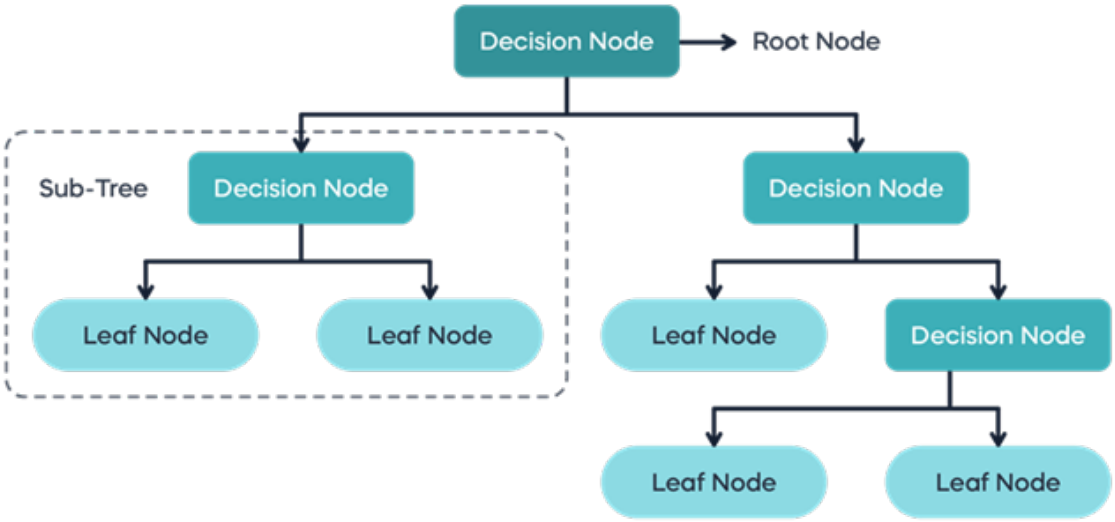


Figure 2: Decision trees

Discriminant analysis, similar to decision trees, is a classification method. It enables the grouping of data based on discriminant functions. The model is trained using previously classified data to predict the classes of new instances. In the quadratic discriminant method, the decision boundaries are quadratic in nature [18].

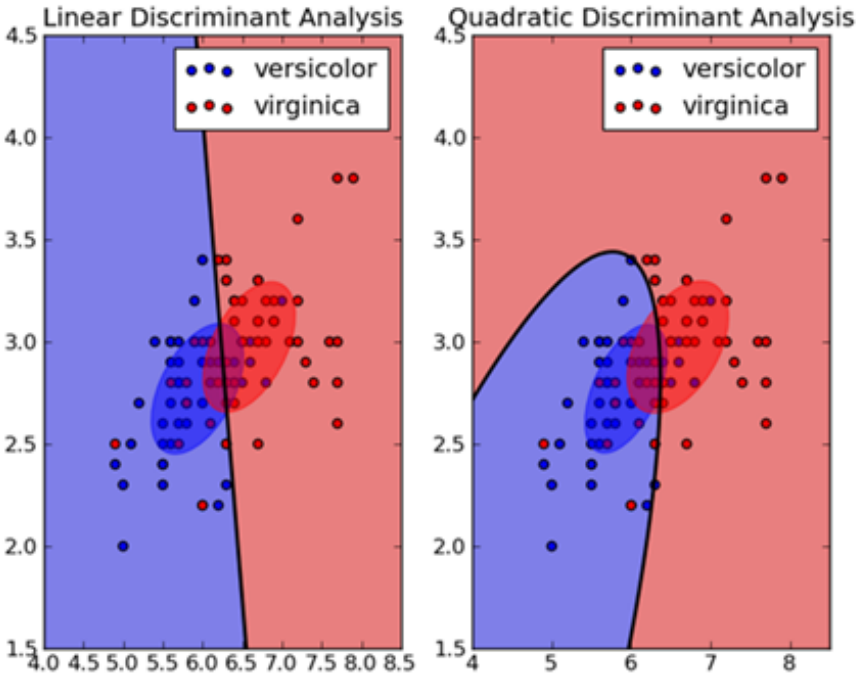


Figure 3: Linear And Quadratic Discriminant Analysis

The K-Nearest Neighbors (K-NN) classifier is a neighborhood-based classification method. It classifies a given data point by identifying its k nearest neighbors. A voting process is then conducted among these selected neighbors. This voting can be performed based on class density, distance-based weighting among the neighbors, or without any weighting at all. Finally, the data point is assigned to the class that receives the majority of votes [18].

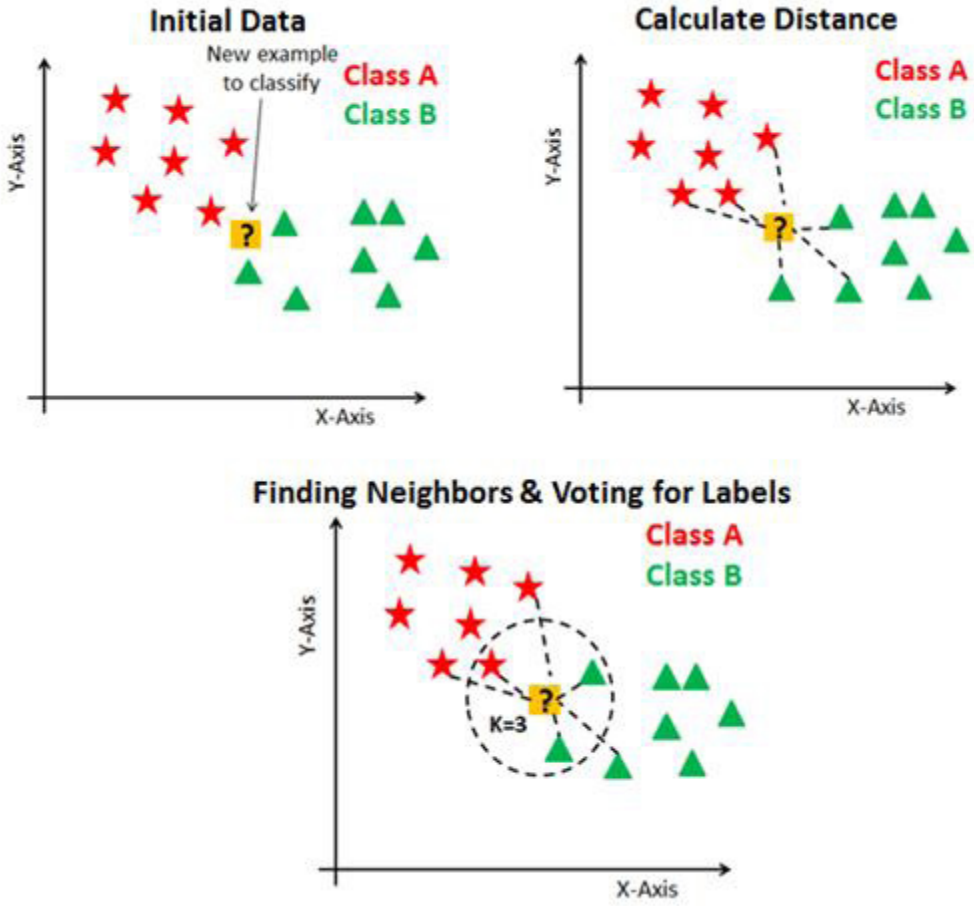


Figure 4: K-Nearest Neighbors

Support Vector Machines (SVM) perform classification by creating boundary lines between data points. These boundaries are defined in such a way that they are drawn at the maximum possible distance from the data points within each class. In the linear SVM model, these boundaries are constructed linearly. During the boundary construction process, lines are drawn through the support vectors in a manner that places them close to both classes. The final decision boundary is determined to lie between these vectors [19].

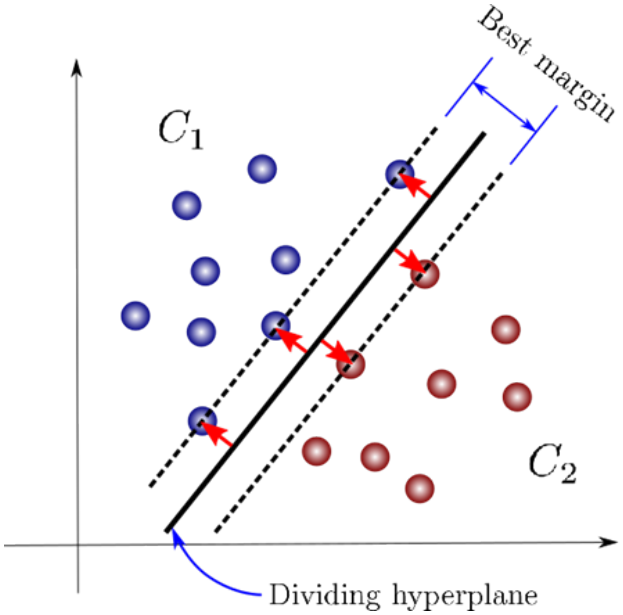


Figure 5: Vector Machines Classifier

AutoML libraries help individuals applying machine learning to identify the best-performing machine learning or deep learning models by experimenting with carefully selected algorithms and performing hyperparameter optimization on the data.

PyCaret is an open-source, low-code AutoML library that simplifies, accelerates, and automates machine learning workflows.

Case Study

We have a dataset of 219 records obtained from an autonomous Audi vehicle, capturing speed, previous distance, and current distance while the vehicle was in motion. Additionally, the Turkish Ministry of Transport has provided data on stopping distances for vehicles traveling on flat, dry asphalt roads (friction coefficient=0.60) based on their speed. Using these datasets, calculations were performed to derive the following variables.

$$TargetSpeed(x) = Speed(x) + (Distance(x) - Distance(x-1)) \times 3{,}6/0{,}2$$

Equation estimates the target speed by accounting for the change in distance over a 0.2-second interval, converting the unit to km/h using the multiplier 3.6.

$$\begin{aligned} StoppingDistance(x) \\ &= (NormalStoppingDistance(NormalizedSpeed(x)) \\ &\times (Speed(x) - NormalizedSpeed(x)) \\ &+ BrakeStoppingDistance(NormalizedSpeed(x) + 5) \\ &\times (NormalizedSpeed(x) + 5 - Speed(x)) \div 5 \end{aligned}$$

Equation calculates an interpolated stopping distance between two known braking distances, weighted by the vehicle's actual speed.

$$BrakingTime(x) = (BrakeTime(NormalizedSpeed(x)) \times (Speed(x) - NormalizedSpeed(x)) + BrakeTime(NormalizedSpeed(x) + 5) \times (NormalizedSpeed(x) + 5 - Speed(x))) \div 5$$

Equation represents the interpolated braking time corresponding to the actual speed of the vehicle.

$$TimeToCollision(x) = Distance(x) \div (Speed(x) - TargetSpeed(x))$$

Equation indicates how long it will take for the vehicle to reach the leading vehicle if the current speed differential is maintained.

$$AggressiveFollowing(x) = \begin{cases} 1, & TimeToCollision(x) - BrakingTime(x) < 0 \\ 0, & TimeToCollision(x) - BrakingTime(x) \geq 0 \end{cases}$$

Equation indicates whether the vehicle is following too closely, leaving insufficient time to brake safely.

$$RiskyProximity(x) = \begin{cases} 1, & StoppingDistance(x) - Distance(x) \geq 0 \\ 0, & StoppingDistance(x) - Distance(x) < 0 \end{cases}$$

Equation identifies cases where the stopping distance required exceeds the current distance to the leading vehicle, representing a high collision risk.

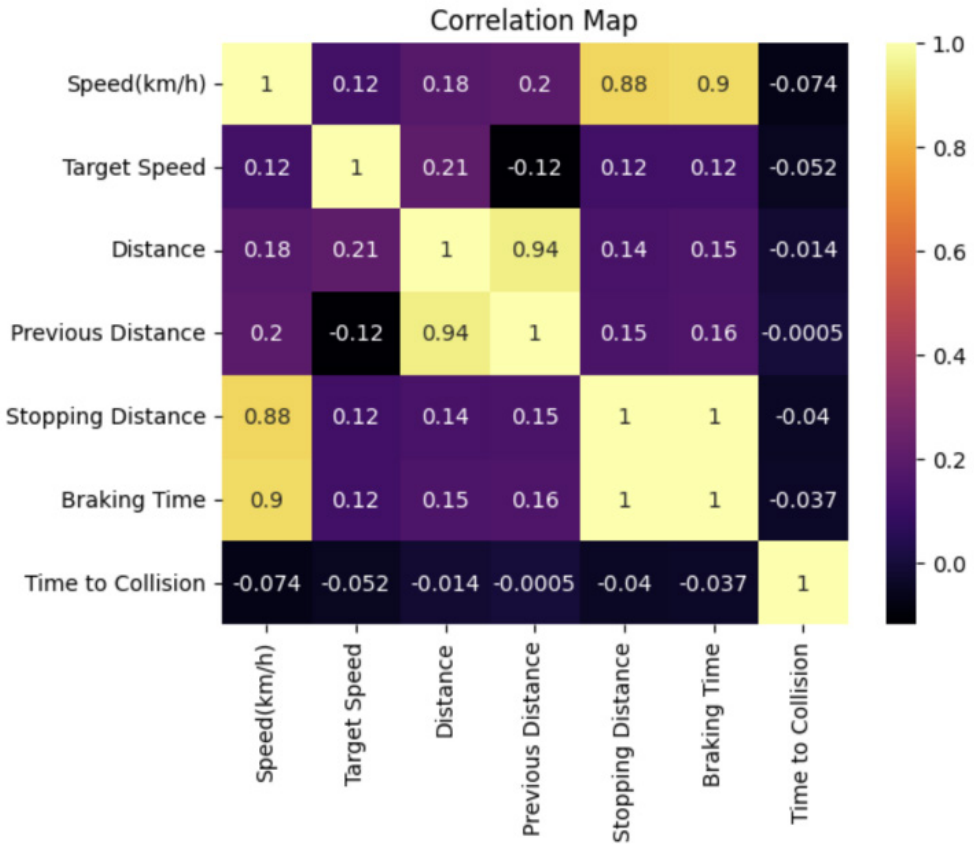


Figure 6: Correlation Map

In order to apply machine learning methods, the dataset previously transformed into a usable format during the data preprocessing steps was transferred to the platform where machine learning techniques were implemented. In all the machine learning methods used, the "distance" and "stopping distance" parameters in the dataset were defined as independent variables, while the "risky proximity" values were designated as the dependent variable. Using these defined dependent and independent variables, the dataset was split into training and testing subsets, with 70% allocated for training and 30% for testing.

A decision tree model was defined to branch at a maximum of 5 levels, using entropy as the decision criterion. The model was trained using the previously defined training data, X and Y. Classification was then predicted on the X test data using the trained model. The decision tree created with the specified features predicted risky convergence situations with an accuracy rate of 96%.

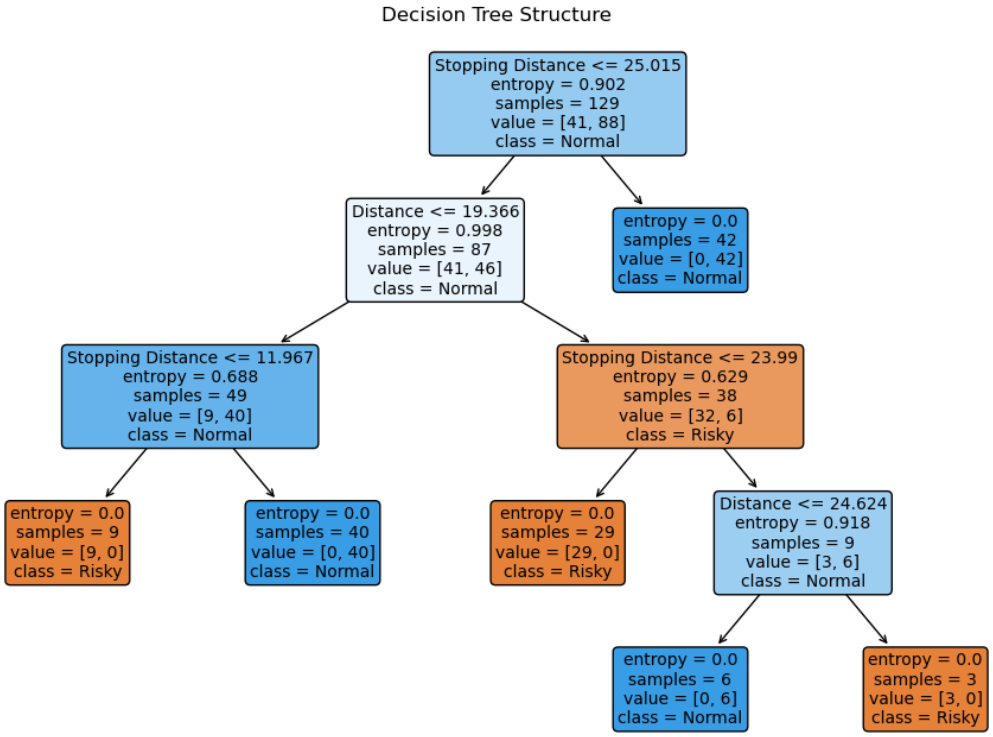


Figure 7: Decision Tree Structure

A new model is created for quadratic discriminant analysis. The model is trained with the specified data. Then, as with decision trees, classification prediction is performed. The prediction made using quadratic discriminant analysis yielded an accuracy rate of 91%.

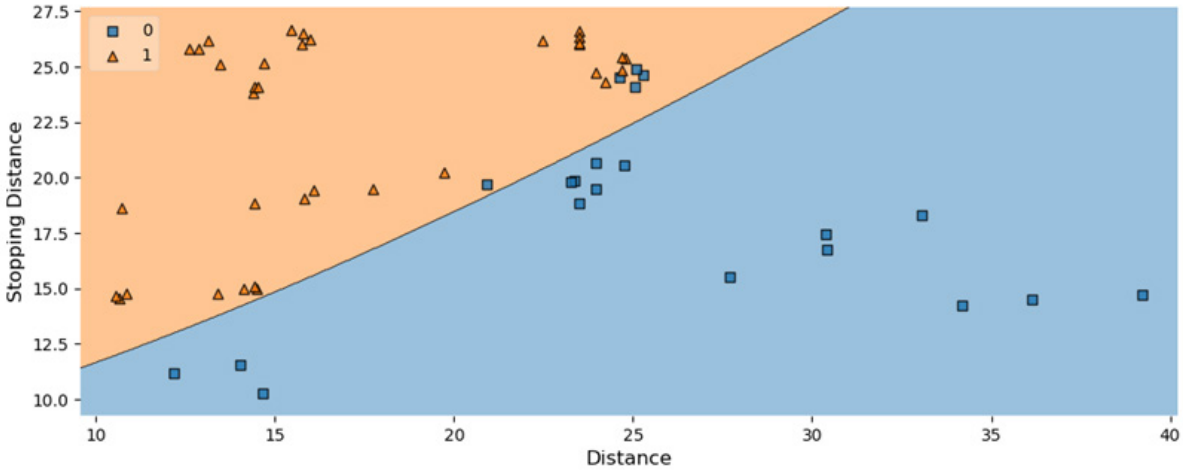


Figure 8: QDA Graph

A model is created for the k-nearest neighbours algorithm. While building the model, the number of neighbors was set to 5. After the model is created, it is trained using the X and Y data, and predictions are made as in the previous models. As a result of these predictions, we achieved an accuracy rate of 96%.

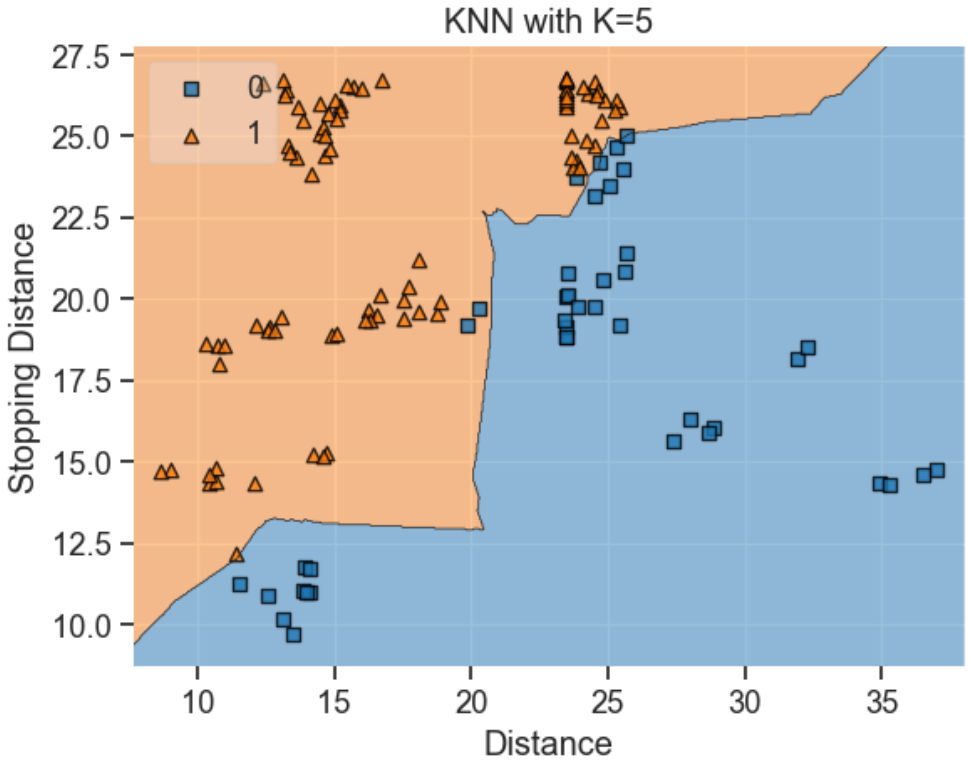


Figure 9: KNN Graph

A model is created for support vector machines. While building this model, the kernel function is chosen to be linear. As a result of training and predictions with our model, an accuracy rate of 93% was achieved.

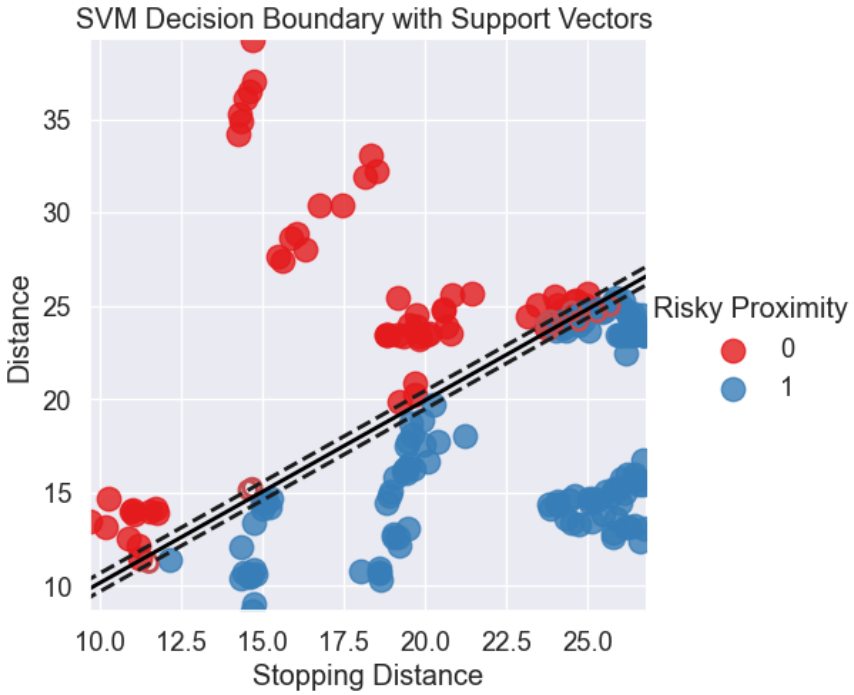


Figure 11: SVM Graph

The artificial neural network model was created with a hidden layer structure of 3 by 3 and 3000 iterations. The solver algorithm was selected as Adam. When the model was trained and used for prediction, it achieved an approximate accuracy rate of 96%.

Through the visualizations generated using the PyCaret library, we obtained the following graphs.

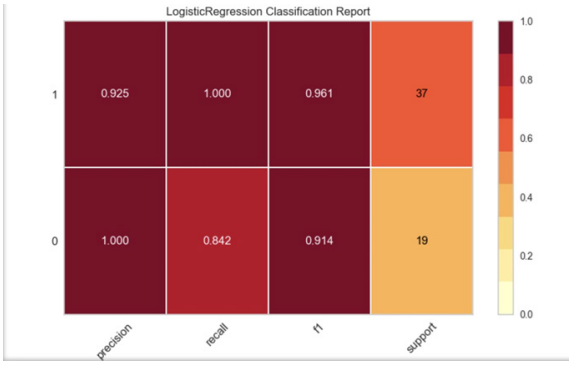


Figure 11: Classification Report

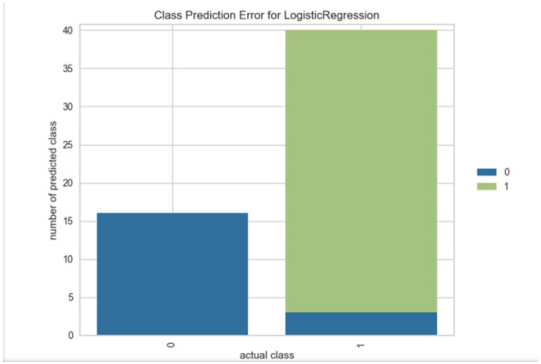


Figure 12: Class Prediction Error

The model's performance is generally balanced across classes; however, improving the recall for class 0 could enhance the model further.

Discussion

The findings of this study highlight the significant role of machine learning and AutoML techniques in enhancing collision detection accuracy for autonomous vehicles. By evaluating multiple supervised learning models, we observed how different algorithms perform under real-world driving conditions.

One of the key insights derived from the study is the superiority of AutoML based optimization over traditional machine learning models. While individual algorithms such as artificial neural networks (ANN), decision trees (DT), and support vector machines (SVM) demonstrated strong predictive capabilities, their performance varied based on dataset complexity and parameter selection. AutoML, on the other hand, automates model selection and hyperparameter tuning, leading to higher accuracy rates with minimal manual intervention.

Another critical observation is the importance of feature engineering in improving prediction performance. Distance based attributes such as stopping distance, braking time, and time to collision proved to be the most influential factors in predicting risky proximity events. This finding aligns with existing literature, which emphasizes the necessity of incorporating real-time environmental data for more precise collision forecasting. Models that lacked optimized feature selection showed relatively lower accuracy rates, reinforcing the need for structured input variables in autonomous vehicle decision-making.

Furthermore, the study confirms that real time data noise and outliers significantly impact model reliability. The initial dataset contained inconsistencies caused by sensor noise and unexpected speed fluctuations, which were mitigated through preprocessing techniques. Filtering out noise and normalizing speed parameters improved classification stability, allowing machine learning models to differentiate between normal and high risk interactions more effectively.

In conclusion, this research establishes a data driven framework for collision detection, demonstrating that AutoML methodologies surpass traditional machine learning techniques in efficiency, accuracy, and adaptability. Future studies should focus on multi-modal data fusion, combining LIDAR, radar, and vision-based inputs to develop more comprehensive autonomous decision-making systems.

Conclusion

Using Method	Algorithm	Accuracy
Machine Learning	Artificial Neural Networks (ANN)	0.964
	Decision Tree (DT)	0.96
	Quadratic Dskriminant Analysis (QDA)	0.91
	K-Nearest Neighbors (K-NN)	0.96
	Support Vector Machines Linear Kernal (SVM Linear Kernal)	0.93
AutoML	Logistic Regression (LR)	0.984

Among the supervised machine learning applications, the most accurate method was the artificial neural network (ANN) classifier with a 3x3 hidden layer structure, achieving 96% accuracy. To further improve predictive performance, Automated Machine Learning (AutoML) was implemented. Although the training and testing data sizes were the same for both ML and AutoML models, AutoML yielded better results. This is because AutoML systematically evaluates multiple machine learning algorithms to determine the most effective model. Through this process, logistic regression was identified as the best performing model, achieving the highest classification accuracy among all tested algorithms.

REFERENCES

[1] Y. Taştan and H. Kaymaz, "Otonom Araçların Yaygınlaşmasının Önündeki Zorluklar Difficulties in the Proliferation of Autonomous Vehicles," Int. J. Adv. Eng. Pure Sci, vol. 2021, no. 2, pp. 195–209, doi: 10.7240/jeps.741594.

[2] G. J. C. L. B. Maurer Markus, "Autonomous Vehicles: Challenges, Opportunities, and Future Directions," Proceedings of the IEEE, vol. 104, no. 6, pp. 1162–1186, Jun. 2016.

[3] G. Bakioğlu and A. O. Atahan, "Otonom Araçların Benimsenmesi ve Güvenlik Algılarının İncelenmesi," European Journal of Science and Technology Special Issue, vol. 32, no. Special Issue, pp. 633–639, Dec. 2021, doi: 10.31590/ejosat.1039725.

[4] F. N. Ortataş and E. Çetin, "Elektrikli ve Otonom Araçlarda Makine Öğrenmesi Kullanarak Trafik Levhaları Tanıma ve Simülasyon Uygulaması," El-Cezeri Journal of Science and Engineering, vol. 8, no. 3, pp. 1081–1092, 2021, doi: 10.31202/ecjse.867733.

[5] K. Baltacıoğlu, M. T. Başar, M. Karaaslan, Z. Özer, and S. Öcal, "Görüntü İşleme Yoluyla Otonom Tren-Hayvan Kazası Önleme Sistemi," Demiryolu Mühendisliği, no. 18, pp. 150–161, Jul. 2023, doi: 10.47072/demiryolu.1300927.

[6] M. S. Bingöl, Ç. Kaymak, and A. Uçar, "Derin Öğrenme Kullanarak Otonom Araçların İnsan Sürüşünden Öğrenmesi," Fırat Üniversitesi Müh. Bil. Dergisi, vol. 31, no. 1, pp. 177–185, 2019.

[7] D.-H. Lee, K.-L. Chen, K.-H. Liou, C.-L. Liu, and J.-L. Liu, "Deep Learning and Control Algorithms of Direct Perception for Autonomous Driving".

[8] M. Hussain and J. Eui-Hong, "DeepGuard: A Framework for Safeguarding Autonomous Driving Systems from Inconsistent Behaviours."

[9] G. Li et al., "Risk assessment based collision avoidance decision-making for autonomous vehicles in multi-scenarios," Transp Res Part C Emerg Technol, vol. 122, p. 102820, 2021, doi: https://doi.org/10.1016/j.trc.2020.102820.

[10] R. Dyro et al., "Autonomous Vehicle Risk Assessment," Sep. 2020.

[11] M. Strickland, G. Fainekos, and H. Ben Amor, "Deep predictive models for collision risk assessment in autonomous driving," Proc IEEE Int Conf Robot Autom, pp. 4685–4692, Sep. 2018, doi: 10.1109/ICRA.2018.8461160.

[12] M. Bilgin, "Gerçek Veri Setlerinde Klasik Makine Öğrenmesi Yöntemlerinin Performans Analizi," Bursa.

[13] M. V. Keskin, "Büyük Veride Makine Öğrenmesi Uygulaması," İstanbul, 2018.

[14] D. R. Dündar, İ. Sarıçiçek, E. Çınar, and A. Yazıcı, "Kestirimci Bakımda Makine Öğrenmesi: Literatür Araştırması," J ESOGU Engin Arch Fac, vol. 29, no. 2, pp. 256–276, 2021, Accessed: Jan. 11, 2025. [Online]. Available: http://orcid.org/0000-0002-3528-7342http://orcid.org/0000-0003-3189-7247http://orcid.org/0000-0001-5589-2032

[15] E. Öztemel, Yapay Sinir Ağları, 3. Basım. İstanbul, Ankara, İzmir, Adana: Papatya Yayıncılık Eğitim, 2012. Accessed: Jan. 11, 2025. [Online]. Available: www.papatya.gen.tr

[16] B. Daş and İ. Türkoğlu, "DNA Dizilimlerinin Sınıflandırılmasında Karar Ağacı Algoritmalarının Karşılaştırılması," Eleco 2014 Elektrik – Elektronik – Bilgisayar ve Biyomedikal Mühendisliği Sempozyumu, Nov. 2014, Accessed: Jan. 11, 2025. [Online]. Available: http://www.ncbi.nlm.nih.gov/Genbank/genomes/bacteria.

[17] G. G. Emel and Ç. Taşkın, "Veri Madenciliğinde Karar Ağaçları ve Bir Satış Analizi Uygulaması," Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi, vol. 6, no. 2, Dec. 2005.

[18] B. Özdet and S. İçer, "Akciğer Bilgisayarlı Tomografi Görüntülerinde Görüntü İşleme Uygulamaları ile Tümörlerinin Tespit Edilmesi," Uludağ University Journal of The Faculty of Engineering, pp. 135–150, Mar. 2022, doi: 10.17482/uumfd.947619.

[19] K. Kardaş and A. Güvenir, "Kısa Sınavların, Ödevlerin ve Projelerin Dönem Sonu Sınavına Olan Etkilerinin Farklı Makine Öğrenmesi Teknikleri ile Araştırılması," EMO Bilimsel Dergi, vol. 10, no. 1, pp. 22–26, Jun. 2020.

COMPARATIVE ANALYSIS OF AI MODELS FOR PREDICTING CUSTOMER PURCHASE BEHAVIOR IN E-COMMERCE

Shefitjon Bregu*
Department of Computer Engineering, Graduate School of Natural and Applied Sciences, Eskisehir Osmangazi University, Eskisehir, Türkiye

Yusuf Kartal
Department of Computer Engineering, Graduate School of Natural and Applied Sciences, Eskisehir Osmangazi University, Eskisehir, Türkiye
*(Corresponding author) *shefitjonbregu@gmail.com*

ABSTRACT

Predicting customer purchasing behavior has become an essential part of modern e-commerce op- erations. By utilizing artificial intelligence (AI) models, businesses can gain insights into customer patterns, enhance marketing strategies, and streamline inventory control. This article outlines and ana- lyzes three major AI approaches commonly used to predict purchasing behavior: decision trees, support vector machines (SVMs), and neural networks. Rather than focusing on specific empirical results, this article provides conceptual guidance and practical recommendations for e-commerce practitioners re- garding model selection. It also discusses future directions in AI-driven customer behavior research, emphasizing the importance of model interpretability, data integration, and ethical implications.

Keywords: AI, decision trees, support vector machines, neural networks, e-commerce, cus- tomer behavior.

Introduction

With the emergence of large-scale data collection, e-commerce platforms are increasingly turn- ing to artificial intelligence (AI) techniques to gain actionable insights into customer behavior. Studies have demonstrated that predicting whether a customer will make a purchase—and if so, when and how much—is critical for optimizing marketing, inventory management, and dynamic pricing. For example, research on customer purchase behavior prediction [4, 5] has shown that machine learning models can capture complex buying patterns. In addition, investigations into dynamic pricing methods [6] confirm that intelligent pricing strategies offer competitive advantages. However, choosing the most appropri- ate AI model is challenging. Simpler models such as decision trees are easy to interpret and quick to train [1], whereas complex models like neural networks may achieve higher accuracy but at the cost of in- creased computational demands and reduced transparency [2, 6]. Support vector machines (SVMs) offer a middle ground by balancing performance and complexity [2,4]. This paper aims to present a high-level overview of

these models for purchase behavior prediction while emphasizing conceptual understanding, methodological design, and practical observations drawn from recent studies.

Literature Review

Importance of AI in E-commerce

The integration of AI-driven tools in e-commerce has been well documented. Researchers have shown that AI-based methods can significantly improve various aspects of online retail, such as cus- tomer segmentation, demand forecasting, recommendation systems, and targeted advertising. Gomes and Meisen [3] review segmentation methods for personalized customer targeting, while studies such as Deniz and Bülbül [4] and Cirqueira et al. [5] illustrate how machine learning models capture nonlinear purchase patterns. Additionally, work on dynamic pricing [6] and inventory forecasting [7] underscores AI's broad potential to transform operational strategies.

Decision Trees and Interpretability

Decision trees have become popular in various business environments due to their simple, rule- based structure. Random forests, as a collection of decision trees, mitigate issues such as variance and overfitting inherent in single-tree models [1]. However, single-decision trees continue to be attractive to smaller businesses due to their simplicity, reduced computational requirements, and high interpretability. They clearly show how a set of features (e.g., age, browsing history, cart additions) lead to a particular purchase prediction, making the output easier to explain to non-technical stakeholders. Huda et al. [1] found that while decision trees might yield slightly lower overall accuracy than ensemble methods or neural networks, their explainability is highly valued in resource-constrained environments.

Support Vector Machines for High-Dimensional Data

Support vector machines (SVMs) have proven effective in handling high-dimensional data through kernel functions. Deniz and Bülbül [4] compared SVMs against other machine learning models on large- scale e-commerce datasets and demonstrated that SVMs can achieve competitive F1-scores, albeit with longer training times that may limit their use in real-time applications.

Neural Networks for Complex Patterns

Neural networks, particularly deep learning models, have received significant attention due to their ability to learn highly complex, nonlinear relationships. Pondel et al. [2] show that deep neural net- works can capture intricate feature interactions in large e-commerce datasets, such as those arising from clickstream and social media data. However, these models typically require substantial computational resources and present challenges in interpretability—a point further noted by Tomitza et al. [6]. Re- cent comparative studies [5] have underscored the trade-offs between accuracy and explainability among these models.

Emerging Trends and Hybrid Methods

Recent literature indicates a growing interest in hybrid methods that combine the strengths of differ- ent approaches. Techniques that integrate decision trees with neural networks can capitalize on both in- terpretability and predictive power. Studies on dynamic pricing [6] and inventory forecasting [7] demon- strate that merging internal transactional data with external signals (e.g., macroeconomic indicators or social media sentiment) can improve model performance. Moreover, research into explainable AI [8] has promoted methods such as SHAP and LIME to render even complex models more transparent.

Comparative Studies and Empirical Results

While many studies focus on a single method, several works conduct direct comparisons of decision trees, SVMs, and neural networks on identical datasets. For example:

- Deniz and Bülbül [4] reported that neural networks often achieve higher prediction accuracy (around 0.92 on certain tasks) than decision trees (around 0.85) or SVMs (around 0.89), albeit at the expense of increased computational complexity.

- Cirqueira et al. [5] provided a comprehensive comparison of purchase behavior prediction models using clickstream data.

- Huda et al. [1] emphasized the advantage of decision trees in interpretability, even if they some- times trail in pure predictive performance.

Table 1: Summary of Key Findings from Literature

Study	Models Compared	Dataset/Focus	Notable Findings
Huda et al. [1]	Decision Trees, etc.	Segmentation and churn classification in e-commerce	High interpretability and fast training
Deniz & Bülbül [4]	Decision Trees, SVM, Neural Networks	Large-scale purchase behavior prediction	Neural networks achieve highest accuracy; SVM is competitive; decision trees are fastest
Cirqueira et al. [5]	Various ML approaches	Clickstream-based purchase prediction	Trade-offs between model accuracy and computational requirements
Tomitza et al. [6]	Dynamic Pricing Models	Dynamic pricing strategies based on AI	Benefits in accuracy and flexibility

As seen in the cited works, neural networks often top the performance metrics but demand more computational resources, whereas SVMs strike a middle ground in both complexity and performance, and decision trees remain favored for their simplicity and ease of explanation. Additionally, based on a local e-commerce dataset of 100,000 transactions, a common trend emerges: neural networks can achieve superior accuracy and AUC with higher computational costs and lower interpretability, while decision trees train more quickly and offer high transparency at some sacrifice in predictive power. SVMs balance accuracy with moderate training times, especially for medium-scale data.

Methodology

This paper discusses a conceptual methodology rather than presenting new experimental results.

Data Collection and Preprocessing

Data is typically collected from an e-commerce platform and may include:

- **Transaction History:** Previous purchase records (order amount, products, timestamps).
- **Browsing Behavior:** Page visits, time spent per page, clickstream patterns.
- **Demographic Data:** Age, gender, location, and other relevant details (subject to ethical guide- lines).
- **User Engagement:** Interactions such as emails opened or social media interactions.

Preprocessing steps include cleaning missing values, encoding categorical variables, and normaliz- ing numeric attributes. Feature engineering may be performed to derive new variables that better capture user intent. For high-dimensional textual data, techniques such as TF-IDF or word embeddings [9] can be used.

Model Selection

This paper focuses on three main AI models:

- **Decision Trees:** Using algorithms like CART, C4.5, or ID3. Pros include high interpretability and relatively fast training; cons include potential overfitting and limited capture of complex interac- tions.
- **Support Vector Machines (SVMs):** These models identify optimal hyperplanes that separate classes. Pros include effectiveness in high-dimensional spaces; cons include slower training times for large datasets and the need for parameter tuning.
- **Neural Networks:** Typically feedforward networks with one or more hidden layers. Pros include the ability to model highly nonlinear relationships; cons include high computational cost and re- duced interpretability.

Training and Validation Approach

- **Data Splitting:** Typically, 70–80% of the data is used for training, 10–15% for validation, and 10–15% for testing.
- **Model Training:** Each model is trained using the training set with appropriate hyperparameters (e.g., maximum tree depth, kernel type, number of neural network layers).
- **Cross Validation:** Techniques such as k-fold cross-validation ensure generalizability.
- **Hyperparameter Tuning:** Methods such as grid search, random search, or Bayesian optimization are used to optimize model configurations.

• **Model Evaluation:** Performance is measured using metrics such as accuracy, precision, recall, F1 score, and the area under the ROC curve (AUC). Computational efficiency is also evaluated.

Discussion

Selecting an appropriate AI model for predicting customer purchase behavior requires balancing multiple factors:

• **Interpretability and Complexity:** Decision trees offer clear, interpretable decision rules, whereas neural networks, despite their predictive power, are more opaque [2, 6].

• **Data Scale and Complexity:** Large-scale datasets may fully leverage the capacity of neural net- works, though at a higher computational cost. SVMs can provide a compromise for medium-scale data [4].

• **Resource Availability:** Organizations with limited computational resources may prefer decision trees or SVMs over computationally intensive neural networks.

• **Data Quality:** Although neural networks can extract complex relationships, they require high- quality, well-preprocessed data. Decision trees and SVMs may be more robust on smaller or moderately sized datasets.

• **Future-proofing and Scalability:** For rapidly growing e-commerce platforms, scalable models that can adapt to increased data volumes are crucial. Hybrid approaches that combine decision trees and neural networks are emerging as promising solutions.

Conclusion

This paper has presented a conceptual framework for comparing three AI models—decision trees, support vector machines, and neural networks—in the context of predicting customer purchase behavior in e-commerce. Each model presents trade-offs between interpretability, predictive accuracy, and com- putational requirements. Decision trees are advantageous for their transparency and rapid deployment, SVMs balance accuracy with efficiency, and neural networks offer high accuracy for complex datasets while demanding more computational resources. Ultimately, model selection should be driven by the specific operational needs and resource constraints of the business.

Future Research Directions

Future research may explore:

• **Hybrid Models:** Combining decision trees with neural networks to balance interpretability and predictive power.

• **Transfer Learning:** Utilizing pre-trained models to reduce data requirements and training time.

• **Explainable AI (XAI):** Integrating methods such as LIME and SHAP to improve the transparency of complex models.

• **Ethical and Privacy Considerations:** Addressing issues related to data privacy, user consent, and data security in AI model design.

• **Integration of External Data:** Enhancing model performance by incorporating social media sen- timent, macroeconomic indicators, or competitor pricing data.

Declaration of Interest: The authors declare that there is no conflict of interest.

Acknowledgements: The authors acknowledge the support of Eskisehir Osmangazi University and thank colleagues for their constructive feedback.

Author Contributions: Shefitjon Bregu led the conceptualization of the study, carried out the comprehensive literature review, designed the methodological framework and Yusuf Kartal contributed equally to the conception, writing, and revision of this paper.

REFERENCES

[1] Huda, I., Suhendra, A. A., & Bijaksana, M. A. (2023). Design of Prediction Model using Data Mining for Segmentation and Classification: Customer Churn in E-Commerce Mall.

[2] Pondel, M., Wuczyn´ski, M., Gryncewicz, W., Łysik, Ł., Hernes, M., Rot, A., & Kozina, A. (2021). Deep Learning for Customer Churn Prediction in E-Commerce Decision Support. *Business Information Systems*, 1, 3–12.

[3] Gomes, M. A., & Meisen, T. (2023). A review on customer segmentation methods for personalized customer targeting in e-commerce use cases. *Information Systems and e-Business Management*, 21, 527–570.

[4] Deniz, E., & Bülbül, S. Ç. (2024). Predicting Customer Purchase Behavior Using Machine Learning Models.

[5] Cirqueira, D., Hofer, M., Nedbal, D., Helfert, M., & Bezbradica, M. (2020). Customer Purchase Behavior Prediction in E-Commerce: A Conceptual Framework and Research Agenda.

[6] Tomitza, C., Ibrahimli, U., & Herm, L. (2024). AI-Based Methods of Dynamic Pricing in E- Commerce: A Systematization of Literature.

[7] Tang, Y., Chau, K. Y., & Lau, Y. (2023). Data-Intensive Inventory Forecasting with Artificial Intelligence Models for Cross-Border E-Commerce Service Automation.

[8] Barredo Arrieta, A., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina, D., Benjakins, R., Chatila, R., & Herrera, F. (2020). Explainable Artifi- cial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82–115.

[9] Mikolov, T., Chen, K., Corrado, G., & Dean, J. (2013). Efficient Estimation of Word Representations in Vector Space. *arXiv preprint* arXiv:1301.3781.

COMPARATIVE ANALYSIS OF BRAIN TUMOR MRI SCANS SEGMENTATION USING SEGFORMER TRANSFORMER AND HYBRID

Muhammad Ahmed Qureshi
FAST NUCES, School of Computing, Islamabad, Pakistan

Mahnoor Tauqeer
MCS, Department of Electrical Engineering, Islamabad, Pakistan

Yashfa Kamran*
NUST, School of Electrical Engineering and CS, Islamabad, Pakistan
(Corresponding author) yashfakamran.92@gmail.com

ABSTRACT

The accurate segmentation of MRI brain tumors is necessary for early diagnosis and selection of specific treatment to increase the survival rates of people having life threatening brain tumors. Incorrect classification of the brain tumor has negative consequences. The manual process of tumor detection takes a lot of time due to various information and types of tumors in brain. The currently existing models such as Convolutional Neural Networks (CNNs) have been widely used for this purpose but that face challenges in correctly capturing complex tumor structures due to the limited receptive field of CNNs. So, an intelligent diagnostic system is needed. To address this challenge, we propose a SegFormer transformer model as an alternative solution for brain tumor segmentation and compare it with hybrid CNN model. It utilizes self-attention mechanism that helps to capture long-range dependencies and detailed spatial features of MRI images. We used the Low-Grade Glioma (LGG) Segmentation Dataset and analyzed the performance of SegFormer by evaluating segmentation accuracy, Dice coefficient, sensitivity, and specificity. Our results indicate that SegFormer outperforms hybrid CNN by generalizing well on test data. This shows great potential of Segformer to improve the diagnosis of brain tumors and encourage its adoption in medical imaging.

Keywords: Brain tumor detection, diagnostic accuracy, MRI images, SegFormer.

Introduction

The brain is a primary and most complicated component of human nervous system which is responsible for managing majority of body functions. A brain tumor is a collection of abnormal cells that develops in the inflexible skull enclosing the brain. These tumors are one of the most complex medical conditions and thousands of people suffer from them globally every year. Around 2.5 lac people are affected by brain tumors yearly. People having these tumors can suffer from severe neurological symptoms like headaches and seizures etc. Their daily life activities can be affected from cognitive impairments such as

memory loss or difficulty in concentration. The tumors in regions of brains that regulate emotions such as frontal lobe might cause sudden mood changes, anxiety, depression and social withdrawal. There are three basic types of brain tumors; benign, pre-malignant and malignant. Among the malignant tumors which are most life threatening for people, gliomas are the most common that may lead to poor prognoses if not treated early. The early diagnosis and treatment planning is necessary to improve the survival rates of patients as these brain tumors are life threatening. Medical Resonance Imaging (MRI) is a helpful tool to visualize brain tumors but the problem lies in the manual analysis of MRI scan. The process is labor-intensive, time taking and there might be error while diagnosing. Also, different doctors might have different suggestions. By automating brain tumor detection using deep learning models, we can reduce error in tumor detection, provide accurate treatments to patients and fasten the diagnosis process especially in countries with limited resources and specialized doctors. People often face great challenges in their diagnosis and treatment as these tumors have intricate morphology and are variable in appearance. The MRI data is quite complex as the scans are different in intensity, resolution and shape of tumors which makes the detection difficult. The model is supposed to be generalized well for different patients and high accuracy should be obtained in training the model. The early diagnosis by accurate segmentation and proper treatment can save the lives of many people and improve their quality of life. There is a growing trend of using AI and machine learning in health sectors and different techniques have revolutionized medical diagnosis of tumors and diseases. However, the use of deep learning in specialized area like brain tumor detection is still emerging. Deep learning models are continuously evolving so there is a great emphasis on improving accuracy of predictions for various datasets. Automated approaches for brain tumor segmentation have revolutionized medical imaging and various research areas are being explored for medical image analysis. Many experts are continuously trying to improve models so that they can accurately detect tumors. Among the techniques, Convolutional Neural Networks (CNNs) are quite popular due to their proven success in analyzing diverse medical images. They have the ability to process and identify complex images having critical patterns in them. U-Net has an encoder-decoder structure and is widely regarded as a benchmark in medical image segmentation. It has skip connections that preserve spatial information to segment images. The 3D U-Net and Attention U-Net have further improved segmentation accuracy by using attention mechanisms to focus on tumor regions of organs. Other notable models include VGGNet, ResNet and DenseNet etc. They utilize deep hierarchical feature extraction to achieve robust performance in identifying tumors. Fully Convolutional Networks (FCNs) have also been used for pixel-wise tumor segmentation. The hybrid models combining CNNs with graph neural networks (GNNs) or recurrent neural networks (RNNs) have addressed complex tumor morphologies. Although, these techniques are successful but CNNs still have limitations in effectively handling certain complexities of tumor segmentation. They face difficulty in capturing long-range dependencies due to their limited receptive fields. They are often unable to capture fine-grained spatial features and fail to generalize well across diverse appearances and characteristics like type, location and size of tumors etc. These limitations motivate the exploration of novel architectures like transformers for image segmentation.

In recent years, transformer-based architectures are preferred instead of traditional CNNs in the field of medical imaging. Initially, they were designed for natural language processing but due to their unique self-attention mechanism, they are being rapidly adopted for medical imaging. The transformer-based models have the property that they can capture long-range dependencies and detailed spatial features or relationships in data. This overcomes the limitations of CNNs that have fixed receptive field. CNNs rely on local feature extraction through convolution layers. On the other hand, transformers process global context i.e. they can effectively understand complex and irregular structures like that of brain tumors in our case. Segformer, which is a transformer-based architecture is proved to be an effective approach for image segmentation tasks. It combines the strengths of transformers and employs a hierarchical structure to extract multi-scale features and utilizes self-attention to maintain spatial coherence across images. Also, it is a lightweight design i.e. it does not rely heavily on extensive computational resources so it is highly efficient and a practical approach for clinical applications. This architecture is suitable for accurate segmentation of diverse structures in medical images but there is limited research available on transformer-based model applications. This gap presents an opportunity to further explore their performance in crucial medical image segmentation tasks. Solving the problem through this technique could reduce the time it takes to detect brain tumors, lessen the workload on doctors, open doors for future research regarding refining the top models to achieve best accuracy and encourage further adoption of AI in medical imaging.

For the application of brain tumor segmentation, the potential of segFormer has not been fully realized. Addressing this gap is necessary to improve the diagnostic tools for enhancing the patient's outcomes. SegFormer has the ability to contribute to more accurate delineation of tumor boundaries with the help of its hierarchical structure and self-attention mechanism. In this study, we evaluate the performance of Segformer in brain tumor segmentation by using the Low-Grade Glioma (LGG) Segmentation Dataset. The data set includes diverse structures and appearances of brain tumors of multiple patients that helps to test the adaptability of segmentation model. Our aim is to highlight the practical benefits of adopting transformer-based architectures in medical imaging by analyzing SegFormer's ability to accurately segment tumors and comparing it to traditional approaches. The contributions of this study are as follows:

- Demonstration of the potential of SegFormer, a transformer-based model, for accurate brain tumor segmentation in medical imaging and assessing its efficiency by focusing on its precision and adaptability in delineating complex and diverse tumor structures.

- Conduct experiment using the LGG Segmentation Dataset obtained from Kaggle which contains brain MRI images along with manual FLAIR abnormality segmentation masks to test the adaptability and generalization of Segformer model in detecting tumors of different appearances and characteristics.

- Compare the Segformer with traditional hybrid CNN model in detection of brain tumors from MRI images. This shows its superiority by depicting more accurate segmentation precision, adaptability and accuracy.

• Our findings show improved accuracy and paving the way for further research and innovations in automated diagnostic tools. This will increase the adoption of transformers in healthcare.

The remaining sections of this paper are arranged as follows: the depth analysis of relevant work related to our problem i.e. the studies on brain tumors using various machine learning models is presented in Section 2. Our proposed work is described in section 3 where we described each step in detail and suggested the model as well. The results and in-depth analysis of the experiment is discussed in section 4. In the last section 5, conclusions of the study and future direction of the research work has been explored.

Existing Technologies

The advancement of deep learning in brain tumor detection has reached a rapid pace, mostly based on new architectures with a hybrid model that have been exploited by many researchers to provide the optimal performance of segmentation first and then classification. While experimental results via CNNs, multimodal approaches, and ensemble methods show promise, the components can be optimized via transformers such as Segformer for state-of-the-art tumor detection and segmentation performance. It is studied by Jahan and Tripathi in [1] that the existing hybrid CNN framework for detection of brain tumor in fMRI images shows interesting results showing pandemic using the deep learning-based hybrid architectures that not only help improve the quality of the extracted features but also assist in better classification. While CNN has plenty of advantages in dealing with high-dimensional fMRI data, we believe that a transformer-based segmentation by an additional model like Segformer could be adapted to better improve detection via precisely extracting complex brain structures. In [2] Srinivasan et al. proposed a hybrid CNN model for multi-class brain tumor classification that addressed image augmentation, and improved differentiation of tumor types. This explicit multiclass distinction underscores the need of models that can be accurately and efficiently trained to multimodal representations, which are further supported by Segformer constituent ability of feature representation and precision in discerning minute differences between tumor types. In [3] Santhosh Kumar et al. worked on multi-modality parallel capsule network for brain tumor segmentation, by leveraging different imaging modalities for advanced feature extraction, this approach maps very well with the segmentation capabilities of Segformer through transformers, which allow details specifically at the edges of tumors to be captured accurately across modalities. It is studied in [4] that brain tumor classification using a deep learning model ensemble and found that aggregating many architectures leads to more robust performance. In [5], the study indicates that the ensemble methods provide error reduction and, therefore, Segformer could be a backbone segmentation model that integrates its outcomes through an ensemble method to encompass more global and local tumor features, Grampurohit et al. investigated CNN architectures (VGG-16 for example) with respect to tumor detection where they concluded that although the deeper networks deliver much rewarding results, they may need more preprocessing to gain from their full capacity. This could indicate a potential in Segformer, which leverages its transformer architecture to accumulate spatially rich yet simple information with minimal preprocessing before segmentation. In [6] Hassan et al. used different machine learning algorithms such

as Gradient Boosting, Random Forest were compared on MRI datasets for detection of brain tumor; used the same techniques. Although classic models achieved competitive results, their inability to construct complex representations from high dimensional MRI data underlies the importance of transformer models such as Segformer which demonstrates strong accuracy-efficiency trade-off for this type of architectures. In [7] Nagaraju et al. worked on machine learning in brain tumor classification — with a focus on radiofrequency. Our results demonstrate the ability of state-of-the-art ML approaches to outperform current classification capabilities, showing promise for tuning Segformer for MRI and radiofrequency images to attain diagnostic accuracy through improved segmentation. Bhagwan et al. Gokila Brindha et. al in [8] showed that while several ML models and features extracted from different modalities were studied, a track toward using more than one modality and feature extraction methods to improve accuracy in tumor detection. Segformer, however, can capitalize on this trend based on the transformer nature to achieve segmentation accuracy across many domains of images. Soomro et al. Kang et al. [9] published a review on image segmentation with specific relevance to MR brain tumor detection and highlighted the need for CNNs to address complications in image features. They further observed that, though CNNs may be adequate for lower-resolution tumors, segmentation of high-resolution and complex tumors might require a model with stronger long-range contextual encoding ability than conventional methods such as CNNs. In a systematic review of the use of different multimedia applications for brain tumor detection, Dhole and Dixit [10] pointed out that advancements in deep learning architectures provide opportunities for leveraging improved segmentation. So Segformer, a transformer-based segmentation model may provide enhanced segmentation performance in other imaging modalities similar to its potential indicated from their study. Both convolutional neural networks (CNNs) with strong local information extraction capacities and Transformers with excellent global representation capacities have achieved remarkable performance in medical image segmentation proposed by Yu Liu and Yize m [11] Ahmad and Chaudary studied that the effectiveness of deep transfer learning networks for brain tumor detection using MRI images has a potential in getting accurate results [12]. The application of Convolutional Neural Networks (CNNs) for brain tumor detection, highlighting their effectiveness in automating the diagnostic process, Aryal et al. [13]. The study by Asif et al. enhances the performance of various deep transfer learning models for brain tumor detection from MRI images, showcasing their improved accuracy in medical image classification [14]. Asiri et al. [15] studied the application of the Swin Transformer for brain tumor detection, demonstrating its potential for achieving high classification accuracy and improved performance in medical imaging. The study IN [16] by Gayathri and Kumar utilizes a CNN pre-trained VGG-16 model for brain tumor segmentation and classification in MRI images, demonstrating its effectiveness in improving detection accuracy. In [17] the paper by Khatib Sulaiman Dalam No et al. (n.d.) provides a review of CNN-based methods for the segmentation and detection of brain tumors in MRI images, highlighting the potential and challenges of using deep learning in medical imaging. In [18], the study by Mallampati et al. combines 3D-UNet segmentation features with a hybrid machine learning model for brain tumor detection, enhancing diagnostic accuracy and performance in medical imaging. In [19], the paper by Meshram et al. (n.d.) explores the use of Swin Transformers for brain tumor detection, emphasizing their potential in enhancing the accuracy and ef-

ficiency of medical image analysis. Musallam et al. [20] introduces a novel Convolutional Neural Network architecture for the automatic detection of brain tumors in MRI images, demonstrating improved accuracy and efficiency in tumor identification. In [21] the study by Neamah et al. leverages an improved ResNet50 architecture for brain tumor classification based on MRI images, showcasing its enhanced performance in medical image analysis. The study in [22] by Prakash Vangala (n.d.) focuses on brain tumor detection in MRI images using a modified CNN and RESNETV50 with transfer learning, enhancing detection accuracy and model efficiency. (Vangala, n.d.). In [23] the paper by Rajendran et al. presents an automated deep learning-based approach for the segmentation of brain tumor MRI images, demonstrating significant improvements in segmentation accuracy. [24], The paper by Rizwan et al. proposes a Gaussian Convolutional Neural Network for the classification of brain tumors and glioma grades, achieving improved classification accuracy in MRI image analysis. It is studied in [25] by Wankhede et al. explores brain tumor detection and classification using an ensemble of deep learning models, including Adjusted InceptionV3, AlexNet, VGG16, VGG19, and ResNet50-152, demonstrating their effectiveness in improving tumor classification accuracy. It is studied in [26] by Zahoor et al. introduces a novel deep residual and regional CNN for brain tumor MRI classification, showcasing its effectiveness in improving classification performance for medical imaging tasks. It is studied in [27] by Saeedi et al. the exploration of MRI-based brain tumor detection using convolutional deep learning methods combined with selected machine learning techniques, highlighting their potential for accurate tumor detection in medical imaging. It is studied in [28] by Xie et al. (2021) the introduction of SegFormer, a simple and efficient transformer-based model for semantic segmentation, offering improved performance and scalability for image segmentation tasks.

Our Proposed Work

In this section, we provide a detailed explanation of the dataset used and our proposed work for brain tumor detection. We introduce a systematic approach for developing and evaluating the MRI brain tumor detection model. We have developed our architecture based on both the transformer based model; Segformer as well as hybrid CNN model. By using this approach, we have combined the strengths of both models to get a comprehensive understanding of tumor structure regardless of its shape, size, intensity or location in brain. This model is designed to effectively capture both local features including edges and textures as well as global contextual information. It will help in segmenting the brain tumors precisely with high accuracy. A brief description of technique for brain tumor detection is shown in Fig. 1. The workflow starts from data collection where MRI scan images and their corresponding masks are gathered. Before model training, the data is preprocessed by resizing and normalizing the images. After preprocessing of data, the next step is segmentation to predict the tumor regions. The data is segmented using a hybrid CNN and Segformer models. Finally, the results are obtained by generating predicted tumor masks from the MRI images and are compared to the original masks to assess the segmentation performance.

A. Dataset

We utilized the LGG Segmentation dataset from publicly available online data from kaggle.com for brain tumor detection. In this dataset, the images were obtained from TCIA (The Cancer Imaging Archive). Both the MRI images and tumor segmentation masks were gathered for patients with brain tumor. The data is of a specific type of brain tumor i.e. lower grade glioma (LGG). The brain MRI images are available in .tif format. Each image consists of 3 channels that are corresponding to different MRI sequences like pre-contrast, post-contrast and FLAIR (Fluid Attenuated Inversion Recovery). There are a total of 110 patient cases. Out of these cases, 101 cases contain all three MRI sequences but for remaining cases, one or two MRI sequences are missing. For 9 cases, the post-contrast sequence is missing and for 6 cases, the pre-contrast sequence is missing. In such cases, the missing sequences are replaced by the FLAIR sequence to maintain consistency of having a uniform three-channel format for all MRI images. For each MRI image of patient, a binary segmentation mask is included in the dataset. These masks indicate the tumor region in the brain and are single-channel images where the tumor region (FLAIR abnormalities) is marked with a value of 1 and the non-tumor regions (background) are marked with 0. For the data of all patients, the dataset is organized into 110 folders. Each folder contains images and masks for that particular patient. All the patients have different number of MRI slices according to the number of slices captured during the MRI procedure. The dataset also contains a data.csv file which provides the information about the tumor genomic clusters for each patient and patient details including age, gender and other demographic data. There are a total of 3929 images available in the dataset including both images and segmentation masks which are labelled and have high quality. So, this dataset is suitable for our training the deep learning model to segment the brain tumors effectively.

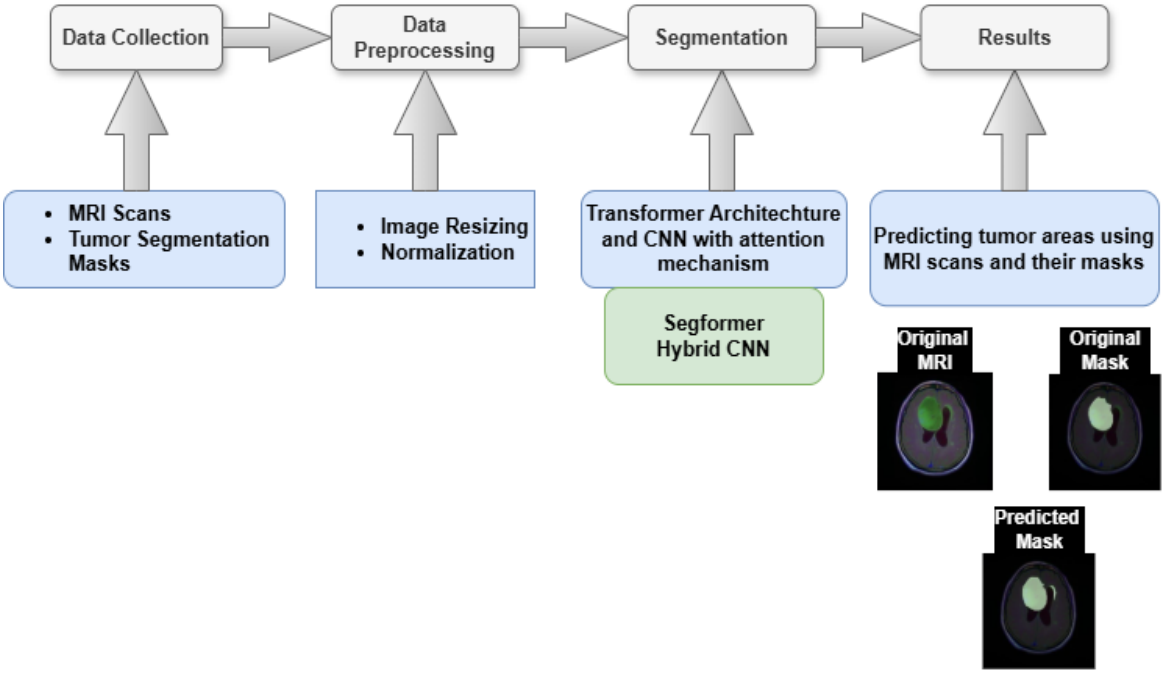


Figure 1: Proposed Methodology

B. Hardware Specification

Table 1 shows the hardware specification for our model development.

Table 1. Hardware Description

Hardware	Description
GPU	1xTesla K80, compute 3.7, having 2496 CUDA cores, 12GB GDDR5 VRAM
CPU	1xsingle core hyper threaded Xeon Processors @2.0Ghz i.e. (1 core, 2 threads)
RAM	11.6 GB Available
Disc	33 GB Available

C. Data Preprocessing

Preprocessing the dataset is an essential step to make it usable for model training. The MRI images are likely to contain noise, inconsistencies and different formats that makes the models unable to learn effectively. So, data preprocessing impacts the performance of model through techniques like normalization, resizing and augmentation. This improves the quality of data so that model is able to generalize well across unseen data. For our study, we implemented the preprocessing of data using the python's Albumentations library which is designed to perform image augmentation. We prepared the data of MRI images and segmentation masks by the following preprocessing steps:

1) Data Resizing and Normalization:

We processed the images and their corresponding segmentation masks through image resizing and normalization. The deep learning models expect a uniform input size for training so that it can be compatible with the model's architecture. The MRI images were different in sizes and resizing them to a standard size was necessary to make sure that the model learn features consistently without being influenced by varying image sizes. For this purpose, we resized the images and masks to a fixed size of 224 x 224 pixels.

In the next step of preprocessing, we normalized the data to adjust the pixel values of the images according to a standard scale. This technique helped the faster training of model with less epochs to converge as the model was able to focus on the relevant patterns in the data. Many models are pretrained on ImageNet. They expect the input data to follow the same normalization pattern for effective transfer learning. We normalized the MRI images to let them have a mean of [0.485, 0.456, 0.406] and a standard deviation of [0.229, 0.224, 0.225]. These are commonly used values from the ImageNet dataset. These values standardize the input images to match the data distribution that many models pretrained on ImageNet expect. Then we normalized the segmentation masks by converting them to binary values (0 and 1). The masks represented the tumor regions as 1 and the non-tumor as 0. We performed the thresholding operation by setting the pixel values greater than a threshold i.e. 0 to 1 and others to 0. This made the masks in a format suitable for segmentation task.

The data was prepared for model training after preprocessing but for visualizing the predictions of model, we converted the images back to their original pixel range. This in-

verse transformation was applied using the invTrans pipeline that undoes the normalization applied earlier and the image was restored to its original color range (0 to 255).

2) Data Augmentation:

To improve the models' ability to generalize well, we expanded the size of our dataset by applying the data augmentation technique to our segmentation dataset. This step is important in medical imaging and it helps the model to learn diverse patterns from large dataset, prevent overfitting and improve its generalization when applied to unseen data. We applied these transformations to our dataset using Albumentations library:

a) Grayscale Transformation

We converted some of the images in our dataset to grayscale by applying the grayscale transformation. We used the probability of 10% so that the color information of most of the images does not change. In this way, we benefitted from grayscale augmentation. In MRI tumor detection, the structural and spatial features are more important than the color. So, we removed the color information occasionally through this transformation which helps the model to focus more on essential details such as tumor's shape or the contrast between tumor and surrounding tissue. This increased the diversity of data and improved the generalization of model.

b) Geometric Transformations

To prepare the data for better performance of model, we also applied multiple geometric transformations randomly to our LGG dataset. The reason behind this was that the model can accurately detect tumors having different orientation or position in the images. This increased the ability of the model to generalize well for different orientations of the brain tumor in MRI images. We used the horizontal and vertical flipping transformation to randomly flip the images either horizontally or vertically. This replicates the scenario where the patient's position might be changed during MRI scan. We rotated the images randomly by 90 degrees using the RandomRotate90 transformation to mimic situations where MRI scan captures the image of brain from slightly rotated angles. We also transposed some images to swap their axes and flip them along the diagonal. We used the probability of 50% in all these transformations so that there is variability in the dataset which will improve the model's segmentation for unseen data.

3) Data Splitting:

We split the dataset into training, validation, and testing subsets using the train_test_split function from scikit-learn. The dataset contained 3929 images and their masks for brain tumor. Approximately 86.7% of data i.e. 3429 samples were used for training. The remaining around 14% of data was used for testing and validation. Both of these sets contained 250 samples each.

D. Model Architectures

1) Hybrid CNN

The hybrid CNN model combines convolutional layers with transformer blocks and channel attention mechanism. Our architecture consists of convolutional layers for initial feature extraction, channel attention to amplify relevant features, transformer blocks to model global dependencies as well as sampling layers for obtaining segmentation maps in the original resolution. The Fig. 3 shows the hybrid CNN with attention mechanism. The CNN blocks act as feature extractors, identifying local patterns in the input images. Each block consists of convolutional layers followed by channel attention, which focuses on the most relevant features in each channel. After feature extraction, MaxPooling reduces the spatial dimensions, making the model more efficient. It employs a Multi-Head Self-Attention mechanism that enables the model to capture long-range dependencies. This facilitates to carry on identifying complex patterns.

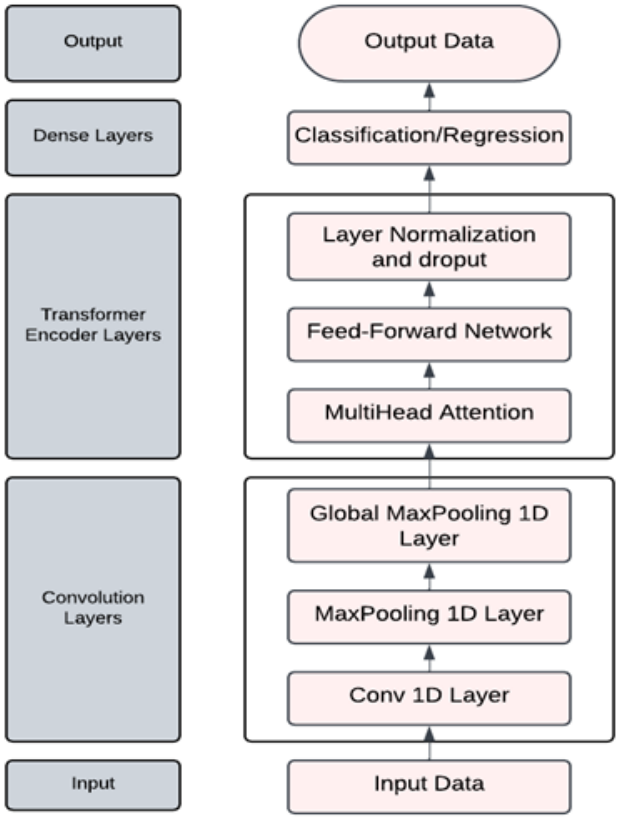


Figure 3: Hybrid CNN with attention Mechanism

The next stage is a Transformer block which captures global context to be able to focus on areas of an image that are relevant after the CNN has capture local features. This code implements Multi-Head Attention as seen in the standard Transformer architecture. This enables it to capture relationships among pixels globally, meaning that the model can attend to multiple regions at once. Its output is shown in Fig. 4.

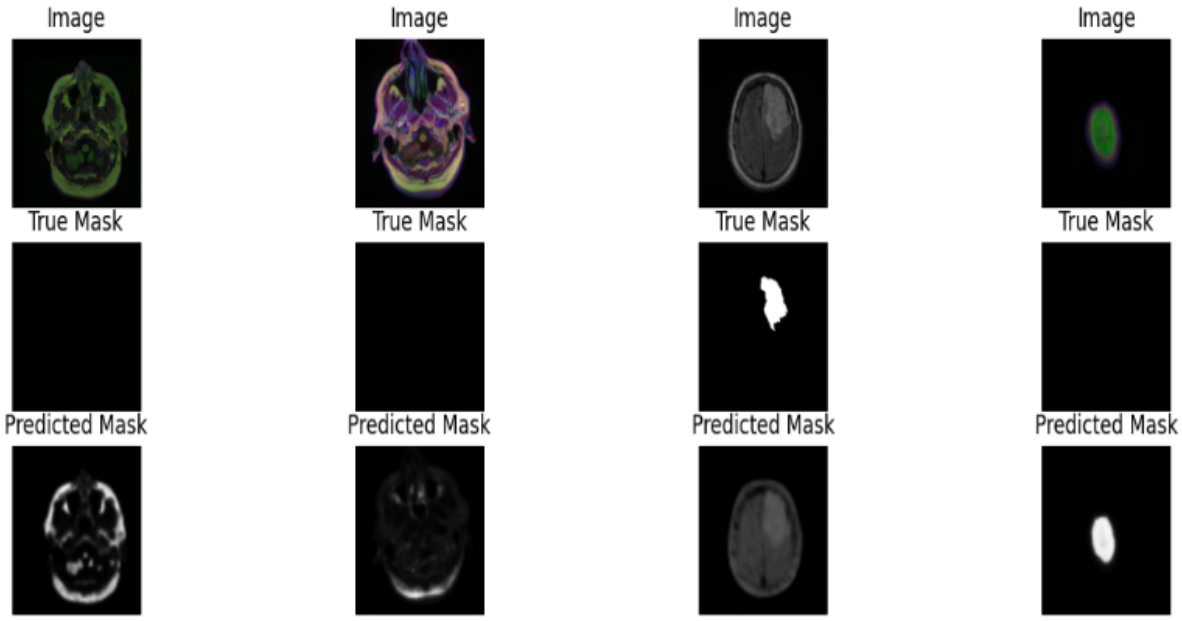


Figure 4: Output Predicted by Hybrid CNN

2) SegFormer

In this paper, we implemented SegFormer for semantic segmentation task. We used encoder with multi-head attention mechanism to extract long-range dependencies from input images. Through overlap patch merging, input images are split into overlapping patches that are processed by convolutional layers to extract multi-scale feature representations. We applied the LayerNorm2d to normalize these features. To enhance the extracted features using depth-wise convolutions and GELU activations, we used a fully convolutional feed-forward network i.e. the MixMLP module. We obtained multi-scale features at different resolutions from the output of encoder. We used the decoder to process the multi-scale features from the encoder to generate high-resolution segmentation masks. It generated the intermediate feature maps that were passed to the segmentation head for final processing. The final binary segmentation output was produced by segmentation head that combines features from all decoder stages. The 1×1 convolution layer reduced the feature maps to a single segmentation map where tumor regions are marked as 1 and the background as 0 for clear identification. We configured the SegFormer model with these parameters:

- Input channels: 3 (pre-contrast, FLAIR and post-contrast MRI sequences)
- Encoder Widths: [64, 128, 256, 512] to handle increasing feature complexity
- Transformer Depths: [3, 4, 6, 3] layers for hierarchical feature extraction
- Number of Attention Heads: [1, 2, 4, 8] (increasing with the feature width)
- Patch Sizes and Overlaps: [7, 3, 3, 3] for patch merging
- Decoder Channels: 256 for merging multi-scale features
- Segmentation Output: A single binary mask

Fig. 5 shows the original SegFormer framework which was described by Xie et al. shows the integration of hierarchical transformers with lightweight decoders to achieve accurate segmentation.

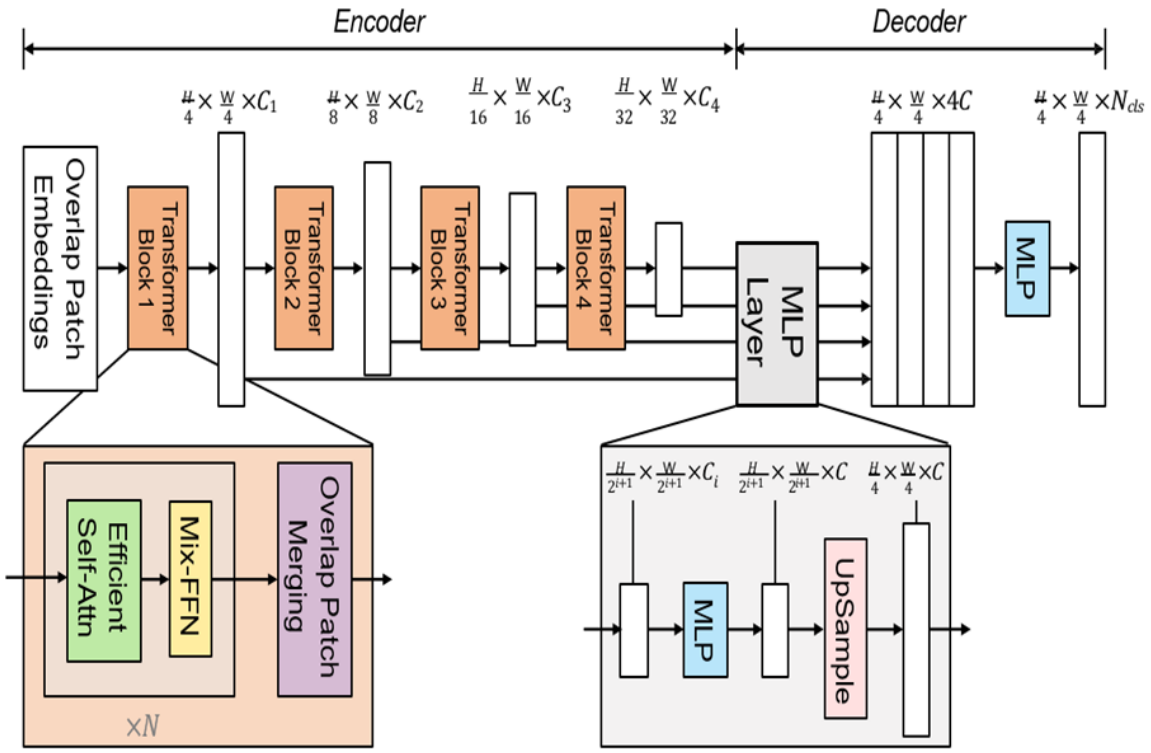


Figure 5: SegFormer Architecture [28]

Few of the outputs predicted by our developed SegFormer model are shown in Fig. 6.

E. Model Training

To achieve accurate segmentation results while minimizing overfitting and unnecessary computational effort, we used these strategies:

1) Loss Function:

We used the binary cross entropy with Logits loss as the loss function along with a sigmoid activation for the segmentation of brain tumors. The BCE loss computed the difference between the predicted logits and the ground truth masks that quantified how well the model identifies tumor regions and background.

2) Optimization

We trained the model using AdamW optimizer which combined adaptive gradient methods with weight decay regularization to prevent overfitting. To further fine tune the training process, we employed a ReduceLROnPlateau scheduler that reduced the learning rate by a factor of 0.1 if the validation loss did not improve for 5 consecutive epochs.

3) Early Stopping

To prevent overfitting to the training data, we implemented early stopping with a patience parameter of 8 epochs. The validation loss was monitored and training was terminated if there was no improvement for the specified patience period. In case of improvement, it acted as a checkpoint for preserving the best-performing model.

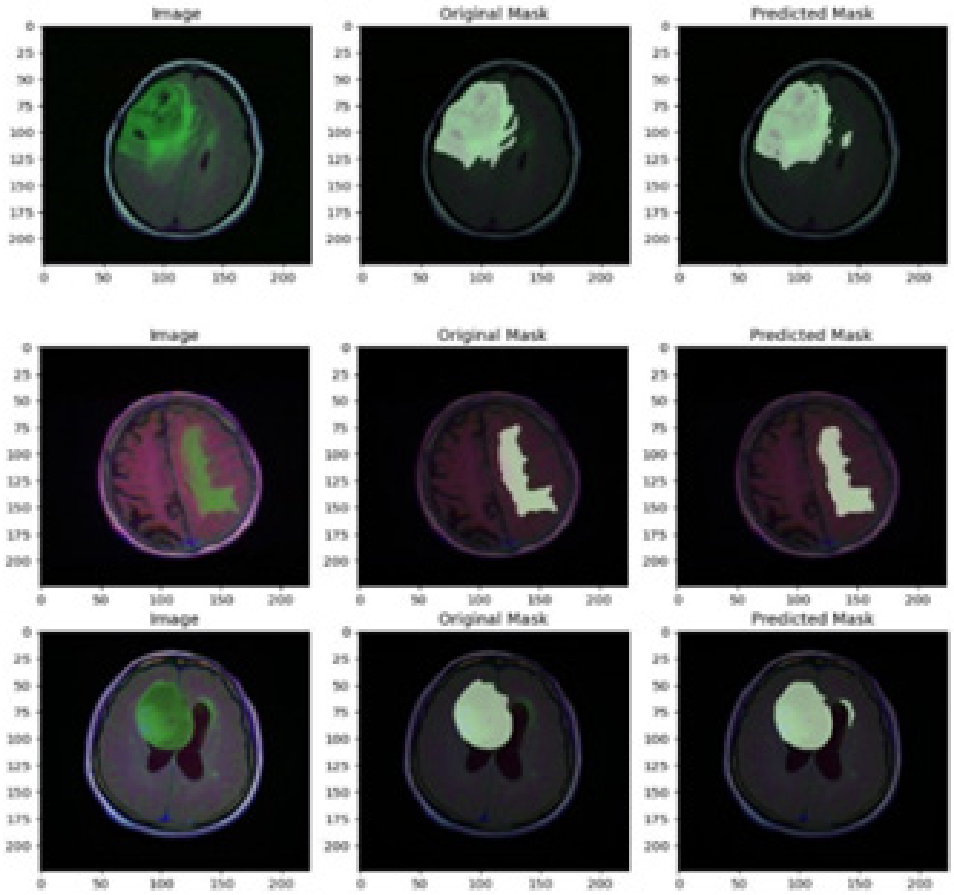


Figure 6: Output Predicted by SegFormer

4. Results

A. Performance Evaluation

Both models were evaluated based on accuracy, the Dice coefficient, sensitivity, and specificity. The comparison strategy involved using the results from the Hybrid CNN-Transformer model as a baseline to underscore the SegFormer's enhanced performance. This approach highlighted the SegFormer's ability to integrate global and local context, yielding superior segmentation accuracy.

1) Accuracy

This tells us how many times our model has predicted the correctly. It is calculated by dividing correct predictions with total number of predictions.

$$Accuracy = \frac{(True\ Positive + True\ Negative)}{(True\ Positive + True\ Negative + False\ Positive\ +\ False\ Negative)} \quad (1)$$

2) Dice Coefficient

Dice Coefficient measure similarity between two sets A and B. The coefficient ranges from 0 to 1. 0 indicates that there is not overlapping between two sets while 1 indicates that two sets are identical.

$$(A, B) = \frac{2(A \cap B)}{(A + B)} \quad (2)$$

3) Sensitivity

Sensitivity tells us how many positive values our model can identify correctly. Sensitivity or SN can be calculated by the number of correct positive predictions divided by the total number of positives.

4) Specificity

This measure tell us how many negative values our model can identify correctly. This measure is also called True Negative Rate (TNR). It is calculated by dividing number of correct negative predictions with total number of negative predictions.

Considering the above performance measures, we calculated the performance of our proposed models i.e. both the segformer and hybrid CNN with attention mechanism to provide a brief analysis on which model performed better and generalizes better according to the input data.

Table 2. Comparative Analysis of SegFormer and Hybrid CNN

Performance Metric	SegFormer Model	Hybrid-CNN Transformer
Accuracy	99%	99%
Dice Coefficient	0.92	0.89
Sensitivity	0.91	0.88
Specificity	0.93	0.90
Computational Efficiency	Faster	Slower
Epoch Duration	20 seconds/epoch	20 seconds/epoch
Segmentation Precision	High	Moderate
Model Robustness	Better Generalization	Moderate Generalization

The Curves in Fig 7. and Fig. 8 provide the training and validation accuracy and loss curves for both SegFormer and Hybrid CNN.

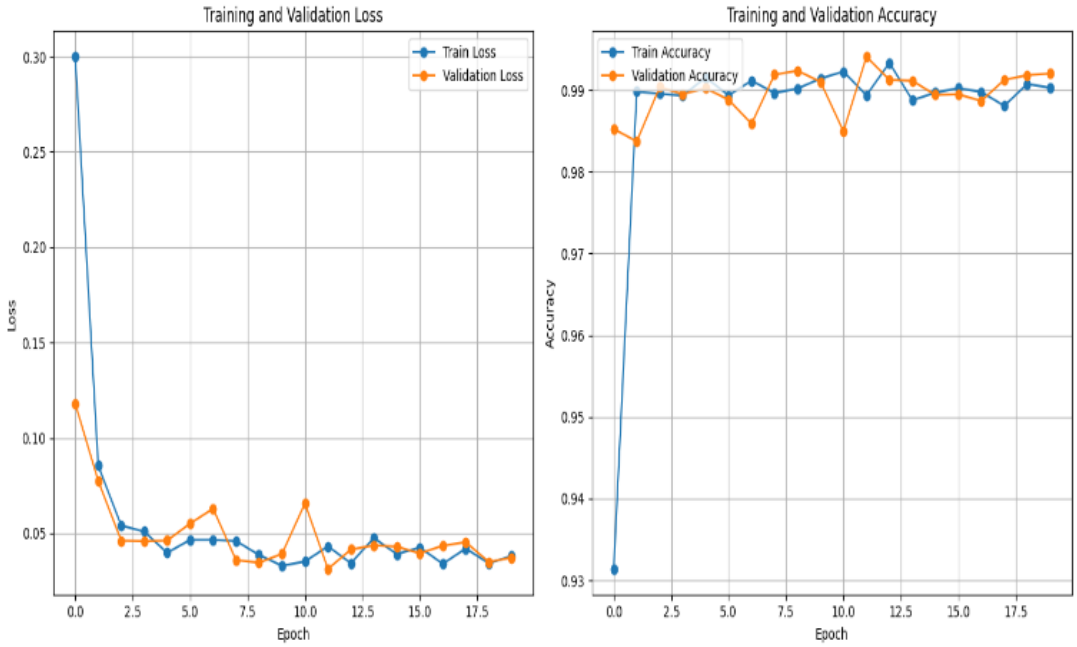


Figure 7: Training and Validation Curves (Hybrid-CNN)

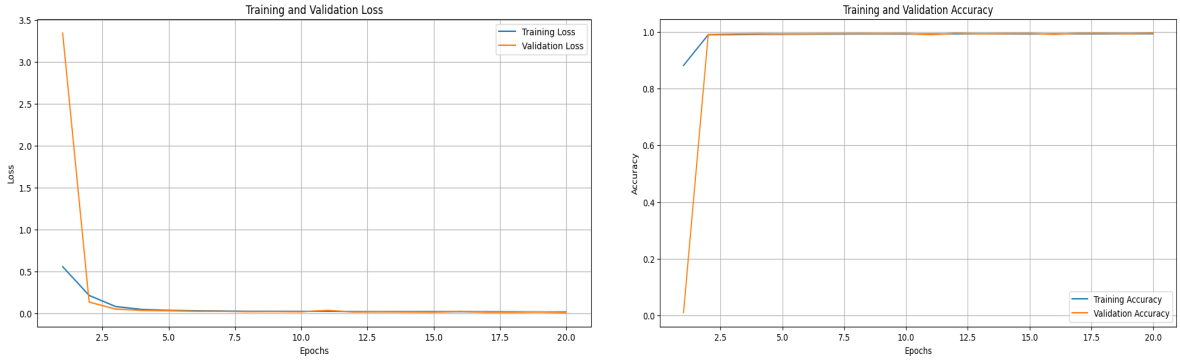


Figure 8: Loss and Accuracy Curves (Segformer)

Discussions

The comparison between the SegFormer and Hybrid CNN-Transformer models for segmentation has presented clearly various behaviors with respect to performance, even though we also confirmed that the Hybrid CNN-Transformer model is a competing method in brain tumor segmentation tasks alongside SegFormer.

• *Accuracy and Efficiency:* In terms of accuracy and efficiency the Segformer model and hybrid CNN model have accuracy of 99% each but hybrid CNN model is not able to perform well on test dataset hence Segformer outperforms with an impressive result to be able to perform well on test data. The architecture of SegFormer is capable of integrating global context with fine-grain features that helped it to achieve high accuracy in the detection of tumor region.

• *Segmentation Performance:* SegFormer was superior to the Hybrid model according to its Dice coefficient which is a common metric for segmentation quality. It even identi-

fied the tumor margins more accurately which are important for diagnosis and treatment planning.

• *Visual Results:* Both models are visualized across image-based MRI scans, refrained from original mask and predicted on the mask (as in block diagram). This visualization comparison illustrates the improved close matching of SegFormer output and ground truth segmentation masks.

• *Computational Efficiency:* When comparing the computational cost, the Segformer was able to train faster (shorter time per epoch) compared to other transformer architectures due to its efficient transformer-based backbone and would be a better candidate for real-time medical imaging applications.

Future work could be done on introducing the variety in dataset by collecting MRI scans of patients from different regions to increase the generalizability.

Conclusion

In this study, we proposed a SegFormer transformer model as an alternative solution for brain tumor segmentation using MRI images. We developed the SegFormer model after preprocessing the LGG segmentation dataset from Kaggle and training the model using optimization strategies. We did a comparative analysis of SegFomer's performance with the hybrid CNN. Experimental findings reveal that the SegFormer and hybrid CNN both have a relatively high accuracy of 99%, but better segmentation precision is achieved by SegFormer. It outperformed the hybrid CNN model in metrics including accuracy, dice coefficient, sensitivity and specificity. The SegFormer effectively utilized self-attention mechanism that helps to capture long-range dependencies and detailed spatial features of MRI images. So, it proved to be a valuable tool for medical imaging. However, the study also has some limitations and need further research.The performance of SegFormer can further be enhanced by tuning the hyperparameters which has not been done in this study and other techniques like GANs can be used to train the model on a larger and broader dataset as this study doesn't includes very large dataset.

In future work, the segmentation accuracy can be improved further by using more diverse and larger datasets. For validating its performance across broader applications, the other type of brain tumors should be included in the dataset. By addressing these challenges, SegFormer can be successfully integrated in health sectors for accurate brain tumor segmentation and treatment planning.

Declaration of Interest: The authors declare that there is no conflict of interest.

Author Contributions: All authors contributed to the work equally.

REFERENCES

[1] Jahan, R., & Tripathi, M. M. (2024). Brain Tumor Detection using Hybrid CNN in fMRI images. Multidisciplinary Science Journal, 6, 2024ss0611. https://doi.org/10.31893/multiscience.2024ss0611

[2] Srinivasan, S., Francis, D., Mathivanan, S.K., et al. (2024). A hybrid deep CNN model for brain tumor image multi-classification. BMC Medical Imaging, 24(21). https://doi.org/10.1186/s12880-024-01195-7

[3] Santhosh Kumar, S., Sasirekha, S. P., & Santhosh, R. (2024). Parallel-way: Multi-modality-based brain tumor segmentation using parallel capsule network. Electromagnetic Biology and Medicine, 43(1), 1-25. https://doi.org/10.1080/15368378.2024.2390058

[4] Patro, S. G. K., Govil, N., Saxena, S., Mishra, B. K., Zamani, A. T., Ben Miled, A., Parveen, N., Elshafie, H., & Hamdan, M. (2024). Brain Tumor Classification Using an Ensemble of Deep Learning Techniques. IEEE Access, 12, 162094-162106. https://doi.org/10.1109/ACCESS.2024.012106

[5] Grampurohit, S., Shalavadi, V., Dhotargavi, V. R., Kudari, M., & Jolad, S. (2020). Brain Tumor Detection Using Deep Learning Models. 2020 IEEE India Council International Subsections Conference (INDISCON), Visakhapatnam, India, pp. 129-134. https://doi.org/10.1109/INDISCON50162.2020.00037

[6] Hassan, S., et al. (2022). Comparative Analysis of Machine Learning Algorithms in Detection of Brain Tumor. 2022 3rd International Conference on Big Data Analytics and Practices (IBDAP), Bangkok, Thailand, pp. 31-36. https://doi.org/10.1109/IBDAP55587.2022.9907433

[7] Nagaraju, G., Nath, R. K., Chinniah, P., Balasubramanian, K., Kirubakaran, S., et al. (2024). Machine Learning Techniques for Brain Tumor Classification. Journal of Electrical Systems, 20(2s), 901-909. https://doi.org/10.1016/j.matpr.2021.03.723

[8] Bhagwan, J., Rani, S., Kumar, S., Chaba, Y., & Godara, S. (2024). Advancements in Brain Tumor Detection Using Deep Learning. Journal of Electrical Systems, 20(3s), 1072-1081. https://doi.org/10.31436/iiumej.v25i2.2963

[9] Gokila Brindha, P., et al. (2021). A Review on Deep Learning Techniques for Brain Tumor Detection. IOP Conference Series: Materials Science and Engineering, 1055, 012115. https://doi.org/10.1088/1757-899X/1055/1/012115

[10] Soomro, T. A., et al. (2023). Image Segmentation for MR Brain Tumor Detection Using Machine Learning: A Review. IEEE Reviews in Biomedical Engineering, 16, 70-90. https://doi.org/10.1109/RBME.2022.3185292

[11] Liu, Y., Ma, Y., Zhu, Z., Cheng, J., & Chen, X. (2024). TransSea: Hybrid CNN-Transformer with Semantic Awareness for 3D Brain Tumor Segmentation. IEEE Transactions on Instrumentation and Measurement.

[12] Ahmad, S., & Choudhury, P. K. (2022). On the Performance of Deep Transfer Learning Networks for Brain Tumor Detection Using MR Images. IEEE Access, 10, 59099–59114. https://doi.org/10.1109/ACCESS.2022.3179376

[13] Aryal, S., Sharma, S., Sedai, S., Aryal, P., & Mansur, J. (2023). Brain Tumor Detection Using Convolutional Neural Networks (CNNs). Kathford Journal of Engineering and Management, 3(1), 134–148. https://doi.org/10.3126/kjem.v3i1.62888

[14] Asif, S., Yi, W., Ain, Q. U., Hou, J., Yi, T., & Si, J. (2022). Improving Effectiveness of Different Deep Transfer Learning-Based Models for Detecting Brain Tumors From MR Images. IEEE Access, 10, 34716–34730. https://doi.org/10.1109/ACCESS.2022.3153306

[15] Asiri, A. A., Shaf, A., Ali, T., Pasha, M. A., Khan, A., Irfan, M., Alqahtani, S., Alghamdi, A., Alghamdi, A. H., Alshamrani, A. F. A., Alelyani, M., & Alamri, S. (2024). Advancing brain tumor detection: harnessing the Swin Transformer's power for accurate classification and performance analysis. PeerJ Computer Science, 10. https://doi.org/10.7717/peerj-cs.1867

[16] Gayathri, T., & Kumar, K. S. (2024). BRAIN TUMOR SEGMENTATION AND CLASSIFICATION USING CNN PRE-TRAINED VGG-16 MODEL IN MRI IMAGES. IIUM Engineering Journal, 25(2), 196–211. https://doi.org/10.31436/iiumej.v25i2.2963

[17] Khatib Sulaiman Dalam No, J., Sadeeq Salih, N., & Mohsin Abdulazeez, A. (n.d.). CNN-Based Segmentation and Detection of Brain Tumors MRI Images: A Review. Indonesian Journal of Computer Science.

[18] Mallampati, B., Ishaq, A., Rustam, F., Kuthala, V., Alfarhood, S., & Ashraf, I. (2023). Brain Tumor Detection Using 3D-UNet Segmentation Features and Hybrid Machine Learning Model. IEEE Access, 11, 135020–135034. https://doi.org/10.1109/ACCESS.2023.3337363

[19] Meshram, P. A., Joshi, S. S., & Mahajan, D. A. (n.d.). Brain Tumor Detection using Swin Transformers. http://www.cancerimagingarchive.net/.

[20] Musallam, A. S., Sherif, A. S., & Hussein, M. K. (2022). A New Convolutional Neural Network Architecture for Automatic Detection of Brain Tumors in Magnetic Resonance Imaging Images. IEEE Access, 10, 2775–2782. https://doi.org/10.1109/ACCESS.2022.3140289

[21] Neamah, K., Mohamed, F., Waheed, S. R., Kurdi, W. H. M., Yaseen, A., & Kadhim, K. A. (2024). Utilizing Deep improved ResNet50 for Brain Tumor Classification Based MRI. IEEE Open Journal of the Computer Society. https://doi.org/10.1109/OJCS.2024.3453924

[22] Prakash Vangala, B. (n.d.). Brain Tumor Detection in MRI Images with Modified CNN and RESNETV50 using Transfer Learning BRAIN TUMOR DETECTION ON MRI IMAGES. https://doi.org/10.13140/RG.2.2.14168.51208

[23] Rajendran, S., Rajagopal, S. K., Thanarajan, T., Shankar, K., Kumar, S., Alsubaie, N. M., Ishak, M. K., & Mostafa, S. M. (2023a). Automated Segmentation of Brain Tumor MRI Images Using Deep Learning. IEEE Access, 11, 64758–64768. https://doi.org/10.1109/ACCESS.2023.3288017

[24] Rizwan, M., Shabbir, A., Javed, A. R., Shabbir, M., Baker, T., & Al-Jumeily Obe, D. (2022). Brain Tumor and Glioma Grade Classification Using Gaussian Convolutional Neural Network. IEEE Access, 10, 29731–29740. https://doi.org/10.1109/ACCESS.2022.3153108

[25] Sushant Wankhede, D., Shelke, C. J., Shrivastava, V. K., Achary, R., & Mohanty, S. N. (2024). EAI Endorsed Transactions Brain Tumor Detection and Classification Using Adjusted InceptionV3, AlexNet, VGG16, VGG19 with ResNet50-152 CNN Model. In EAI Endorsed Transactions on Pervasive Health and Technology (Vol. 10).

[26] Zaitoon, R., & Syed, H. (2023). RU-Net2+: A Deep Learning Algorithm for Accurate Brain Tumor Segmentation and Survival Rate Prediction. IEEE Access, 11, 118105–118123. https://doi.org/10.1109/ACCESS.2023.3325294

[27] Saeedi, S., Rezayi, S., Keshavarz, H., & R. Niakan Kalhori, S. (2023). MRI-based brain tumor detection using convolutional deep learning methods and chosen machine learning techniques. BMC Medical Informatics and Decision Making, 23(1). https://doi.org/10.1186/s12911-023-02114-6

[28] Xie, E., Wang, W., Yu, Z., Anandkumar, A., Alvarez, J. M., & Luo, P. (2021). SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers. http://arxiv.org/abs/2105.15203

COMPARATIVE ANALYSIS OF LARGE LANGUAGE MODELS

Buse Cansu Yiğit
Isparta University of Applied Sciences, Graduate Education Institute, Computer Engineering, Isparta, Türkiye

Ahmet Ali Süzen
Isparta University of Applied Sciences, Faculty of Technology, Computer Engineering, Isparta, Türkiye

Remzi Gürfidan*
Isparta University of Applied Sciences, Yalvaç Vocational School of Technical Sciences, Computer Programming, Isparta, Türkiye
(Corresponding author) remzigurfidan@isparta.edu.tr

ABSTRACT

Large language models have made great progress in the field of natural language processing in recent years. Today, many models have been developed in the field of artificial intelligence and these models address different purposes. Some of these uses include language processing, text generation, translation, question-answering, summarization, coding and many more. This paper presents a comparative analysis of today's widely used large language models (DeepSeek, Gemini, GPT-4, Cursor, Claude, Llama, Mistral, GPT-4, Cursor, Claude, Llama, Mistral), examining their effectiveness on different tasks, the size and quality of training data, privacy and security issues, multilingual capabilities, and explainability. The findings of the study reveal the strengths and weaknesses of each model and provide recommendations for future research. This study aims to provide guidance in the development and implementation of LLMs.

Keywords: Large language models, natural language processing, performance analysis, security, explainability, multilingual models.

Introduction

Large language models are deep learning models pre-trained using machine learning techniques and various algorithms on extensive datasets [1]. These models have demonstrated significant technological advancement, substantially simplifying users' daily activities. Consequently, the question "Will artificial intelligence replace human lives?" has emerged as a topic of discussion. This development yields positive outcomes in certain contexts while simultaneously generating concern among various individuals. Numerous large language models employed across diverse domains have emerged in contemporary times, with their numbers continuously increasing. Initially, these models were primarily utilized for tasks such as text generation, translation, summarization, and question-answering; however, their ap-

plications have progressively diversified to include code generation, image creation, visual interpretation, and numerous additional functions.

The accuracy of large language models correlates directly with the magnitude and quality of the training dataset. Nevertheless, increased data volume may simultaneously produce adverse effects, complicating data management and introducing security vulnerabilities (Yao et al., 2024). Additionally, the enhancement of multilingual capabilities will facilitate the global proliferation of large language models, enabling them to address a broader user demographic. The capacity to meaningfully translate content from one language to another constitutes a significant indicator of these models' capabilities [2].

When examining large language models from a comprehensibility perspective, the degree of accuracy demonstrated by the model represents a critical concern for users. In this context, explainable artificial intelligence approaches offer substantial solutions for enhancing model transparency [1]. Recently, researchers have addressed security and privacy risks associated with large language models. The potential leakage of data utilized during the training process poses a significant risk regarding the protection of user information [3].

This study had been analyze and comparatively evaluate large language models including DeepSeek, Gemini, GPT-4, Cursor, Claude, LLama, and Mistral, identifying their respective strengths and limitations. The comparison examines characteristics such as training data volume, privacy and security measures, multilingual capabilities, and explainability features. Through this comparative analysis, the research aims to provide guidance to users and researchers by elucidating the strengths and weaknesses of each model.

Methodology

In this study, the size and quality of the training dataset, privacy and security, multilingual capabilities and explainability are analyzed comparatively.

Model Selection

DeepSeek, Gemini, GPT-4, Cursor, Cursor, Claude, Llama and Mistral are the most widely used big language models today. Although these models have been designed by different developers, each of them exhibits different performance characteristics in various application domains due to their unique architectures, training methods and optimization techniques, which necessitates a comprehensive benchmarking study so that users can choose the most suitable model for their needs. Some of these application domains include text-based question and answer systems, multilingual translation applications, summarization, coding, text generation, and visual content generation. These tasks and the language processing capabilities of each model are discussed and analyzed in a comparative manner.

Performance Assessment

Different natural language processing methods were used to evaluate the performance of the models. Within the scope of these tests, different usage areas were considered. The performance of each model was evaluated by considering features such as accu-

racy values, model response time, semantic accuracy. In this way, comparable data was obtained not only in terms of technical aspects but also in terms of user experience and practical performance in real-world applications. For example, it was observed that the slow response time of a model, despite its high accuracy, may limit its usability in real-time applications. Finally, these evaluations are compared and tabulated, and the strengths and weaknesses of the models are discussed.

Size and Quality of Training Data

The quality of models is usually directly proportional to the size of the datasets used, as more data improves the learning process of the model and leads to more consistent results. However, as the datasets grow larger, different negative situations may arise. For example, problems such as increased noise in the data sets, increased labeling errors or unbalanced distributions can negatively affect the performance of the model [4]. It is very important to be able to balance this. For this reason, the training data set size, diversity and quality of each model are considered separately. Cleaning the datasets, removing unnecessary or inconsistent data, completing missing values with appropriate methods and keeping the data up-to-date are among the critical factors that directly affect the accuracy and reliability of the model. Considering all these factors, the performances of different models were analyzed and compared in detail. Based on the results obtained, it is interpreted which model is more suitable for a specific use case and these interpretations are presented to guide users in the model selection process.

Privacy and Security

The protection of user data and the risks of misuse of the model are as important as the function of the model. The ethical and secure use of such technologies directly affects not only the privacy of individuals but also public trust. Therefore, security measures and data privacy practices are evaluated and continuously improved from the design phase of each model. In particular, how the models process user data, what data protection protocols they implement and what measures they take against possible data leaks are issues that need to be analyzed in detail. In addition, the potential for artificial intelligence systems to generate false or misleading information should not be ignored, as this can have negative consequences not only for individuals, but also for organizations and even society at large. As a result, it is necessary to be aware of the risks that technology brings with it as well as the opportunities it offers and to work to minimize these risks [5].

Multilingual Capabilities

Multilingual capabilities are very important for the model to reach more users and produce more accurate information in these languages. In this direction, different models are considered and the performance of each large language model is tested according to the answers given in different languages and evaluated in various tasks such as language translation, language content, ability to understand the given text, text generation, and translation [2]. As a result, it is very important that the model not only produces answers but also produces consistent and meaningful answers. At the end of these analyses, the multilingual performances of each model are tabulated and compared, and suggestions

are presented on which model can be preferred by identifying its weaknesses and advantages.

Explainability and Transparency

The explainability of models is crucial for users to be able to understand the model's response and evaluate the accuracy of these outputs [6], because trusting a model's outputs without understanding how it makes decisions can pose great risks for users. Therefore, this study analyzes the explainability and transparency properties of various models by using explainable artificial intelligence (XAI) techniques to analyze how understandable and traceable their decision-making processes are [7]. Through these analyses, users are able to see more clearly how the model arrives at certain conclusions, which helps them come to a more robust conclusion about the accuracy and reliability of the model. Therefore, this approach is not only a technical examination, but also a step towards increasing users' confidence in AI systems, as understanding how a system works is a fundamental condition for using it correctly. Within the scope of all these, research has been conducted and the results are interpreted and presented in tables.

Comparative Analysis

In previous studies, the performance of models has been examined in detail under the headings such as the quality of training data sets, security measures, multilingual capabilities and explainability. Because these factors directly affect the applicability and theoretical soundness of a model. The findings not only provide a comparative analysis of the strengths and weaknesses of each model, but also the results are interpreted and their suitability for different usage scenarios is discussed. For example, while some models may exhibit superior performance in multilingual tasks, others may be limited in terms of explainability.

The main purpose of these comparisons is to objectively assess the potential and limitations of the models. However, these analyses also helped to identify areas for improvement. This is because a model's weakness in a particular area can be instructive for future work. As a result, this study not only finds the current state of the art but also provides recommendations for researchers and users.

Results

In this paper, we compare the performance, training data quality, security metrics, multilingual capabilities and explainability of various large language models. However, since each of these criteria affects the effectiveness of the models in different ways, the analyses are interpreted in a way that indicates what should be used where.

The findings reveal the strengths and weaknesses of each model in quantitative and qualitative terms and are detailed in the tables below. This is because these comparative analyses highlight not only existing performance differences, but also critical areas for improvement. Hence, the data generated during the study is presented in tables to provide the reader with a holistic view of the advantages and disadvantages of the models.

Performance Assessment

In this study, model performance is evaluated using a multidimensional approach to comprehensively measure the core capabilities of large language models. In particular, we focus on basic natural language processing tasks such as text generation, language translation, text summarization, and question answering, as these tasks holistically reflect both the language understanding and generation capabilities of models.

However, considering that performance on different types of tasks may vary, separate assessment protocols were applied for each task. For example, while fluency and coherence criteria are more important in text generation tasks, accuracy and contextual relevance are more important in question answering tasks. Therefore, appropriate metrics were selected for each task type and standardized test data sets were used in the evaluation process.

Furthermore, all models were evaluated under the same test conditions and with similar hardware configurations to increase the reliability of performance comparisons. As a result, the per-test accuracy rates presented in the table below clearly demonstrate the performance differences of the models across tasks.

Table 1: Model Performance Assessment

Model	Text Generation (%)	Language Translation (%)	Summarization (%)	Q&A (%)
GPT-4	95	93	92	94
Claude	90	89	88	91
Gemini	92	91	90	93
DeepSeek	89	90	87	90
Mistral	88	88	86	89
Llama	87	87	85	88
Cursor	91	90	89	92

According to the data presented in Table 1, there are significant differences between the performance of the large language models on different tasks.Overall, GPT-4 stands out as having the highest accuracy rates in all categories, followed by Claude and Gemini, which perform in the 90-91% range. However, even among these top models, there are small differences: for example, GPT-4's 5% advantage in text generation may indicate that it is more proficient in creative language use.

On the other hand, models like DeepSeek and Mistral, while lagging behind in the overall ranking, perform better in certain areas, especially language translation. DeepSeek's 90% in language translation is only 3% lower than GPT-4, suggesting that it could be a strong alternative in this area.

Size and Quality of the Training Data

The size of the training dataset is an important factor affecting the performance of the models. In addition, data diversity, the frequency of data updates and the quality of the data also significantly affect the model performance. The factors examined within the scope of the quality of the data can be expressed as increased noise in the data sets, increased labeling errors or unbalanced distributions [4]. As a result of all these, an evaluation is made and shown in the table below.

Table 2: Size and Quality of Training Data

Model	Training Data Size	Data Variety	Frequency of Updating	Data Quality
GPT-4	500B+	Yüksek	Sürekli	Çok Yüksek
Claude	400B+	Yüksek	Sürekli	Yüksek
Gemini	250B+	Orta	Ara Sıra	Orta
DeepSeek	200B+	Düşük	Düşük	Orta
Mistral	150B+	Düşük	Düşük	Düşük
Llama	220B+	Orta	Ara Sıra	Orta
Cursor	280B+	Yüksek	Sürekli	Yüksek

The data presented in Table 2 reveals critical differences in the training processes of large language models. In general, there is a direct correlation between the size, variety, frequency of updates and quality of the training data and model performance. GPT-4 and Claude perform well on a wide range of tasks, thanks to the fact that they are trained with huge datasets such as 500B+ and 400B+. This confirms the advantage of big data, especially for complex and multidimensional language tasks. Models using relatively smaller datasets, such as Mistral and DeepSeek, are limited in some tasks. However, it also shows that data size alone is not decisive; data quality and diversity are at least as important as data size.

Privacy and Security

Security and privacy measures of large language models are critical for both protecting user data and preventing the misuse of models [5]. These security measures are presented in Table 3, where the performance of the models in terms of data encryption, access control, malicious content filtering, privacy policy, open source codes are analyzed in detail.

Tablo 3: Comparative Analysis of Security Measures

Model	Data Encryption	Access Control	Damaging Content Filtering	Privacy Policy	Open Source Status
GPT-4	Very High	Very High	Very High	Solid	Closed Source
Claude	High	High	High	Solid	Closed Source
Gemini	Middle	Middle	Middle	Middle	Partially Open
DeepSeek	Middle	Low	Middle	Flexible	Open Source
Mistral	Low	Low	Low	Weak	Open Source
Llama	Middle	Middle	Middle	Middle	Open Source
Cursor	High	High	High	Solid	Closed Source

Table 3 shows that there are significant differences between the security and privacy performance of the models [3]; GPT-4 stands out from the other models with very high values in all categories, while Claude and Cursor also perform well, indicating that closed source models offer more secure solutions in general. In terms of data encryption, GPT-4 provides above-standard protection, while Claude and Cursor use high-level but less comprehensive encryption methods; in contrast, open source models DeepSeek and Llama perform moderately, while Mistral performs poorly, suggesting that these models may be particularly risky. Table 4 shows that Mistral and DeepSeek have serious vulnerabilities in terms of access security compared to other models. In malicious content filtering, GPT-4 performs well, Claude and Cursor follow with similar but less advanced systems, while open

source models fall short in this area; Mistral's poor performance suggests that it may be vulnerable to malicious use.

In terms of privacy policies, GPT-4, Claude and Cursor stand out with their strict policies, Gemini and Llama are in the middle level, and Mistral stands out with its weak policies; this reveals that open source models have serious deficiencies in data privacy. As a result, it can be concluded that closed source models are generally more secure, while open source models, despite the advantage of transparency, have security gaps.

Multilingual Skills

The multilingual capabilities of large language models play a critical role in overcoming language barriers on a global scale, and this study systematically evaluates their language coverage and multilingual performance [2]. The analysis reveals the competency levels of the models in different language families and is presented in detail in the table below.

Table 4: Comparative Analysis of Multilingual Capabilities

Model	Multilingual Performance (High/Low)	Coverage in Languages
GPT-4	High	50+
Claude	High	40+
Gemini	Middle	30+
DeepSeek	Low	20+
Mistral	Low	15+
Llama	Middle	25+
Cursor	High	45+

When the data presented in Table 4 is analyzed in detail, GPT-4 and Claude show high performance with more than 40-50 language support. The reason for this is that the model is trained with huge training datasets and advanced language processing algorithms. On the other hand, Gemini and Llama models, which perform at medium level, can meet basic needs with 25-30 languages. The limited support in 15-20 languages offered by models in the low performance category, such as DeepSeek and Mistral, can only provide a basic level of functionality in their most common languages, while their poor performance in text greatly reduces their practical use value, and despite being open source, their lack of language support is clearly visible.

Explainability and Transparency

The explainability of models is an important factor that allows users to clearly understand the output produced by the model [6]. This is because these properties vary from model to model. The lower the explainability of a model, the more likely it is to provide false or misleading information. Considering all these possibilities, the following table was obtained by analyzing the models.

Table 5: Analysis of the Explainability and Transparency of the Model

Model	Explainability (High/Low)	Transparency Features
GPT-4	High	Detailed explanations and visualizations
Claude	High	Detailed explanations and visualizations
Gemini	Middle	Restricted disclosures
DeepSeek	Low	Limited disclosures
Mistral	Low	Limited disclosures
Llama	Middle	Restricted disclosures
Cursor	High	Detailed explanations

As seen in Table 5, GPT-4 and Claude are considered to be the most successful models in terms of explainability and transparency, which allows users to clearly understand the output of the models more effectively and minimizes confusion. However, when we look at DeepSeek and Mistral, user satisfaction is negatively affected due to low explainability and limited explanations, because the more limited the information obtained from the model's output, the more difficult it is for the user to understand.

Conclusions

In this study, GPT-4, Claude, Gemini, DeepSeek, Mistral, Llama and Cursor, the most widely used large language models today, are comprehensively analyzed and their performance is compared according to various critical metrics. In the overall performance evaluations, GPT-4 leads in almost all categories; however, this superiority is directly related to the model's massive training data and continuously updated architecture. Claude and Cursor show similarly high performance, but fall behind GPT-4, especially in complex language processing tasks. On the other hand, the balanced performance of Llama, one of the open source models, is noteworthy, but it should be noted that these models are trained with more limited resources compared to their closed source competitors.

As a result, there is no single ideal solution when choosing large language models; instead, a balance needs to be struck according to the needs of the users. GPT-4, with its wide range of applications and high level of reliability, is an ideal choice, especially for corporate users. However, for researchers and developers who need open source solutions, models such as Llama should also be considered. However, the shortcomings of open source models in terms of security and language support should be carefully evaluated before use. This study is a guiding resource for model developers and users for future research.

Declaration of Interest: The authors declare that there is no conflict of interest.

Author Contributions: The authors contributed equally to the research, writing, model development and discussion sections of the study.

REFERENCES

[1] Alberts, I. L., Mercolli, L., Pyka, T., Prenosil, G., Shi, K., Rominger, A., & Afshar-Oromieh, A. (2023). Large language models (LLM) and ChatGPT: what will the impact on nuclear medicine be?. *European journal of nuclear medicine and molecular imaging*, *50*(6), 1549-1552.

[2]Huang, K., Mo, F., Zhang, X., Li, H., Li, Y., Zhang, Y., ... & Liu, Y. (2024). A survey on large language models with multilingualism: Recent advances and new frontiers. *arXiv preprint arXiv:2405.10936*.

[3]Gerasimenko, D. V., & Namiot, D. (2024). Extracting Training Data: Risks and solutions in the context of LLM security. International Journal of Open Information Technologies, 12(11), 9-19.

[4]Pouransari, H., Li, C. L., Chang, J. H., Anasosalu Vasu, P. K., Koc, C., Shankar, V., & Tuzel, O. (2024). Dataset decomposition: Faster llm training with variable sequence length curriculum. *Advances in Neural Information Processing Systems*, *37*, 36121-36147.

[5]Yao, Y., Duan, J., Xu, K., Cai, Y., Sun, Z., & Zhang, Y. (2024). A survey on large language model (LLM) security and privacy: The good, the bad, and the ugly. High-Confidence Computing, 100211.

[6]Zhao, H., Chen, H., Yang, F., Liu, N., Deng, H., Cai, H., ... & Du, M. (2024). Explainability for large language models: A survey. *ACM Transactions on Intelligent Systems and Technology*, *15*(2), 1-38.

[7]Mumuni, F., & Mumuni, A. (2025). Explainable artificial intelligence (XAI): from inherent explainability to large language models. *arXiv preprint arXiv:2501.09967*.

KALP YETMEZLİĞİ HASTALIĞININ TAHMİNİNDE MAKİNE VE DERİN ÖĞRENME MODELLERİNİN KARŞILAŞTIRMALI ANALİZİ

Taner Erol*

Bolu Abant İzzet Baysal Üniversitesi, Mühendislik Fakültesi, Elektrik ve Elektronik Mühendisliği Bölümü, Bolu, Türkiye

Kemal Polat

Bolu Abant İzzet Baysal Üniversitesi, Mühendislik Fakültesi, Elektrik ve Elektronik Mühendisliği Bölümü, Bolu, Türkiye

**(İletişimden sorumlu yazar) kpolat@ibu.edu.tr*

ÖZET

Dünya genelindeki toplam ölümlerin önemli bir kısmı kardiyovasküler hastalıklar nedeni ile ortaya çıkmaktadır [1]. Geleneksel olarak, standart istatistiksel yöntemler ve doktor sezgisi, bilgisi ve deneyimi prognoz ve hastalık riski tahmini için kullanılmıştır ancak bu uygulama genellikle istenmeyen önyargılara, hatalara ve yüksek masraflara yol açar ve hastalara sağlanan hizmetin kalitesini olumsuz etkiler [2]. Bu bağlamda, mevcut kardiyovasküler rahatsızlığı olan veya risk altındaki bireylerde erken teşhis büyük bir öneme sahiptir. Sağlık verilerinin çıkarımında elektroniğin kullanılabilirliğinin artması, daha sağlam ve gelişmiş hesaplamalı yaklaşımlar ile hastalık tahmini alanında uygulanılmasını ve araştırılmasını daha pratik hale getirmiştir [3]. Risk altındaki bireyler için öngörü geliştirmede makine öğrenimi ve derin öğrenme (DL) modellerinin gelişimi ve çeşitlenmesi bir çok hastalıkta olduğu gibi kardiyovasküler sağlık sorunları için de umut verici bir yaklaşım sunmaktadır. Bu çalışma, kardiyovasküler hastalıklar ile ilişkilendirilen 11 farklı özelliği içeren kapsamlı bir veri setini kullanılarak en iyi öngörüyü elde etmeyi amaçlamaktadır. Bireylerde kardiyovasküler sorunların tahmini için ML ve DL metodolojilerinden yararlanmak bu çalışmanın temelini oluşturmaktadır. Geliştirilen 5 farklı modelin; Rastgele Orman (RF), Destek Vektör Makinesi (SVM), Karar Ağacı (DT), Çok Katmanlı Algılayıcı (MLP) ve Konvolüsyonel Sinir Ağı (CNN). performans metrikleri ortaya koyulmaktadır. 11 semptom üzerinden kalp yetmezliğini tahmin etmede doğruluk, kesinlik, duyarlılık ve f1-score ölçülmüştür. Analiz sonucunda seçilen modellerin dogruluk parametrelerinin en düşüğü dahi %80 üzerinde sonuçlanmış ve %97,28'lık oran ile MLP en yüksek sonuçları üretmiştir. Buradan ML ve DL algoritmalarının klinik çalışmalara katkısının oldukça yüksek olabileceği açıkça görülmektedir. Bu çalışma, doğru tahmin potansiyeli sağlaması ile yapay sinir ağı algoritmalarının özelliklede uygun hiperparametre kombinasyonları veya katman derinlikleri kazandırmış DL algoritmaların kardiyovasküler sağlık alanında tıbbın ilerlemesine katkıda bulunacağını göstermiştir.

Anahtar Kelimeler: Derin öğrenme, erken teşhis, kardiyovasküler hastalıklar, makine öğrenimi.

Giriş

Dünya Sağlık Örgütü'nün (WHO) sağladığı verilere göre, nüfusun neredeyse %31'inin küresel ölümleri kalple ilgili hastalıklardan kaynaklanmaktadır [4]. Erken tespit ve yönetim sayesinde hastalıklara karşı mücadele daha etkin hale gelebilmektedir. Sinir ağları genellikle kalp hastalığı ve beyin hastalığı gibi hastalıkların tahmini için en iyi araç olarak kabul edilir [5]. Bu durumda ML ve DL modellerinin çok yardımcı olabileceği düşünülmektedir. Bu çalışma kalp yetmezliği hastalığına neden olan 11 sempton ile oluşturulmuş veri seti kullanılarak yürütülecektir.

Literatür taraması

Dutta ve arkadaşları [6], çalışmalarında, klinik verilerdeki sınıf dengesizliğini ele almak için evrişimli katmanlara sahip etkili bir sinir ağı önermişlerdir. CNN modelin diğer makine öğrenimi modellerine kıyasla sınıf dengesizliğine karşı daha dirençli ve kararlı bir performans sergilediğini ortaya koymuşlardır. Önerdikleri CNN mimarisinin Koroner Kalp Hastalığı (KKH) varlığını %77, KKH olmayan durumları ise %81,8 doğruluk oranlarıyla sınıflandırabildiğini yapılan testler sonucunda göstermişlerdir.

Nagavelli ve arkadaşları [7], çalışmalarında kalp hastalığının erken teşhisinin sağlık hizmetlerinde kritik bir konu olduğuna değinmişlerdir. Ve çalışmalarını erken teşhis verisi sağlamak amacı ile farklı makine öğrenimi teknolojileri kullanarak gerçekleştirmişlerdir. Kalp hastalığını öngörmek için Naive Bayes ağırlıklı bir yaklaşım kullanılırken sınıflandırma ve kalp yetmezliğini otomatik tanımlama için SVM tercih etmişlerdir. Son olarak aykırı değerlerin tespiti ve ortadan kaldırılması için etkili bir kalp hastalığı tahmin modeli (HDPM) kullanmışlardır. Çalışma sonucunda performans parametreleri olarak hassasiyet, doğruluk, f1 ölçümü ve duyarlılık açısından dört farklı ML model karşılaştırılmıştır.

Olsen ve arkadaşları [8], çalışmalarında ML ve yapay zeka algoritmalarının, kalp yetmezliğinin teşhisi ve tedavisi de dahil olmak üzere, hasta bakımının kişiselleştirilmesi ve iyileştirilmesi için tıpta büyük bir potansiyel kullanım alanı bulunduğunu ve kalp yetmezliğinin teşhisi, sınıflandırılması ve tahmininde makine öğreniminin mevcut uygulamalarını değerlendirmişlerdir.

Wang [9], çalışmasında sıklıkla görülen dengesizlik sınıfı problemi için, kalp yetmezliği tahmini üzerinden 18 popüler ML modelinin, z-skoru veya min-maks normalizasyon yöntemleri ve Sentetik Azınlık Aşırı Örnekleme Tekniği (SMOTE) ile karşılaştırmalı bir çalışma sunuyor. Sonuç olarak kalp yetmezliği tahmini için z-skoru normalizasyonu ve SMOTE kullanımının üstünlüğünü ortaya koyuyor.

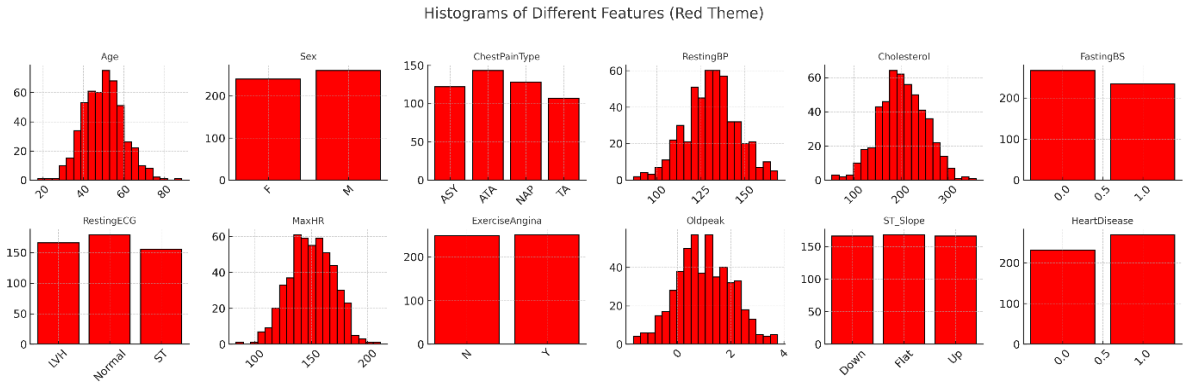
Bharti ve arkadaşları [10], makine öğrenme ve derin öğrenme algoritmaları ile "UCI Machine Learning Heart Disease Dataset" girdisi üzerinden çalışmalarını yürütmüşlerdir. Veri seti, analizi gerçekleştirmek için 14 ana özelliği ele almışlardır. çalışmalarının sonucu olarak performans metriklerini ve derin öğrenme yaklaşımı kullanılarak %94,2 doğruluk elde edildiği paylaşılmıştır.

Yöntem

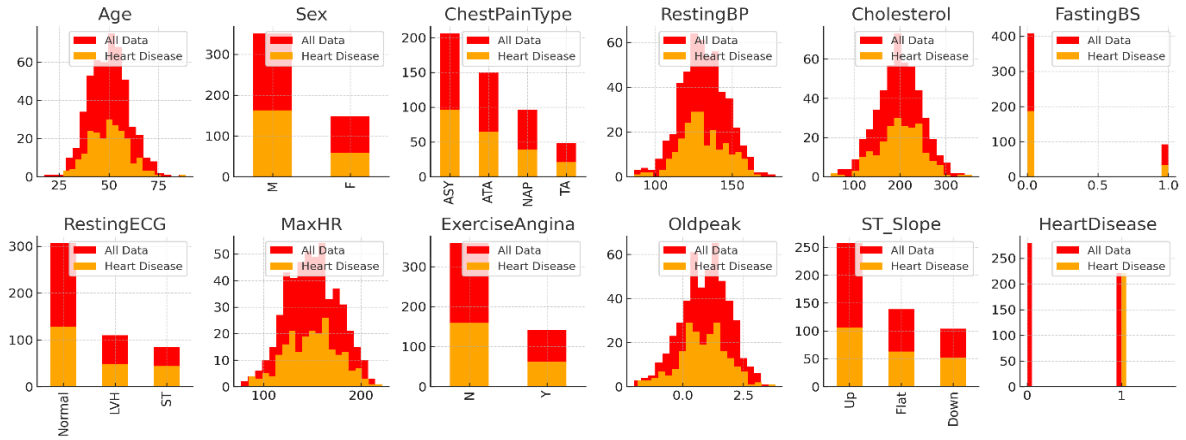
Bu çalışma, 2021 tarihinde Fedesoriano tarafından açık kaynak olarak sunulan "Heart Failure Prediction Dataset" adlı veri seti kullanılarak gerçekleştirilmiştir [11]. Veri seti, bir kalp hastalığını tahmin etmek için kullanılabilecek 11 özelliği içermektedir. Tablo 1 ile veri seti yapısı ve katagorilerin detayları paylaşılmıştır. Çalışmanın analiz ve görselleştirme işlemleri JetBrains tarafından desteklenen PyCharm programı üzerinde yürütülmüş ve kod dizini Python programlama dili kullanılarak oluşturulmuştur. Çalışmanın kapsamını genişletmek amacı ile Kalp hastalığı tahmini için geleneksel ML algoritmalarından Kara Ağacı (RT), Rasgele Orman (DT), Destek Vektör Makineleri (SVM) ve DL algoritmalarından Çok Katmanlı Algılayıcı (MLP), Konvolüsyonel Sinir Ağı (CNN) gibi farklı algoritmalar kullanılarak sonuçlar arası performans metrikleri karşılaştırılmıştır.

Veri seti analizi

Kullanılan veri seti, 11 farklı kategori üzerinden elde edilen 918 kişinin verilerini içermektedir. Bu veriler, kalp yetmezliği hastalığına sahip olma durumunu ortaya koymaktadır. Şekil 1'de sunulan görsel, 11 kategori için kategorinin özelliğine göre kişi dağılımlarını ve kalp yetmezliği hastalığına sahip olma durumuda eklenerek 12 farklı histogram olarak göstermektedir. Şekil 2'de verilen görsel ise bu 12 kategori için kategorinin özelliğine göre kalp yetmezliği hastalığı bulunan kişilerin dağılımlarını Şekil 1 üzerine yerleştirilerek oluşturulmuştur.



Şekil 1: 12 kategori için kişi dağılım histogramları

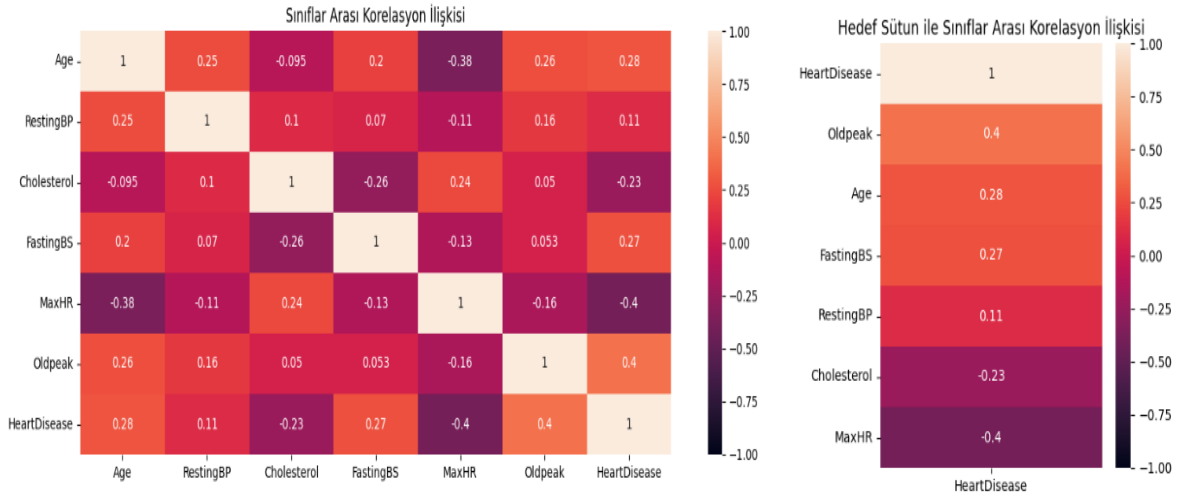


Şekil 2: Kategorilere göre kalp yetmezliği bulunan kişilerin oransal dağılım histogramları

Tablo 1: Veri seti yapısı

Veri sayısı	Öznitelikler			
	Yaş (Age)	**Cinsiyet (Sex)**	**Göğüs ağrısı (ChestPainType)**	**Sistolik kan basıncı (RestingBP)**
1	Sayısal	M	ATA	Sayısal
2	-	F	NAP	-
3	-	-	ASY	-
4	-	-	TA	-
	Serum kolesterol değeri (Cholesterol)	**Açlık kan şekeri (FastingBS)**	**İstirahat halindeki EKG (RestingECG)**	**Egzersiz kalp atımı (MaxHR)**
1	Sayısal	0	Normal	Sayısal
2	-	1	LVH	-
3	-	-	ST	-
4	-	-	-	-
	Egzersizle tetiklenen anjina (ExerciseAngina)	**Egzersiz sonrası depresyon (Oldpeak)**	**ST segment eğimi (ST_Slope)**	**Hedef değişken (HeartDisease)**
1	0	Sayısal	Yükselme	0
2	1	-	Düz	1
3	-	-	Alçalma	-
4	-	-	-	-

Korelasyon matrisi, değişkenler arasındaki güçlü pozitif veya negatif ilişkileri belirlemekte etkilidir. Değişkenler arasındaki ilişkiyi anlamak, özellikle "HeartDisease" hedef sütunu ile diğer sayısal özellikler arasındaki ilişkileri incelemek için, bu çalışmada Pearson korelasyon katsayıları kullanılarak bir korelasyon matrisi hesaplandı. Hesaplanan matris, Şekil 3a'da görselleştirildi. Ayrıca, "HeartDisease" hedef sütunu ile diğer sayısal özellikler arasındaki korelasyonu daha açık bir şekilde göstermek amacıyla Şekil 3b kullanıldı. Renkli hücreler, ilgili iki değişken arasındaki korelasyonun gücünü göstermektedir. Pozitif korelasyonlar genellikle pozitif renk tonları ile temsil edilirken, negatif korelasyonlar negatif renk tonları ile temsil edilir.



(a) (b)

Şekil 3: Korelasyon matris çıktısı (a) her bir kategorinin birbiri ile ilişkisi (b) kalp yetmezliğine sahip olma durumu ile diğer sayısal özellikler arasındaki korelasyon ilişkisi

Veri seti üzerinde gerçekleştirilen veri analizi ve görselleştirmelerin ardından, sinir ağlarını kullanarak model oluşturma işlemine geçilmiştir. Öncelikle tüm veri seti kategorileri eğitim esnasında kullanılacak ise veri seti kategorilerinin hem sayısal hem metinsel bilgi ile oluşturulmuş olması durumuna dikkat edilmelidir. Daha doğru sonuçlar elde etmek için metinsel verilerin sayısal olarak ifade edilmesi ve eğitime eklenmesi gerekmektedir. Burada tüm sınıflar sayısal niteliklere dönüştürülerek modellerin eğitimine dahil edilmiştir. Bu aşamaların ardından sırası ile model oluşturma ve eğitimlerine geçilmiştir.

Algoritma analizi

Destek vektör makinaları (SVM) algoritması

Doğrusal olmayan verilere sahip veri setlerinde Radial Basis Function (RBF), Lineer çekirdeğin aksine çok daha iyi genelleme yapmaktadır. [12]. Buradan hareketle model oluşturulurken karmaşık veri setlerinde Lineer SVM modele göre daha kararlı sonuçlar veren Radial Basis Function (RBF) Çekirdek SVM model tercih edilmiştir.

Karar ağacı (DT) algoritması

DT, en eski ve en yaygın makine öğrenme algoritmalarından biridir. Bir DT, veri öğelerinin sınıflandırılması için sonuçları bir ağaç gibi bir yapıya değerlendiren ve eşleştiren bir şekilde karar mantığını tasarlar [13]. DT sınıflandırma modeli oluşturken modelin hiperparametrelerinin en iyi şekilde ayarlanması gerekmektedir. Burada Grid Search yöntemi ile çeşitli hiperparametre kombinasyonları denenerek, en iyi performansı sağlayan hiperparametre ayarlamaları elde edilmeye çalışılmıştır. Veri seti eğitim ve test veri setlerine ayrılmış ve ardından karar ağacı modeli oluşturulmuştur. Son olarak, elde edilen en iyi modelin karar ağacı ve performans değerlendirmeleri görselleştirilecektir.

Rasgele orman (RF) algoritması

RF , DT'ye dayalı topluluk öğrenmesiyle verileri sınıflandırmak için kullanılan bir yöntemdir [14] [15]. Bu algoritmanın kullanımı için ilk olarak modelin yapısını oluşturmak üzere RandomForestClassifier sınıfı, en iyi performans değerleri için GridSearchCV fonksiyonu ile hiperparametre seçimi ve sınıf dengesizliğini gidermek için SMOTE fonksiyonu kullanılarak yeni sentetik rnekler oluşturulmuştur.

Test seti üzerinden modelin performansı değerlendirilmiştir. Ayrıca modelin doğruluk oranını hesaplanmış, gerçek ve tahmini sınıflar arasındaki karşılaştırmalar görselleştirlmiştir.

Çok katmanlı algılayıcı (MLP) algoritması

MLP algoritması, girdi verisi kümelerini uygun çıktı kümesine eşleyen ileri beslemeli bir yapay sinir ağı modelidir [16]. Bu çalışmada yapay sinir ağı yapısı olusturabilmek için Keras kütüphanesinden yararlanılmıştır. Modelin derlenmesi sırasında, çok sınıflı bir sınıflandırma görevi olduğu için çapraz entropi kaybı (categorical crossentropy) kullanılmıştır. Optimizasyon algoritması olarak 'rmsprop' seçilmiş ve metrik olarak doğruluk ölçütü kullanılmıştır. Ayrıca, eğitim sırasında aşırı uyumdan kaçınmak ve en iyi modelin kaydedilmesini sağlamak amacıyla EarlyStopping adlı bir geri çağrı fonksiyonu eklenmiştir.

Konvolüsyon sinir ağı (CNN) algoritması

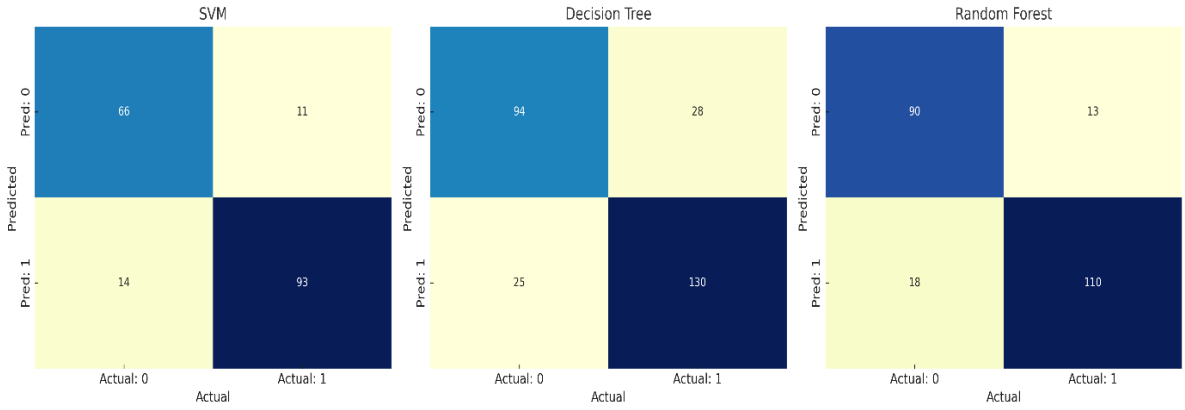
Bu algoritma derin sinir ağlarının, yani birçok gizli katmana sahip sinir ağlarının temsilleridir. Evrişimli sinir ağları, girdi olarak 2 boyutlu görüntüler biçiminde milyonlarca parametre içeren büyük veri kümelerini eğitebilen ve istenen çıktıları üretmek için bunları filtrelerle evriştirebilen derin öğrenme algoritmalarıdır [17]. Bu çalışmada derin öğrenme modelinin inşasında TensorFlow ve Keras kullanılmıştır. TensorFlow açık kaynaklı bir platformdur ve çoklu makine öğrenmesi için kullanılır. Keras ise üst düzey bir sinir ağı kütüphanesidir ve TensorFlow üzerinde çalışır. TensorFlow 2.x sürümleriyle birlikte, TensorFlow içinde Keras'ı "tf.keras" olarak entegre ederek, Keras ve TensorFlow arasındaki sınırları daha da belirsizleştirmiştir.

Modelde 1D konvolüsyon katmanları(Filtre, stride, feature map), dropout katmanları, tam bağlantılı dense katmanlar ve bir sigmoid aktivasyon fonksiyonuna sahip çıkış katmanı bulunur. Ayrıca, eğitim sırasında aşırı öğrenme oranını düşüren ve erken durdurma gibi callback'ler kullanılmıştır.

Eğitim sırasında elde edilen kayıp ve doğruluk metrikleri grafiğe dökülerek eğitim süreci, en iyi epoch'un vurgulanması ve test verisi üzerinde modelin değerlendirilmesi gerçekleştirilmiştir.

Sonuç

Bu aşamaya kadar farklı modellerin tanımlaması ve derlemesi sağlandı. Burada ise her model orijinal veri seti aynı olmak kaydı ile belirli bir veri kümesi üzerinden eğitilerek performans metrikleri ortaya koyulmuştur.

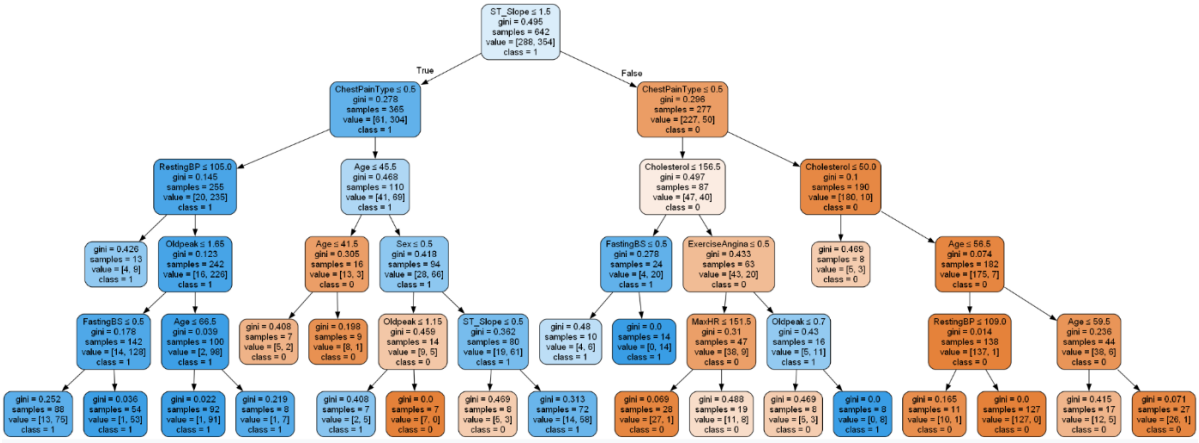


Şekil 4'de gösterilen konfüzyon matrisleri, SVM, DT ve RF algoritmalarının sınıflandırma performanslarını görsel olarak karşılaştırmalı şekilde sunmaktadır. Her bir matris, modelin iki sınıflı bir problemde yaptığı doğru ve yanlış sınıflamaları içermektedir. Satırlar model tahminlerini (predictions), sütunlar ise gerçek sınıfları (actual) temsil etmektedir.

Şekil 4: ML modelerin konfüzyon matrisleri

SVM, DT ve RF algoritmalarına ait konfüzyon matrisleri ve Tablo 2'deki performans metrikleri birlikte ele alındığında, sınıflandırma performansları açısından önemli farklılıklar gözlenmektedir. Özellikle SVM algoritması, pozitif sınıfı tanıma (duyarlılık: 0.894) konusunda en yüksek başarıyı sağlamıştır. Bu durum, modelin pozitif sınıfları doğru şekilde tahmin etme kabiliyetinin oldukça yüksek olduğunu göstermektedir.

RF algoritması ise kesinlik (0.894) metriğinde en başarılı model olarak öne çıkmakta; bu yönüyle yanlış pozitifleri azaltma açısından etkili bir performans sergilemektedir. RF'nin genel doğruluk (0.865) ve F1-Score (0.876) açısından da güçlü bir denge kurduğu görülmektedir.



Şekil 5: DT modelin karar ağacı çıktısı

DT modeli için Grid Search fonksiyonu kullanılarak toplam 143 farklı hiperparametre kombinasyonu bulunmuştur. Her bir hiperparametre kombinasyonu için veri seti 5 katmanlı Cross-Validation ile ayrılmış ve 5 farklı model eğitilmiştir. En başarılı modelin elde ettiği doğruluk, duyarlılık ve F1-score Tablo 2'de ve karar ağacı çıktısı Şekil 5'te görülebilmektedir. DT algoritması hem kesinlik, hem de duyarlılık değerlerinde diğer iki ML modele göre en düşük performansı sağlayan model olmuştur. Bu durum, DT tabanlı modelin overfitting'e eğilimli olduğunu düşündürmektedir.

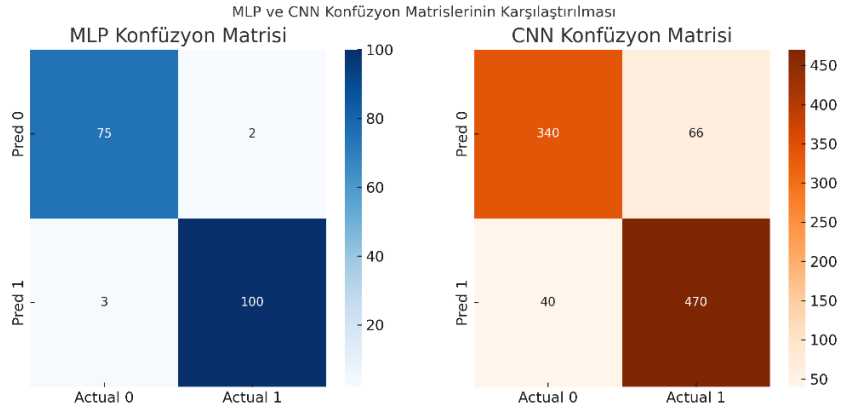
Tablo 2: ML modellerin performans metrikleri

Algoritma	Doğruluk	Kesinlik	Duyarlılık	F1-Score
SVM	0.864	0.869	0.894	0.881
DT	0.803	0.823	0.839	0.831
RF	0.865	0.894	0.859	0.876

Sonuç olarak, SVM modeli pozitif sınıfı yakalama başarısı ile öne çıkarken, RF modeli genel performans açısından daha dengeli bir profil sergilemektedir. DT modeli ise kıyaslanan algoritmalar arasında en zayıf performansa sahiptir.

Derin öğrenme (DL) algoritmaları

Geliştirimiş hali ile DL tabanlı modeller olarak kabul edilebilen MLP, %97.2 doğruluk oranı ve %97.6 F1-Score değeriyle oldukça yüksek bir performans ortaya koymuştur. Şekil 6'da gösterilen konfüzyon matrisleri MLP ve CNN algoritmalarının sınıflandırma performansları ve Tablo 2 ise performans metriklerini karşılaştırmalı şekilde sunmaktadır.



Şekil 6: DL modelerin konfüzyon matrisleri

Buradan görüldüğü üzere modelin duyarlılık değeri olan 0.980, tahmin edilen pozitif sınıfların neredeyse tamamının doğru sınıflandırıldığını göstermektedir. Benzer şekilde, kesinlik değerinin 0.971 olması, pozitif örneklerin büyük çoğunluğunun başarıyla tespit edildiğine işaret etmektedir. Bu sonuçlar, MLP modelinin, uygun hiperparametre ayarlamaları ve dengeli bir veri seti ile son derece etkili hale getirilebileceğini göstermektedir.

Tablo 3: DL modellerin performans metrikleri

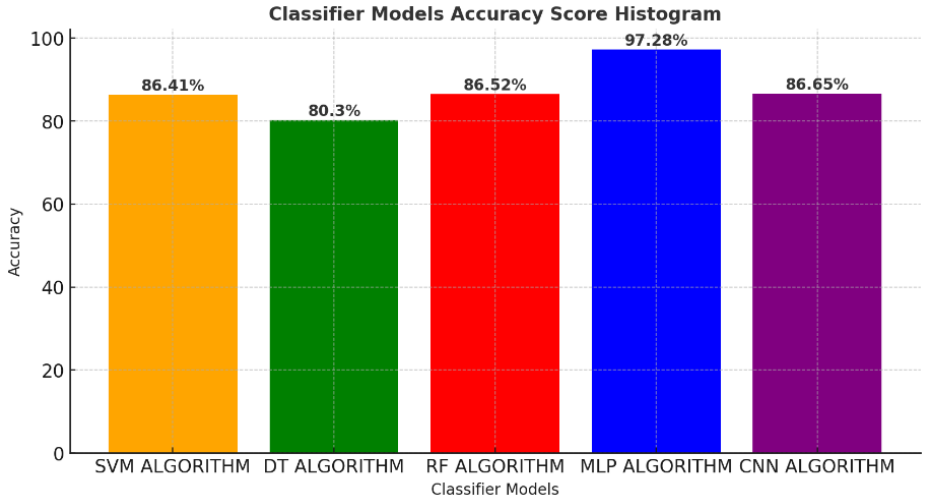
Algoritma	Doğruluk	Kesinlik	Duyarlılık	F1-Score
MLP	0.972	0.971	0.980	0.976
CNN	0.866	0.922	0.877	0.899

Öte yandan, CNN modeli, daha çok görüntü tabanlı veriler üzerinde güçlü bir öğrenme yeteneği sergilemesiyle bilinmektedir. Bu çalışmada CNN, doğruluk oranı açısından %86.6 ile MLP'nin gerisinde kalmakla birlikte, duyarlılık değeri olan 0.877 ile pozitif örnekleri tanıma konusunda tatmin edici bir başarı ortaya koymuştur. CNN'in kesinlik değeri 0.922 olup, bu metrik MLP'ye göre daha düşük düzeydedir; bu durum, modelin zaman zaman yanlış pozitif sınıflandırmalar yapabildiğini göstermektedir. Bu farklılıklar, CNN'in karmaşık örüntüleri öğrenme gücüne rağmen, sınıf dengesizlikleri ya da modele özgü aşırı öğrenme gibi problemlerden etkilenebileceğini göstermektedir.

Sonuç olarak, MLP modeli genel metriklerde CNN'e kıyasla daha üstün bir performans sergilemiştir. Ancak CNN'in yapısal avantajları, özellikle büyük ve karmaşık veri kümelerinde daha belirgin hale gelebileceğinden, uygulama bağlamına göre tercih edilmesi gereken model farklılık gösterebilir.

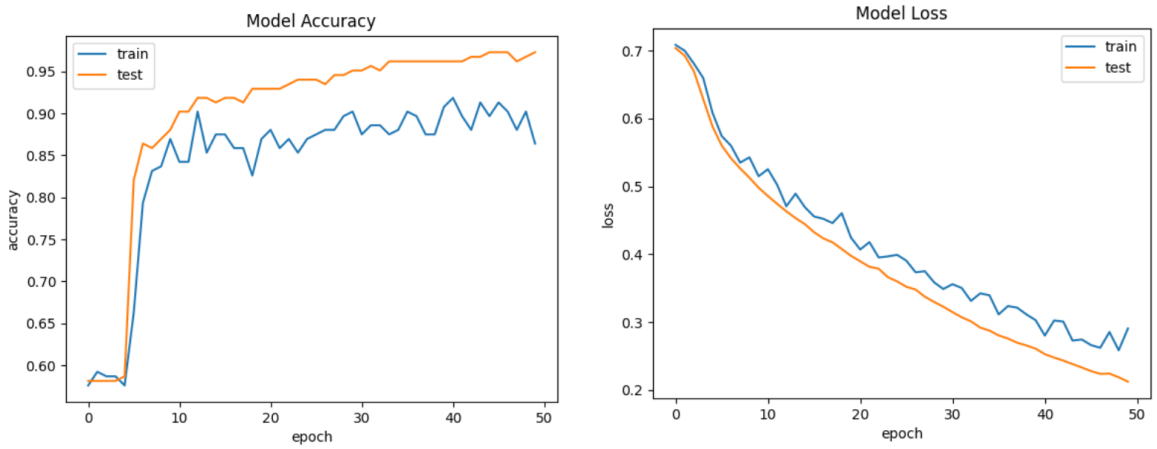
Tartışma

Bu çalışmada, hem klasik ML algoritmaları (SVM, RF, DT) hem de geliştirilmiş şekli ile MLP gibi DL tabanlı yaklaşımlar (MLP, CNN) karşılaştırılmıştır. Elde edilen bulgular Şekil 7'de verilmiştir. Buradan hareketle her iki yaklaşımın performans metrikleri açısından belirgin farklılıklar gösterdiğini ve DL modellerin genel olarak daha üstün sonuçlar elde ettiği söylenebilmektedir.



Şekil 7: ML ve DL modellerin sınıflandırma performansları

CNN gibi derin yapılar, görüntü tabanlı örüntüleri tanımada daha etkin sonuçlar verirken, MLP gibi daha sade yapılar küçük ve dengeli veri kümelerinde overfitting riskini azaltarak stabil sonuçlar elde edebilmektedir. Bu çalışmada açık ara en iyi performans sonuçlarını ortaya koyan MLP algoritmasının train ve test aşamalarındaki doğruluk oranları Şekil 8a'da ve kayıp oranları Şekil 8b'de gösterilmiştir. Genel olarak DL yaklaşımları, daha yüksek doğruluk ve duyarlılık gibi performans göstergeleriyle klasik ML algoritmalarını geride bırakmayı başarmıştır.



(a) (b)

Şekil 8: MLP algoritmasının train ve test oranları (a) doğruluk oranları (b) kayıp oranları

Sonuç olarak, bu çalışma, klasik ML ve DL yöntemlerinin güçlü ve zayıf yönlerini ortaya koyarak, algoritmaların uygun hiperparametre kombinasyonları veya katman derinlikleri kazandırma gibi yöntemlerle daha etkin hale getrilebileceğini kanıtlarken veri kümesinin büyüklüğü ve karmaşıklığına bağlı olarak en uygun modelin seçilmesine yönelik de önemli bir kılavuz sunmaktadır.

Çıkar Beyanı: Yazarlar bu çalışmada herhangi bir çıkar çatışmasının bulunmadığını beyan eder.

Teşekkür: Yazarlar, bu çalışmanın hazırlanmasında doğrudan destek sağlayan herhangi bir kurum veya kuruluş olmadığını belirtir.

Yazar Katkıları: Taner EROL: Çalışmanın kavramsallaştırılması, veri toplama, veri analizi, grafik oluşturma, teknik düzenlemeler ve yazım süreci. Kemal POLAT: Literatür taraması, sonuçların yorumlanması ve düzeltme süreci.

REFERANSLAR

[1] F. S. Alotaibi, "Implementation of machine learning model to predict heart failure disease," *International Journal of Advanced Computer Science and Applications*, vol. 10, no. 6, 2019, DOI: 10.14569/IJACSA.2019.0100637.

[2] S. Palaniappan and R. Awang, "Intelligent heart disease prediction system using data mining techniques," 2008 IEEE/ACS International Conference on Computer Systems and Applications, Doha, Qatar, 2008, pp. 108-115, DOI: 10.1109/AICCSA.2008.4493524.

[3] S. Uddin, A. Khan, M. Hossain, and others, "Comparing different supervised machine learning algorithms for disease prediction," BMC Medical Informatics and Decision Making, vol. 19, no. 281, 2019, DOI: 10.1186/s12911-019-1004-8.

[4] Katarya, R. ve Meena, S.K. "Machine Learning Techniques for Heart Disease Predictio: A Comparative Study and Analysis," Health Technol, cilt 11, ss 87–97 2021, DOI: https://doi.org/10.1007/s12553-020-00505-7

[5] S. Mohan, C. Thirumalai and G. Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques," in IEEE Access, vol. 7, pp. 81542-81554, 2019, DOI: 10.1109/ACCESS.2019.2923707.

[6] Aniruddha Dutta, Tamal Batabyal, Meheli Basu ve Scott T. Acton, "An Efficient Convolutional Neural Network for Coronary Heart Disease Prediction," Expert Systems with Applications, cilt 159, sayı 1, ss 113408, Kasım 2020, DOI: 10.1016/j.eswa.2020.113408.

[7] U. Nagavelli, D. Samanta, ve P. Chakraborty, "Machine Learning Technology-Based Heart Disease Detection Models," Journal of Healthcare Engineering, cilt. 2022, Makale kimliği 7351061, Şubat 2022, DOI: 10.1155/2022/7351061.

[8] C. R. Olsen, R. J. Mentz, K. J. Anstrom, D. Page, ve P. A. Patel, "Clinical applications of machine learning in the diagnosis, classification, and prediction of heart failure," American Heart Journal, cilt. 229, sayı. 1, ss. 1-17, Kasım 2020, DOI: 10.1016/j.ahj.2020.07.009.

[9] J. Wang, "Heart Failure Prediction with Machine Learning: A Comparative Study," in 2nd International Conference on Signal Processing and Computer Science (SPCS 2021), Qingdao, China, Ağustos 2021, DOI: 10.1088/1742-6596/2031/1/012068.

[10] R. Bharti, A. Khamparia, M. Shabaz, G. Dhiman, S. Pande, ve P. Singh, "Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning," Computational Intelligence and Neuroscience, Makale kimliği 8387680, Temmuz 2021, DOI: 10.1155/2021/8387680.

[11] Fedesorıano, "Raster: Heart Failure Prediction Dataset" Eylül 2022, DOI: https://www.kaggle.com/fedesoriano/heart-failure-prediction

[12] Apostolidis-Afentoulis, Vasileios, ve Konstantina-Ina Lioufi. "SVM classification with linear and RBF kernels." Temmuz 2015, DOI: http://www. academia. edu/13811676/SVM Classification with Linear and RBF kernels.

[13] M. M. Ali, B. K. Paul, K. Ahmed, F. M. Bui, J. M. W. Quinn, and M. A. Moni, "Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison," *Computers in Biology and Medicine*, vol. 136, p. 104672, 2021, DOI: 10.1016/j.compbiomed.2021.104672.

[14] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5–32, 2001, DOI: 10.1023/A:1010933404324.

[15] M. M. Ali, B. K. Paul, K. Ahmed, F. M. Bui, J. M. W. Quinn, and M. A. Moni, "Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison," *Computers in Biology and Medicine*, vol. 136, p. 104672, 2021, DOI: 10.1016/j.compbiomed.2021.104672.

[16] M. k. Alsmadi, K. B. Omar, S. A. Noah ve I. Almarashdah, "Performance Comparison of Multi-Layer Perceptron (Back Propagation, Delta Rule and Perceptron) algorithms in Neural Networks,"2009 IEEE International Advance Computing Conference, Patiala, India, 2009, pp. 296-299, DOI: 10.1109/IADCC.2009.4809024.

[17] R. Chauhan, K. K. Ghanshala ve R. C. Joshi, "Convolutional Neural Network (CNN) for Image Detection and Recognition," 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC), Jalandhar, India, 2018, pp. 278-282, DOI: 10.1109/ICSCCC.2018.8703316

COMPARING THE NFT MINTING PROCESS: ETHEREUM VS SOLANA

Gjergji Mulla
University of Tirana Faculty of Economics, Department of Statistics and Applied Informatics, Tirana, Albania

Kristi Gorea*
University of New York Tirana, Department of Information and Intelligent Systems, Tirana, Albania
(Corresponding author) kgorea@unyt.edu.al

ABSTRACT

After Satoshi Nakamoto made the Bitcoin paper public, blockchain has become a very popular technology in 15 years. Its decentralized ledger allows for the storing of transaction records safely. It is the technology that gave rise to Bitcoin. One of the derivates of this technology are non-fungible tokens. Non-fungible tokens are a tool that allow the tokenization of unique digital and physical assets, due to their three defining characteristics of uniqueness, indivisibility and irreplaceability. The leading blockchain ecosystems in non-fungible tokens are Ethereum and Solana. This study compares the technical implementation of non-fungible tokens minting (creation process) on the Ethereum Sepolia testnet and Solana devnet. Using JavaScript-based scripts, we minted a single non-fungible token on each blockchain, uploaded the assets on the InterPlanetary File System (IPFS) and verified the results in MetaMask and Phantom wallets. The process in Ethereum involved a custom smart contract for the non-fungible token minting, Infura for blockchain access and Pinata for IPFS storage, while in Solana the process was streamlined by utilizing the Metaplex framework. Findings indicate that the process is more complex and slower in Ethereum, whereas Solana offers faster transaction and better wallet integration. The results provide the advantages and disadvantages between the two blockchains for non-fungible token development.

Keywords: Blockchain, non-fungible, tokens, minting, Ethereum, Solana.

Introduction

The distributed ledger technology, or, as it is technically known, blockchain, has revolutionized digital transactions by offering secure and transparent peer-to-peer communication[1]. Blockchain was first revealed when Satoshi Nakamato made the debut of Bitcoin as a digital currency designed with the purpose of enabling the decentralization of payments[2]. Since then this technology has evolved and many derivatives have been created with different use cases. One of these transformative derivatives are non-fungible tokens. Due to their three defining characteristics of uniqueness, indivisibility and irreplaceability, non-fungible tokens have enabled the creation and ownership of digital and physical assets through tokenization[3].

Ethereum, one of the second generation blockchains introduced by founder Vitalik Buterin in 2015, has been the dominant ecosystem for the development of non-fungible tokens as a result of its smart contract functionality and wider adoption[4]. The ERC-721 and ERC-1155 token standards provide the technical template for creating non-fungible tokens, allowing developers to generate unique tokens with associated metadata stored on the InterPlanetary File System (IPFS)[5]. However, Ethereum has demonstrated some scalability issues such as high gas fees and slow transaction times that have paved the way for the exploration and adoption of alternative blockchains[6]. Solana ,launched in 2020, is another blockchain that addresses this challenges by using a Proof-of-History consensus mechanism that enables thousands of transactions per second at minimal costs[7]. Solana's ecosystem has gained traction in non-fungible tokens creation also due to tools like Metaplex[8] that serve as decentralized platform designed to facilitate this process.

The choice of blockchain for non-fungible tokens minting involves different advantages and disadvantages when we take into consideration the ecosystem maturity, cost, speed and development experience. On one hand, Ethereum's established infrastructure and development community make it a reliable choice, but the costs can be too high for smaller projects. On the other hand, Solana's efficiency and affordable costs may have a better appeal to developers who want to have rapid development, though its ecosystem is not as established and reliable as Ethereum's. To evaluate these trade-offs, testnets such as Ethereum's Sepolia testnet[9] and Solana's devnet are utilized to replicate mainnet functionalities in a controlled manner and without the associated transactional costs.

The main objective of this paper is to compare the technical implementation of minting a non-fungible token on the Ethereum Sepolia testnet and Solana devnet with a focus on factors such as setup complexity, transaction speed, resource requirements and wallet support. By documenting this processes on both ecosystems, the paper aims to highlight the key differences between the two blockchain for non-fungible token development.

Methodology

This paper uses a comparative analysis to evaluate the technical implementation of non-fungible tokens minting on both blockchains. This is done by minting a single non-fungible token on the Ethereum Sepolia testnet and Solana devnet and thus allowing us to compare in every detail the procedures and technicalities involved in the process. The analysis is made on specific criteria that include setup complexity, the execution steps, transaction speed, costs and wallet support.

Prerequisites

In order to conduct the experiment an initial setup was required to ensure a consistent process on both blockchains. The process was conducted on a Linux-based system, as development on Windows is not supported for the Solana blockchain. Additionally, Node.js and npm were installed to support JavaScript scripts that will be used to create the minting functions on both ecosystems. To create and modify scripts the nano text editor offered by Linux was used. Two blockchain wallets, respectively MetaMask for Ethereum Sepolia and Phantom for Solana devnet were implemented to facilitate the importation

and visualization of the minted non-fungible tokens. Also, an image example, which was uploaded in a decentralized storage, was used.

Ethereum NFT minting process

Setup

The Ethereum minting process begun with the creation of a project directory to organize the scripts and dependencies. A Node.js project was initialized and the three required dependencies were installed using the *npm install ethers axios form-data* method. The ethers library is used to facilitate the interaction with the Ethereum blockchain, while axios and form-data are used to upload files to IPFS via Pinata (a cloud-base IPFS service). An Infura account was created to provide a RPC endpoint that enables interactions with Ethereum Sepolia testnet without needing to run our own node on the blockchain. Metamask wallet was configured to operate on the Sepolia testnet, and funded with Sepolia ETH via a faucet[10] to cover gas fees.

Smart contract

A pre-deployed ERC-721 smart contract was utilized for the minting process. The contract includes key functions for NFT minting such as:

- mintNFT(address to) – mints a non-fungible token to a specified address
- setTokenURI(uint256 tokenId, string memory uri) – assigns the metadata to the non fungible token
- tokenURI(uint256 tokenId) - retrieves the metadata for verification

Execution

To automate the entire process that includes the image upload, non-fungible token minting and metadata assignment, a *mint-nft.js* JavaScript script was used. The script was created using the *nano mint-nft.js* command and executed using the *node mint-nft.js* command. The process began with uploading the image to IPFS via Pinata's API, resulting in this image URI: https://ipfs.io/ipfs/QmQwnyPmm7D2i9vfNYVYpWoL3tz9VtsisSRxyd8He9UtD3. After that the Metadata was generated and uploaded to IPFS, yielding this Metadata URI: https://ipfs.io/ipfs/QmSwuCdu1yNZHk5p85zimG1gmEofxw4Uf1iX3r5eNHcNfG.

{"name":"My First Ethereum NFT","symbol":"MFEN","description":"A unique NFT minted on Ethereum","image":"https://ipfs.io/ipfs/QmQwnyPmm7D2i9vfNYVYpWoL3tz9VtsisSRxyd8He9UtD3","attributes":[{"traitType":"Type","value":"Art"},{"traitType":"Edition","value":"1"}],"sellerFeeBasisPoints":500}

Figure 1: Metadata URI in IPFS

The script then established a connection to the Sepolia testnet using the ethers library, while making use of Infura's RPC endpoint and the wallet's private key. The NFT is finally minted by calling the mintNFT function, targeting the wallet address with a transaction hash as shown in the below picture:

```
kgorea@KristiGorea:~/metaplex-ethereum-project$ node mint-nft.js
Image URI: https://ipfs.io/ipfs/QmQwnyPmm7D2i9vfNYVYpWoL3tz9VtsisSRxyd8He9UtD3
Metadata URI: https://ipfs.io/ipfs/QmSwuCdu1yNZHk5p85zimG1gmEofxw4Uf1iX3r5eNHcNfG
NFT Minted! Transaction: 0xbcb1e31af8edb06aa58af0489c0bf732039e3092a5aa1db57ca8f15da88aeaa0
Token URI set! Transaction: 0xea990752ff70fadf268f922fac5933c877aebb4af1201f7e831cc817637f870b
Token URI: https://ipfs.io/ipfs/QmSwuCdu1yNZHk5p85zimG1gmEofxw4Uf1iX3r5eNHcNfG
```

Figure 2: NFT minting process in Ethereum Sepolia

Verification

The minted non-fungible token was verified by multiple methods to confirm that the process has completed successfully. The mint and set token URI transactions were confirmed on Sepolia Etherscan with the mint transaction hash as: 0xbcb1e31af8edb06aa58af0489c0bf732039e3092a5aa1db57ca8f15da88aeaa0 and the set token URI transacation hash as: 0xea990752ff70fadf268f922fac5933c877aebb4af1201f7e831cc817637f870b. The non-fungible token was imported into the MetaMask wallet using the contract address: 0xcd65588Eb7089BF37C283D90C901eF4fEAc97Dc2 and token ID 0, where it displays with the name "My first Ethereum NFT" as it is shown in the below picture:

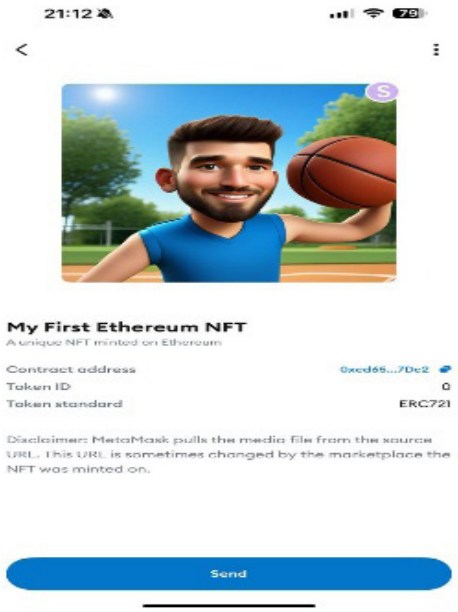


Figure 3: Imported NFT in Metamask

Solana NFT minting process

Setup

Solana makes use of the Solana CLI and the Umi framework of Metaplex to facilitate the creation of non-fungible on the devnet. The setup begin with the installation of Solana CLI (Command Line Interface) using *sh -c "$(curl -sSfL https://release.solana.com/stable/install)"* command. After the installation the Solana CLI was configured to operate on the Solana devnet. A Solana keypair was generated afterwards to create a wallet that was going to be used for transaction tests. A Phantom wallet address was created and configured on the Solana devnet after importing this keypair value and was funded using a Solana faucet[11] to cover the transaction costs. The Umi framework, part of the broader

Metaplex ecosystem, is a JavaScript library that offers a developer-friendly interface to interact with the Solana blockchain, especially in non-fungible token operations. Umi ABSTRACTs away the complexities involved in Solana Rust-based programs by offering the ability to work in a JavaScript environment[12]. This framework was installed by using *npm install @metaplex-foundation/umi-bundle-defaults* command. Also the mpl-token-metadata, responsible for metadata and ownership, and Irys uploader used for asset storage were installed by using the following commands *npm install @metaplex-foundation/umi-uploader-irys @metaplex-foundation/mpl-token-metadata.*

Execution

The minting process on Solana, in a similar way to Ethereum, involved preparing the metadata, uploading the image asset to a decentralized storage and minting the non-fungible token. The key difference here is that all these steps are executed by using the Umi framework of Metaplex. The same image used in the Ethereum example was also used in Solana. The image asset was uploaded to Irys (a decentralized storage commonly used for Solana NFTs) resulting in this image URI: https://gateway.irys.xyz/AB71xN8imQ8DQ3xrW97ASoss6EmRsuD6mUzyqFCQ8U43. Afterwards the metadata file was uploaded to Irys as well producing this metadata URI: https://devnet.irys.xyz/BrSb1noiAPRFMarAxSmar6SyHNiSisKsu6CVEUxvwPgT.

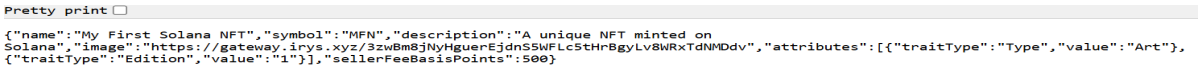


Figure 4. Metadata *URI in Irys*

In the final step the non-fungible token was minted using the *createNft* function from mpl-token-metadata, specifying the metadata URI, name, symbol, and seller fee. The script was executed in the same manner as in Ethereum by executing the *node mint-nft.js* command with an output that includes the non-fungible tokens address and a Solana explorer link for verification as shown below:

Figure 5. NFT minting process in Solana devnet

Verification

The minted non-fungible token was successfully verified on the Solana devnet. The token address 7UxTHZb63ikcvvMVSqdZBcp5MD7LcG7zJ5f5yPxxzDXp, was confirmed on this Solscan URI: https://solscan.io/token/7UxTHZb63ikcvvMVSqdZBcp5MD7LcG7zJ5f5yPxxzDXp?cluster=devnet. The token was later imported in the Phantom wallet we created previously with the address 9BVQH5PLuZc7zD7EGR8tpU9giVF3eTw3ReBxZNxmnxy4 and it displays with the name “My first Solana NFT” as it is shown in the picture below:

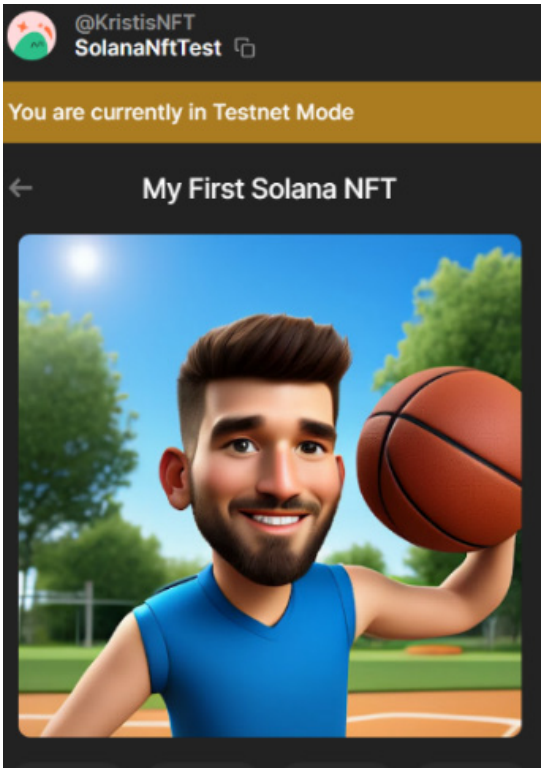


Figure 6. Imported NFT in Phantom

Results

In this section we discuss the key differences highlighted on the non-fungible token minting process on both Ethereum Sepolia testnet and Solana devnet. This discussion interprets the results based on the evaluation criteria defined in the Methodology section.

Comparative analysis

The analysis in this paper is based on the setup complexity, execution steps, transaction speed, costs and wallet support.

In the setup complexity criteria, both blockchains required moderate setup efforts that involved the installation of CLI tools, JavaScript libraries and wallet configurations. Ethereum required additional services like Infura for connecting to the blockchain and Pinata IPFS as a decentralized storage for the assets, compared to Solana whose process was streamlined due to Umi and Irys integrated setup. However, Solana required familiarity with it CLI and devnet environment that may pose a learning curve to new developers.

The process in both cases involved five execution steps that include dependency installations, asset and metadata upload, blockchain connection and non-fungible token minting. Because of the ERC-721 contract template, Ethereum involved an extra step to set the token URI post minting, whereas Solana streamlines the process due to the createNft function that handles metadata during the minting call.

As for transaction speed Solana demonstrated a notable advantage with a confirmation time of under one second, reflecting the capability of its architecture to handle

thousands of transactions per second. Ethereum Sepolia testnet recorded a confirmation time of around 15 second for minting and another 12 seconds for setting the token URI, for a total confirmation time of around 25-30 seconds which confirms the slower transaction times on this blockchain. This makes Solana more suitable for use cases that require rapid transactions.

The transaction costs on Solana for the non-fungible token minting were 0.00001 Sol, which, based on today's exchange rate, is around $0.0014. Ethereum's transaction costs is calculated with the following formula:

Gas used by transaction X Gas price = Transaction fee (1)

Based on this formula, the total cost to mint a non-fungible token in the Ethereum Sepolia testnet was approximately 0.0000001 Sepolia ETH, making the costs significantly lower to Solana in testnet environment. However a mainnet scenario would reverse the situation, as Ethereum's gas price on mainnet would be 20-50 gwei[1] (0.00100001 Gwei in testnet), resulting in a higher cost compared to Solana's fixed low-cost model.

Both wallets provided similar support for non-fungible token visualization, where a manual import of the token to the wallet was needed. This is case only when using the testnet of both blockchains. However the Phantom wallet has the capability to also import Ethereum wallets and store their non-fungible tokens, while in Metamask the Solana blockchain wallets are not supported.

Findings

The results suggest that Solana has better transactions speed and a more streamlined approach, making it a favourable choice for rapid and effortless non-fungible tokens minting, for developers that prioritize speed and wallet integration. However, Solana supports development in Linux based systems only and requires knowledge in Solana CLI which can create initial challenges for developers unfamiliar with the ecosystem.

On the other hand, Ethereum has a mature ecosystem with a stable platform, making it a perfect choice for projects that need advanced smart contract features and wider marketplace compatibility such as OpenSea. However, Ethereum's slower transaction speed and reliance on external services such as Infura or Pinata may be challenging for developers working in low-cost environments.

As for the transaction costs, Solana has relatively low costs on both the devnet and mainnet due to its fixed low-cost model, while Ethereum has lower costs in testnet than Solana, but in the mainnet this situation is flipped due to the high gas prices. Ethereum tries to lower this cost by using layer 2 solutions, but this increases the complexity of implementation for developers.

Conclusions

This paper aimed to compare the technical implementation of non-fungible token minting on the Ethereum Sepolia testnet and Solana devnet, focusing on factors such as setup

[1] 1 Gwei = 0.000000001 ETH

complexity, execution steps, transaction speed, costs and wallet integration and support for evaluation. The experiments successfully minted a non-fungible token on each platform, showcasing the advantages and disadvantages of using either blockchain.

On Ethereum, an ERC-721 smart contract template was used for minting, while the ethers library enabled the interaction with the blockchain and Pinata IPFS was used to upload the asset and metadata. The minting process had a total transaction time of around 25-30 seconds and an approximate cost of 0.0000001 Sepolia ETH, which is significantly lower than Solana but only in the testnet use case.

On Solana the minting process was streamlined thanks to the Umi framework of Metaplex. The createNft function specified the metadata URI, name, symbol, and seller fee of the token, while Irys was used as a decentralized storage to upload the assets and metadata. The process had transaction speed of under one second and relative low costs, making it a favourable option for quick minting and low-cost projects.

To summarize, Solana offers a more efficient minting process in terms of speed and costs, particulary for developers who prioritize speed, low-costs and a streamlined approach. However the requirement of a Linux based system and Solana CLI may add initial challenges to new developers in the ecosystem. Ethereum, while slower and more complex, remains a great option for projects that require advanced smart contract capabilities and ecosystem support.

Future work may include comparing the non-fungible token minting process on other blockchain platforms such as Polygon, Cardano, Binance chain or compare the same blockchains but on a mainnet setting.

Declaration of Interest: The authors declare that there is no conflict of interest.

REFERENCES

[1] L. Abbas, 'Exploring Decentralized Blockchain Technologies: Analyzing Bitcoin, Ethereum, and Solana', MZ Comput. J., vol. 5, no. 1, Art. no. 1, Jan. 2024.

[2] S. Nakamoto, 'Bitcoin: A Peer-to-Peer Electronic Cash System', p. 9.

[3] Q. Wang, R. Li, Q. Wang, and S. Chen, 'Non-Fungible Token (NFT): Overview, Evaluation, Opportunities and Challenges', Oct. 25, 2021, arXiv: arXiv:2105.07447. doi: 10.48550/arXiv.2105.07447.

[4] V. Buterin, 'A Next Generation Smart Contract & Decentralized Application Platform'.

[5] 'ERC721 - OpenZeppelin Docs'. Accessed: Apr. 18, 2025. [Online]. Available: https://docs.openzeppelin.com/contracts/3.x/erc721

[6] J. Sedlmeir, H. U. Buhl, G. Fridgen, and R. Keller, 'The Energy Consumption of Blockchain Technology: Beyond Myth', Bus. Inf. Syst. Eng., vol. 62, no. 6, pp. 599–608, Dec. 2020, doi: 10.1007/s12599-020-00656-x.

[7] A. Yakovenko, 'Solana: A new architecture for a high performance blockchain'.

[8] 'Metaplex Developer Hub'. Accessed: Apr. 18, 2025. [Online]. Available: https://developers.metaplex.com/

[9] 'Goerli vs. Sepolia Testnet | QuickNode Guides'. Accessed: Apr. 18, 2025. [Online]. Available: https://www.quicknode.com/guides/ethereum-development/getting-started/goerli-vs-sepolia-a-head-to-head-comparison

[10] 'Ethereum Sepolia Faucet'. Accessed: Apr. 19, 2025. [Online]. Available: https://cloud.google.com/application/web3/faucet/ethereum/sepolia

[11] 'Solana Devnet Faucet - Airdrop SOL'. Accessed: Apr. 19, 2025. [Online]. Available: https://faucet.solana.com/

[12] 'Overview | Umi'. Accessed: Apr. 19, 2025. [Online]. Available: https://developers.metaplex.com/umi.

CONDİLLAC VE YAPAY ZEKÂ: ALGIDAN ZİHİNSEL YAPIYA BİR YOLCULUK

Hilal Kahraman
İzmir Kâtip Çelebi Üniversitesi, Felsefe Bölümü, İzmir, Türkiye
**(İletişimden sorumlu yazar) hilal.kahraman@ikc.edu.tr*

ÖZET

Etienne Bonnot de Condillac, 18. yüzyılın Fransız Aydınlanması düşüncesinde önemli bir figürdür. Felsefi çalışmaları, özellikle insan zihninin doğası ve bilgi edinme süreçleri üzerine geliştirdiği teorilerle tanınmaktadır. Condillac'ın, John Locke'ın empirizminin etkisiyle savunmuş olduğu zihnin doğuştan boş bir levha (tabula rasa) olduğu ve tüm bilgi ediniminin yalnızca deneyim yoluyla mümkün olduğu yönündeki görüşü, dönemin epistemolojik anlayışlarını önemli ölçüde şekillendirmiştir. Condillac'a göre duyular, her türlü bilgi edinmenin temel kaynağıdır; hiçbir şey bilmediğimizde, ancak duygularımızla tanıyabiliriz. Bu felsefi perspektif, yapay zekâ (YZ) ile ilgili günümüzdeki tartışmalara ışık tutabilecek nitelikte bir temel oluşturmaktadır. Zira yapay zekâ, insan zihninin bilişsel işlevlerini taklit etmeye yönelik bir teknoloji olarak, Condillac'ın düşüncelerinin bir tür dijital yansımasıdır. Condillac'a göre, insan zihninin temel işlevi, duyusal verileri işleyerek anlam üretmektir. Algı, dış dünyadan gelen verilerin zihinsel yapılarla şekillenmesinin ilk adımıdır. YZ, benzer bir biçimde, dış dünyadan aldığı verileri işleyerek anlam üretmektedir, ancak burada kullanılan yöntem ve mekanizma farklıdır. YZ sistemleri, bilgisayar algoritmaları aracılığıyla öğrenir, analiz eder ve kararlar alır. Bu yazı, Condillac'ın epistemolojik anlayışını ve onun algı teorisini, yapay zekâ ile ilişkilendirerek her iki alanın epistemolojik ve ontolojik yansımalarını inceleyecektir.

Anahtar Kelimeler: Algı, Condillac, yapay zekâ.

Giriş

Bilgi felsefesi, insan zihninin bilgiyi nasıl elde ettiğini ve bu sürecin temel dinamiklerini inceleyen disiplinlerden biridir. 18. yüzyıl Fransız düşünürü Étienne Bonnot de Condillac, bu alanda özellikle duyulara dayalı bilgi edinme anlayışıyla dikkat çekmektedir. Condillac'ın epistemolojisi, deneyimi ve duyusal algıyı merkeze alarak, zihnin doğuştan boş bir levha (tabula rasa) olduğunu ve çevresel uyarıcılara maruz kaldıkça geliştiğini savunur. Bu yaklaşım, günümüzde yapay zekânın öğrenme ve veri işleme yöntemleriyle anlamlı bir karşılaştırma yapmaya imkân tanımaktadır. Bu çalışmada, Condillac'ın duyusal algı ve bilinç anlayışıyla, yapay zekâ sistemlerinin bilgi edinme süreçleri arasındaki epistemolojik ve ontolojik ilişkiler tartışılacaktır.

Duyusal Algının Zihinsel Süreçlerdeki Rolü

Condillac'ın bilgi teorisinin temelinde, insan zihninin tüm bilgi edinme sürecini, dış dünyadan gelen duyusal verilerle şekillendirdiği fikri yer almaktadır. Locke'un *tabula rasa* anlayışını bir adım öteye taşıyan Condillac, doğuştan boş bir zihnin, çevreden gelen uyarıcılara tepki vererek bilgi edinmeye başladığını savunmaktadır. Condillac'a göre, bilgi yalnızca deneyim yoluyla edinilebilir ve her bilgi parçası, duyusal algıların ürünü olarak zihinde şekillenmektedir. Ona göre, "İnsan zihni, dış dünyadan gelen algılara dayalı olarak gelişir. Bu algılar, insan düşüncesinin ilk temellerini atar." [1] Burada dikkat çeken nokta, algının, yalnızca pasif bir duyusal algılama değil, aktif bir zihinsel süreç olduğudur. Condillac, insanın dünyaya dair bilgiye nasıl ulaştığını yaratıcı bir şekilde tanımlarken, mermer bir heykel (ya da "hayali heykel") örneğini kullanmaktadır. [2] Condillac, bu heykel aracılığıyla, bilginin edinilmesinin sürecini, koku duyusundan başlayarak her bir duyuyu ayrı ayrı ele alarak açıklamaktadır. Ayrıca, dokunma duyusu olmaksızın, nesnelerin niteliklerinin tam anlamıyla kavranamayacağını ileri sürmektedir. Bu bağlamda, zihnin bilgi ve yeteneklerinin çoğunlukla dokunma duyusunun ürünü olduğuna dikkat çekmektedir. Dış dünyayı kavrayışımızın, özellikle dokunma ve görme duyularının gelişimiyle şekillendiğini belirtirken, dokunma duyusunun diğer duyuların "öğretmeni" olduğunu savunmaktadır. [3] Condillac, duyumlamanın sabit bir varlık değil, dış dünyaya etkin bir katılım olduğunu vurgulamakta ve bu anlayışla birlikte ilk dışsal algı fenomenolojisini geliştirmektedir. Condillac'a göre, algı birden fazla duyu organının etkileşimiyle gerçekleşmektedir ve bu etkileşim, zihnin aktif bir şekilde yapılandırılmasına olanak tanımaktadır. Örneğin, bir nesneye bakıldığında, görsel algı yalnızca ışık dalgalarının göz tarafından algılanmasından ibaret değildir. Aynı zamanda bu görüntüyle ilgili anlamlar, kavramlar ve geçmiş deneyimler de zihinde bir araya gelmektedir. [4] Bu, Condillac'ın düşüncelerine göre, zihinsel yapının temellerinin atılmasına olanak vermektedir. Condillac'a göre "Algı, duyu organlarımızın sadece bir sonucu değil, zihinsel faaliyetlerin başladığı bir noktadır." [5] O böylece duyusal algıların zihnin işlevsel yapısının temellerini attığını vurgulamaktadır.

Yapay zekâ, Condillac'ın duyusal algılarla şekillenen bilgi edinme modelinin bir teknolojik simülasyonudur. Yapay zekâ sistemleri, çevreden aldıkları verileri işlemekte ve bu verilerle anlamlı sonuçlar üretmektedir. YZ, görsel, işitsel veya dokunsal veriler gibi çevresel uyarıcılara tepki vermektedir ve bu tepkiyi, algoritmalar aracılığıyla işlemleyerek anlam üretmektedir. Bu süreç, Condillac'ın epistemolojik teorisiyle paralellik gösterir; zira her iki durumda da bilgi, dış dünyadan gelen verilerle şekillenmektedir. [6] Ancak, Condillac'ın felsefesindeki bu süreç, insanın bilinçli deneyimlerini ve duygusal algılarını içerirken, yapay zekâ tamamen veri odaklı bir öğrenme sürecine dayanmaktadır. YZ, insan benzeri bir bilinçten yoksundur ve yalnızca mekanik bir işleyişle öğrenmektedir. YZ'nın bu öğrenme ve algılama süreçleri, Condillac'ın "Algı, insanın yalnızca dünyayı anlamakla kalmayıp, aynı zamanda ona dair bir farkındalık geliştirmesinin temel aracıdır." [7] şeklindeki düşüncesiyle de örtüşür. Her iki sistem de dış dünyadan gelen uyarıcılara dayalı olarak bilgi oluşturma kapasitesine sahiptir. Ancak, burada temel fark, yapay zekânın algoritmalara ve veriye dayalı bir işleyişle bu bilgiyi üretirken, insanın bilinçli algı ve duygusal deneyimlerini işleyerek anlam üretmesidir.

Bilinç ve Öznellik

Condillac'ın algıya dayalı bilgi edinme teorisi ile yapay zekâ arasında benzerlikler derinlikli bir düşünme konusu olabilmektedir. Her iki sistem de çevreden gelen verileri işleyerek anlam oluşturma yeteneğine sahiptir. Ancak, bu süreçte bir temel felsefi fark bulunmaktadır: Condillac'a göre, algı yalnızca bir bilişsel etkinlik değil, aynı zamanda bir bilinç durumunun yansımasıdır. İnsan zihni, çevresel uyarıcıları işleyerek bir tür öznellik geliştirmekte ve bu öznellik, anlam üretme sürecine uyum sağlamaktadır. Ona göre bilinç, yalnızca algıların toplamı değil, onların içsel olarak işlenmesiyle ortaya çıkan bir niteliktir. [8] Bu duruma bağlı olarak Condillac, algının duygusal ve bilişsel bir etkileşimin ürünü olarak, dünyayı anlamlandırmada temel bir rol oynadığını öne sürmektedir. [9] Bu nedenle, algı sadece bilgi edinmenin bir aracı değil, aynı zamanda bilinçli deneyimlerin bir parçasıdır. Bununla birlikte, yapay zekâ tamamen farklı bir düzlemde çalışmaktadır. YZ, çevresel verileri işlemek için algoritmalar kullanmaktadır, ancak bu süreçte herhangi bir bilinçli deneyim bulunmamaktadır. YZ sistemleri, duygusal, bilinçli veya öznel bir deneyimi işlemeksizin, yalnızca objektif verilerle çalışmaktadır. Yapay zekâ Condillac'ın bahsettiği türden bir öznel farkındalık geliştirmemektedir. Verileri işleyip, öğrenip, karar vermektedir; ancak bunu bilinçli bir içsel deneyim olmadan yapmaktadır. Dolayısıyla yapay zekâ ile insan zihni arasındaki temel fark epistemolojik değil, ontolojiktir. YZ, veri işleyebilir ama anlam "yaşayamaz". Bu bağlamda, yapay zekânın bilinci olup olamayacağı sorusu, Condillac'ın algıdan bilinçli düşünmeye geçişine dair felsefi soruları gündeme getirmektedir. YZ'nın potansiyel bilinç durumu, insan zihninin dış dünyayla kurduğu bilinçli etkileşimin, yapay bir sistemde simüle edilip edilemeyeceği sorusuna dayanmaktadır.

Condillac'ın algıdan bilince geçişi, yapay zekâ tartışmalarını derinleştiren önemli felsefi sorular doğurmaktadır. Bir yapay zekânın insan benzeri bir bilinç geliştirmesi mümkün müdür? Bu, yapay zekânın bilinçli bir düşünceye ulaşma kapasitesinin olup olmadığını sorgulayan bir sorudur. Condillac'a göre, bilinç, dış dünyadan gelen uyarıcıların bir birleşimiyle değil, bu uyarıcıların içsel bir biçimde işlenmesiyle ortaya çıkmaktadır. O bu düşünceleriyle bilincin yalnızca algılarla şekillenen bir süreç olduğunu savunmaktadır. Bununla birlikte, yapay zekânın ontolojik olarak bilincin ötesine geçip geçemeyeceği sorusu, sadece teknik bir mesele değil, aynı zamanda insan bilincinin doğasına dair bir felsefi çözümleme gerektirmektedir. Yapay zekânın öğrenme süreci, Condillac'ın epistemolojik perspektifiyle örtüşse de, bu süreç bilinçten yoksundur. YZ'nın bilincini tartışırken, insan zihninin öznel deneyimlerinin doğasını daha derinlemesine incelemek gerekecektir. Eğer bilinç, Condillac'ın belirttiği gibi içsel bir süreçse, algoritmaların bunu simüle etmesi mümkün müdür? Bilgi edinme süreçleri benzer olsa da, bilinçli deneyimlerin eksikliği, yapay zekânın insan benzeri bilinç geliştirmesinin önünde felsefi bir engel olarak durmaktadır. [10] Condillac'ın felsefesi, yapay zekânın epistemolojik temellerine dair önemli bir perspektif sunmaktadır. Ancak yapay zekânın gelişimindeki ontolojik sorular, zihnin sadece algı ve öğrenme süreçlerinden ibaret olup olmadığına dair daha derin bir sorgulamayı da gerektirmektedir. Yapay zekânın bilinçli olup olamayacağı sorusu, Condillac'ın zihnin dış dünyayla ilişkili yapısını simüle etmekle kalmayarak, aynı zamanda insan zihninin öznel deneyimlerine dair felsefi bir tartışma açmaktadır.

Sonuç

Sonuç olarak, 18. yüzyıl düşünürü Étienne Bonnot de Condillac'ın duyusal algıya dayalı bilgi edinme teorisinin yapay zekâ sistemlerinin veri işleme süreçleriyle karşılaştırmalı olarak ele alınması felsefi düzlemde ilişkilendirilebilecek iki zihin açıklamasının ayrıntılarıyla incelenmesini gerektirmektedir. Condillac'ın bilgi anlayışı, zihnin dış dünyadan gelen duyusal verilerle yapılandığını savunurken; yapay zekâ, çevreden aldığı verileri algoritmalar aracılığıyla işleyerek anlam üretmektedir. Her iki sistemin bilgi oluşturma süreçleri arasında epistemolojik benzerlikler bulunsa da, bilinç ve öznellik düzeyinde önemli ontolojik farklar bulunmaktadır. Bu bağlamda, yapay zekânın insan benzeri bir bilinç geliştirip geliştiremeyeceği sorusu, Condillac'ın algıdan bilince geçiş anlayışı temelinde yeniden değerlendirilmektedir. Condillac'ın algıdan bilince olan geçişini ve yapay zekânın potansiyel bilinç sorunlarını ele alırken, bu iki alanın epistemolojik ve ontolojik boyutlarını anlamak, yapay zekânın felsefi geleceği hakkında daha geniş bir tartışma başlatmaya imkân sunmaktadır. Yapay zekanın yalnızca işlevsel bir model değil, aynı zamanda insan bilincine dair bir simülasyon olup olamayacağı sorusu, modern felsefenin ve teknolojinin birleşim noktalarından biridir. Bu tartışma, insan zihninin doğası ve yapay zekânın geleceği hakkında daha derin bir anlayış geliştirmemize olanak sağlayacaktır.

REFERANSLAR

[1] É. B. de Condillac, "Traité des sensations", I. Kitap, Paris: 1754.
[2] É. B. de Condillac, "Traité des sensations", I. Kitap, Paris: 1754.
[3] É. B. de Condillac, "Traité des sensations", II. Kitap, Paris: 1754.
[4] É. B. de Condillac, "Traité des sensations", III. Kitap, Paris: 1754.
[5] É. B. de Condillac, "Traité des sensations", IV. Kitap, Paris: 1754.
[6] J.-C. Pariente, Condillac et la naissance du langage et de la conscience, 63-67. Paris: PUF, 1985.
[7] É. B. de Condillac, "Traité des sensations", IV. Kitap, Paris: 1754.
[8] M. Malherbe, "Condillac: L'expérience et la connaissance", s. 101-104, Paris: Vrin, 1995.
[9] É. B. de Condillac, "Traité des sensations", III. Kitap, Paris: 1754.
[10] T. Metzinger, "The Ego Tunnel: The Science of the Mind and the Myth of the Self", 42-50. New York: Basic Books, 2009

CONNECTOGRAM: A NOVEL TIME-AWARE GRAPH REPRESENTATION FOR EEG-BASED EPILEPTIC SEIZURE DETECTION

Serkan Aksu*
Bartın University, Department of Computer Technologies, Bartın, Türkiye

Bayram Akgül
Bartın University, Department of Computer Engineering, Bartın, Türkiye

Fatih Dinç
Bartın University, Department of Computer Technologies, Bartın, Türkiye
*(Corresponding author) aksu@bartin.edu.tr

ABSTRACT

Different methods have been used to analyze and classify time series such as electroencephalography (EEG) signals. One of the proposed fast and useful methods is visibility graphs in the time domain. In this study, single-channel EEG signals for epileptic seizure detection are transformed into graph representations consisting of images using an innovative method. The connectogram representation, a colored graph that we propose in this study, is an innovative approach that extends existing graph transformation methods used in time series by adding the time dimension. Thus, we contribute to the literature with an innovative time-convexity based graph representation.

In the connectogram representation time series of EEG signals are represented by 3-layer RGB images. When the images generated by this method were classified using the Xception Transfer Learning method, 98.67% classification success was obtained for normal, epileptic interictal and epileptic ictal conditions. The results obtained in this study show the effectiveness of the proposed innovative method in diagnosing epileptic seizures and detecting other brain activities compared to other methods.

Keywords: EEG signals, epilepsy, time series, visibility graphs, connectogram, deep learning.

Introduction

Epilepsy is a neurological disorder caused by unusual and excessive electrical activity of nerve cells in the brain, causing mental problems such as sensation and awareness and physical problems such as muscle contraction. About one percent of the world is affected by epilepsy, a baffling neurological disorder characterized by seizures [1]. Over the last decade, graph-based time series analysis methods have made great progress, both in statistical comparisons with classical graph metrics and in modern Deep Learning (DL) integrations. Initially comparing epileptic and healthy brain networks to identify unique network biomarkers, the field is now evolving towards artificial intelligence systems capa-

ble of real-time seizure detection and prediction. These developments have brought the concept of network science to a central position in epilepsy research and created a new paradigm in epilepsy diagnosis.

Diagnosis of epileptic disorders usually requires long-term examination of Electroencephalogram (EEG) signals, which are caused by the magnetic effects of electrical currents in the brain. The large number of EEG signals and signal lengths make this diagnosis costly in terms of time and processing power [2]. Brain activity in the ictal state (during a seizure) can differ significantly from normal activity in terms of the frequency and pattern of triggering in neurons. Furthermore, given factors such as the excessive presence of myogenic artifacts, seizures can be difficult to detect by visual inspection of the EEG, even by a trained neurologist. Predicting future seizures is even more difficult as there is little confirmed information on the exact mechanism responsible for the seizure [2]. There are effective algorithms in the literature for automatic seizure detection and prediction for epilepsy diagnosis and treatment.

Graph-based analysis of electroencephalography (EEG) signals aims to make sense of multichannel data by modeling the brain as a network (graph). According to this approach, the brain regions to which each EEG electrode is connected are represented as nodes, and the interactions between these regions are represented as edges [3]. A connectivity matrix is created by calculating the functional connectivity strength (e.g. correlation, coherence or phase matching) between all electrode pairs over a given time interval. This matrix transforms the time series of EEG into a network structure; the strength of interactions between nodes is reflected as weights on the edges of the graph. Especially in neurology, these functional connectivity networks are often modeled as weighted or binary [3].

There are alternative methods for analyzing EEG signals by transforming them into graphs, such as the visibility graph approach. Visibility Graphs, take each data point in the time series of a single channel as a node and form a graph by defining a geometric visibility relationship between these points. The visibility graph allows to represent the structure of the signal with complex network theory and is known for being robust to noise and not too sensitive to parameter choices [4]. Figure 1 shows an example of VG obtained from a randomly generated time series [5].

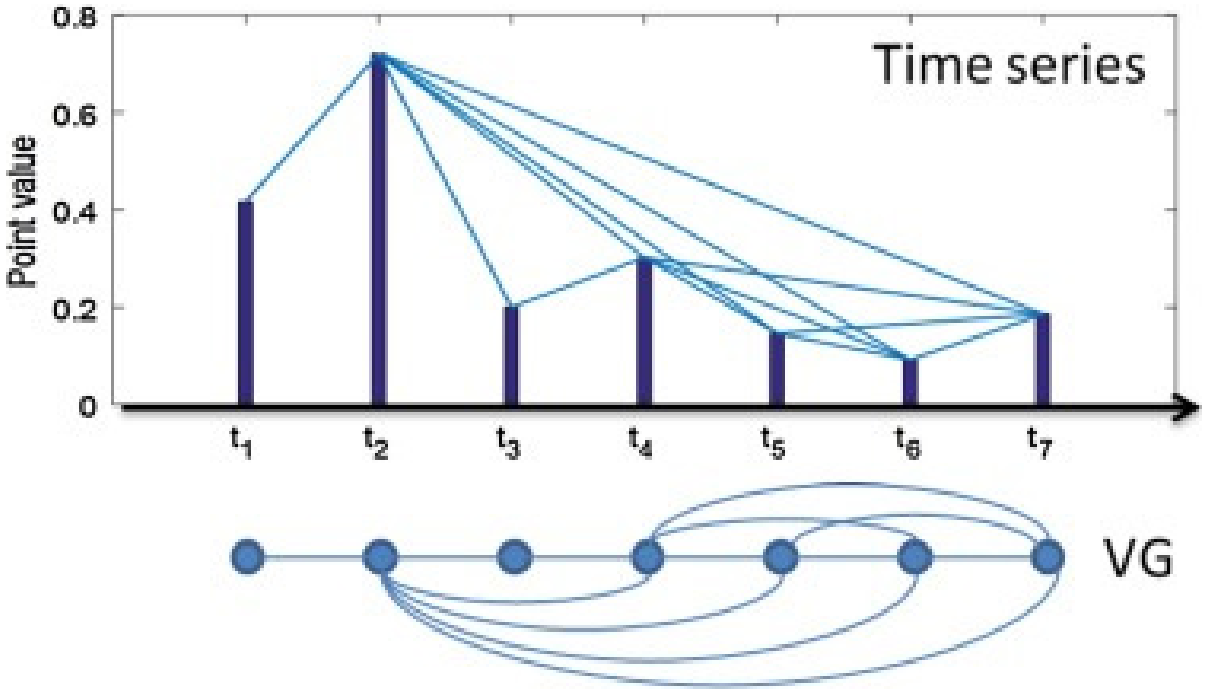


Figure 1: An example of a visibility graph obtained from the connection between seven points of a randomly generated time series.

In the diagnosis of epileptic disorders, graph theory-based analyses offer a number of important advantages. First, these methods provide a network-level perspective, focusing on the global connectivity patterns of the brain, not just individual electrode activity. Epilepsy is seen as a common network disorder, especially in refractory cases. Indeed, studies in temporal lobe epilepsy have shown that the epileptic brain exhibits "more organized" (more lattice-like) network structures in the interictal period, and that the balance of local segregation and global integration is abnormally altered [6]. This indicates an abnormal topological organization in the epileptic brain. Graph analysis can identify network features (e.g. abnormal hub nodes, regions of high connectivity density) that may be conducive to the onset and spread of seizures. Such findings suggest that graph methods can capture pathological brain activity in a more holistic perspective [6].

Second, graph-based analyses have the potential to reveal foci that may remain hidden in clinical EEG. Epileptic foci may not always be in the area with the most prominent activity on the EEG; the clinical signs of a seizure may be caused by activity in another part of the brain. In such cases, focal detection becomes difficult with classical methods, whereas graph theoretical analysis of EEG can capture the indirect effects of epileptogenic regions [3].

In the literature, it has been emphasized that epileptic foci may be far from the area where clinical symptoms are seen during a seizure and that the integration of EEG + graph theory is an important step in detecting these foci that are not evident in simple EEG examination [3].

Graph theory analyses are also advantageous for predicting seizures. Changes in connectivity patterns in the brain shortly before epileptic seizures occur can be captured

by graph-based metrics. Recent studies have shown that deviations in dynamic network properties can be harbingers of seizures. For example, Bomela et al. (2020) extracted time-varying functional networks from EEG data from 23 epilepsy patients and evaluated the topological properties of each moment, successfully detecting 93.6% of seizures in a realistic patient dataset and reporting a low false positive alarm rate of only 0.16 per hour (FP/h) [7]. Moreover, this study was able to observe pre-ictal period signals by detecting topological changes just before the seizure in some cases. Such findings suggest that graph-based analytics may form the basis of future systems that can warn patients before the onset of a seizure. Seizure prediction is a potentially groundbreaking advantage in epilepsy management, offering the patient and physician the opportunity to take proactive measures before a seizure starts [3].

Another important advantage of graph-based methods is the interpretability and transparency of the data obtained with this method. While classical DL approaches often act as a "black box", the outputs of graph-based models can be correlated with concrete changes in the brain network. For example, an increase in the weight of certain edges in a graph model may indicate that two regions in the brain are abnormally synchronized; or if the centrality of a particular node is high, it may indicate that this region plays a key role in the distribution of epileptic activity. In the graph convolutional network-based model proposed by Jia et al. in 2022, it was reported that the possible location and types of abnormal discharges could be predetermined by examining the edge features, thus making the model more clinically understandable [8]. Similarly, other studies have shown that graph metrics (e.g., "centrality of control") can be used to label brain regions as "synchronizing" or "desynchronizing," providing insight into which regions play a role in triggering or inhibiting seizures [9].

A recent work on the combination of DL and graph analysis includes hybrid models such as GCN-LSTM. In 2023, Kuang et al. presented a model that combines graph convolution layers and long-short-term memory networks for EEG-based seizure prediction. This model learns the spatial connectivity structure of EEG with GCN and processes its temporal patterns with LSTM. In experiments on the CHB-MIT dataset, an average accuracy of 99.39% was obtained for two-class (seizure vs. normal) discrimination and 98.69% for three-class (preictal, ictal, interictal) discrimination [10].

Methods such as GCN-LSTM are an indication of how fast graph-based DL approaches are evolving and how AI-supported network analysis has become an important area of research in epilepsy detection in the last 10 years.

Data and Methods

In this study, we used the Bonn dataset created by the Epilepsy Center of the University of Bonn, Germany, which consists of five single-channel EEG signal sets [11]. Sets A and B consist of EEG signals from five healthy subjects with eyes open and eyes closed, respectively. Sets C and D consist of interictal (the phase between seizures, when the person is not actively having a seizure.) EEG signals from the hemisphere of the epileptogenic region and the opposite hemisphere. Set E consists of ictal (the phase during a seizure) EEG signals.

Temporal lobe epilepsy with hippocampal formation was diagnosed for all epileptic signals. Each signal set contains 100 EEG segments with a duration of 23.6 s and a sampling frequency of 173.61 Hz. To remove noise from the signals, a Wavelets band-pass filter was applied and the spectral band range was set from 0.5 Hz to 85 Hz. An example of the EEG signals of healthy (normal: Z and O group) [A], epileptic but seizure-free (interictal: N and F group) [B] and epileptic seizure (ictal: S group) [C] subjects, respectively shown in Figure 2.

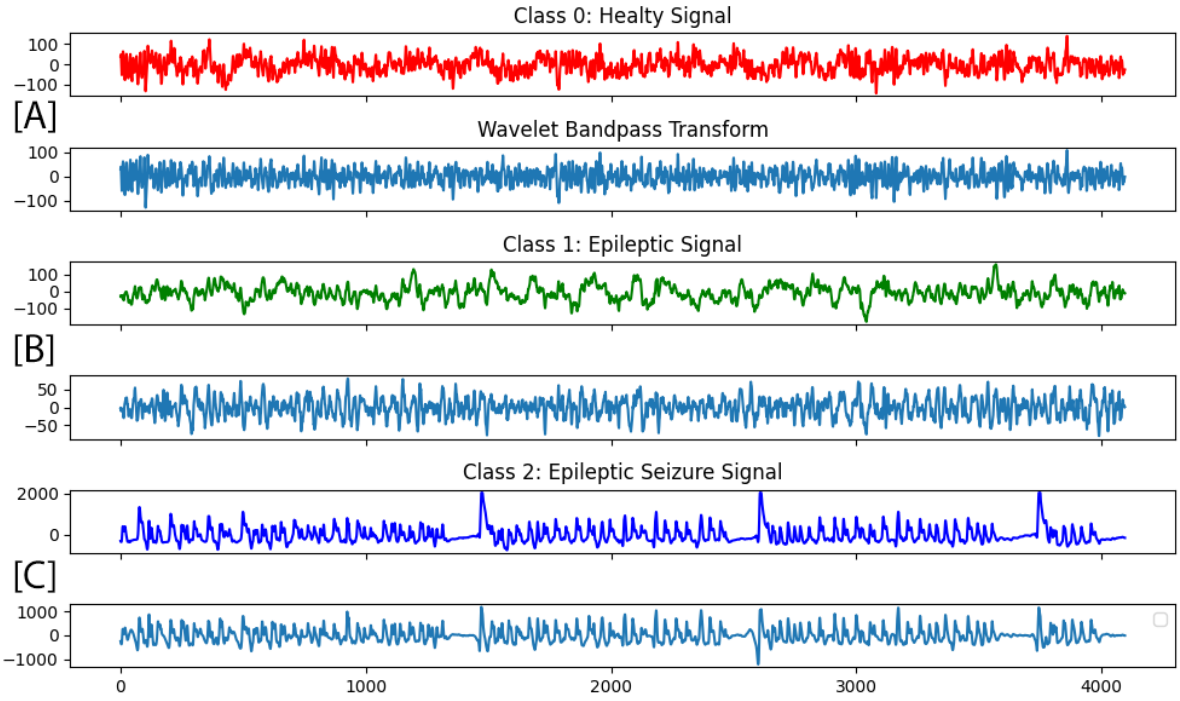


Figure 2: EEG signals of healthy [A], interictal epileptic patient [B] and ictal seizure patient [C], respectively.

The original Bonn data set contains 200 healthy, 200 epileptic interictal and 100 epileptic ictal data. Since this dataset is not sufficient for DL, we first doubled the data size by adding Gaussian noise at the signal level. Then, in the feature extraction phase with connectogram, jump parameters [1,3,5], [3,5,7] and [5,7,11] were selected and images were created for different jump options and the data size was tripled. Thus, the total data size was increased to 3000, including 1200 normal, 1200 interictal and 600 ictal images.

Graph Representation for EEG Signals

Time series data, such as EEG signals, consist of discrete real-valued data with varying amplitude ranges. Additionally, the peaks within these signals can manifest as either positive or negative alternations. This amplitude variation remains consistent across consecutive frames in the time series, thereby preserving the accuracy and coherence of graphical representations generated from each frame within the sequence.

The proposed methodology for converting time series data into graph representations processes each individual frame of the original EEG signal autonomously. In the context of the network theory approach, the amplitude values of the denoised signal, after under-

going normalization and quantization, are treated as discrete nodes within the graph. The edges between these nodes are defined by the spatial proximity of their values within a defined neighborhood window. While the method is capable of generating an arbitrarily large number of network layers, it has been specifically optimized to yield exactly three layers per frame, which are subsequently merged to produce RGB images. This framework is built upon the definition of a multilayer network model, as detailed in the subsequent sections [12].

$$\mathscr{M} = (\mathscr{G}, \mathscr{C}) \tag{1}$$

where G constitutes the layers of the network:

$$\mathscr{G} = \{G_\alpha; \alpha = 1, \cdots\cdots, M\} \tag{2}$$

also defined as a family of graphs as:

$$G_\alpha = (V_\alpha, E_\alpha) \tag{3}$$

and

$$\mathscr{C} = \{E_{\alpha\beta} \subseteq V_\alpha \times V_\beta; \alpha, \beta \in [1, 2, \cdots., M], \alpha \neq \beta\} \tag{4}$$

The approach proposed in this paper does not consider directional connections within the special layers $G\alpha$ and $G\beta$ with the condition $\alpha \neq \beta$ specified in Equation 4, which defines inter-layer connections in multilayer networks. Instead, it only establishes intra-layer connections within each layer based on the connection procedure [13,14].

EEG signal time series are composed of discrete data with varying amplitude ranges. The peak values of these signals can alternate between positive and negative. The range of fluctuations between the negative and positive peak values are normalized to the [0-1] interval. This range of variation is maintained consistently across successive frames, thereby preserving the coherence and accuracy of the graph representations generated from the frames in the time series.

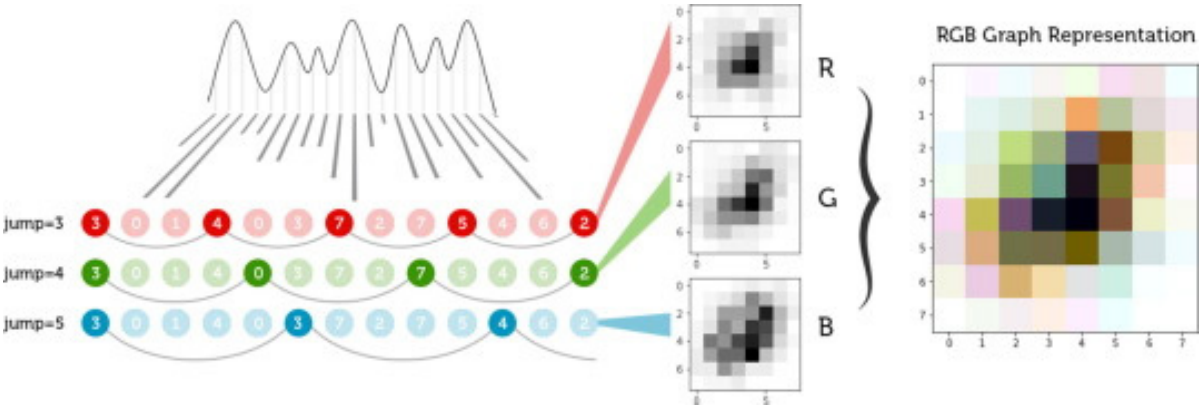


Figure 3: Time-series to graph conversion using the connectogram method [12].

The quantization procedure begins by determining a fixed bit depth, which establishes the dimensions (rows and columns) of the adjacency matrices used to construct the graph representations of the time series. The representation of the time series to graph transformation shown in Figure 3 uses a depth of 3 bits. This creates a quantization depth (a parameter called scale in the main function) of 23 = 8 [12].

The adjacent matrix representations are then compressed into a single column to create the connectogram. Thus, for this example, a connectogram of height 23 × 23 = 64 (scale-squared) is included. For example, in this study, where we keep the bit depth value small, the optimum value of this variable turns out to be 4, resulting in a 16×16 graphic representation and a connectogram height of 256 pixels.

In this study, a hopping parameter is used between consecutive EEG signals to obtain different graph results from a single EEG signal. The skip parameter in question is used as a sequence of 3 different numbers such as [1,3,5]. Thus, the RGB layers of the connectogram image are obtained with 3 different connection options. For each skip parameter applied to an EEG frame, a subsampling process proceeds by skipping intermediate nodes, as shown in Figure 3. These values are set as primes to avoid overlap between them.

Ultimately, the windowSize parameter is employed to generate the connectogram. This parameter defines the proximity required for nodes to be considered neighbors and connected. When set to 2, it indicates that the active node will be linked to its two nearest neighbors, as determined by the jump parameters. The parameter values listed in Table 1 were utilized to produce the connectogram images.

Table 1: Determined along with the values used in the tests. connectogram parameters.

Parameter	Value	Definition
Bit-Depth	4	The individual frames of the original EEG signal.
Windowsize	2	Each node is linked to the subsequent two adjacent nodes.
Sr	173.6 Hz	Sampling frequency applied to the original signal.
Jump	[1,3,5]	Jumping parameters.
Neighdiffthreshold	-1	The smallest threshold for connecting neighboring elements. A value of -1 allows nodes of equal amplitude to be connected.
Winlength	128	Width of signal frames.
Hoplength	64	The overlap between consecutive frames.

In this proposed connection procedure, a connection is established only between the amplitude nodes of the neighborhoods in the forward direction, creating a unidirectional graph. The main differences between the Visibility Graph and our proposed graph transformation method are the following:

(i) RGB images are created from adjacency matrices by establishing links with neighbors with varying subsampling rates,

(ii) The visibility criterion is ignored and all nodes within this node distance are included in the calculation.

As can be seen from these points, the graph conversion procedure in this proposed method has fundamental differences with the Visibility Graph approach. However, the time

dependency of this connection scheme is unique for the present study, as described in the method steps.

The process of converting time series into graph structures outlined earlier produces matrix-based representations for each EEG frame. In the subsequent step, these graph representations are integrated to form a time-link representation, which illustrates the temporal connectivity patterns across the entire EEG signal.

By generating RGB graph representations of consecutive overlapping EEG frames with predefined window sizes, these square graphs are transformed into single columns. These columns are then arranged side by side to create a Connectogram image, as shown in Figure 4 [12].

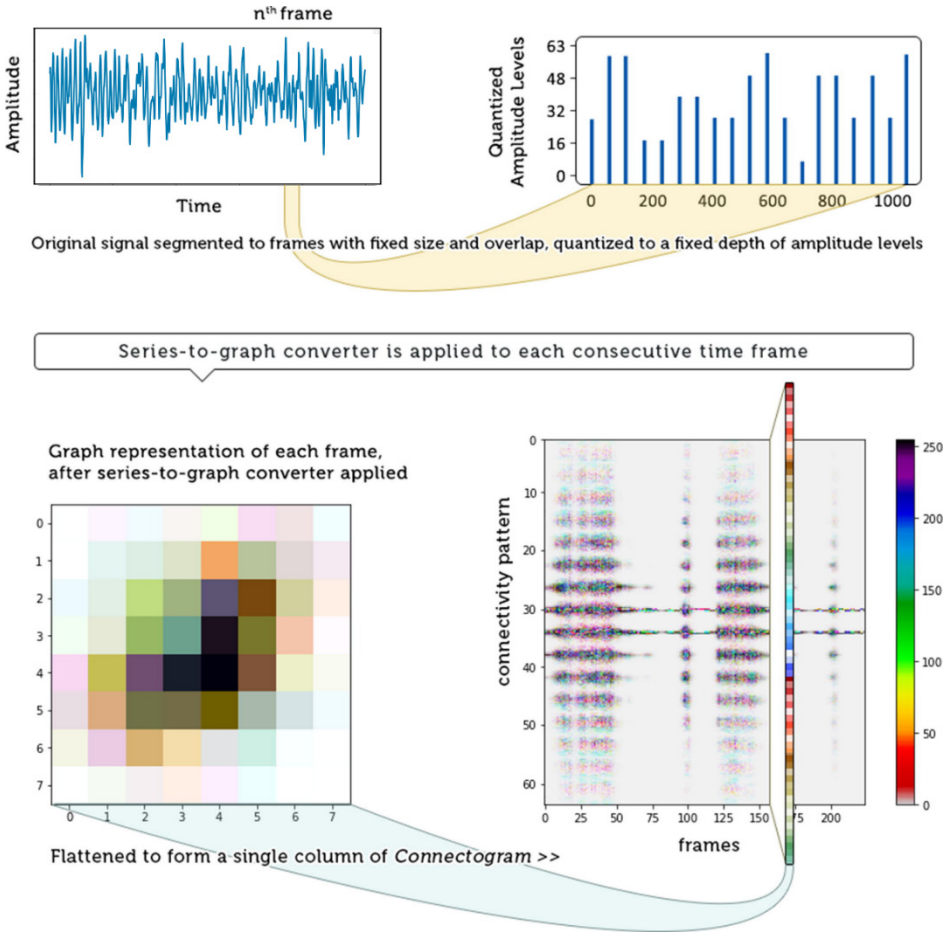


Figure 4: The process steps of the procedure for creating a Connectogram [12].

Deep Learning Model

In this study, since RGB images were created for graph representation of time series, an image-based Deep Learning (DL) model was used for the classification phase. In our study, we preferred Transfer Learning (TL) models such as **VGG**, **ResNet, Xception**, which are pre-trained DL architectures that provide results in a shorter time and better classification performance compared to their competitors [15–17].

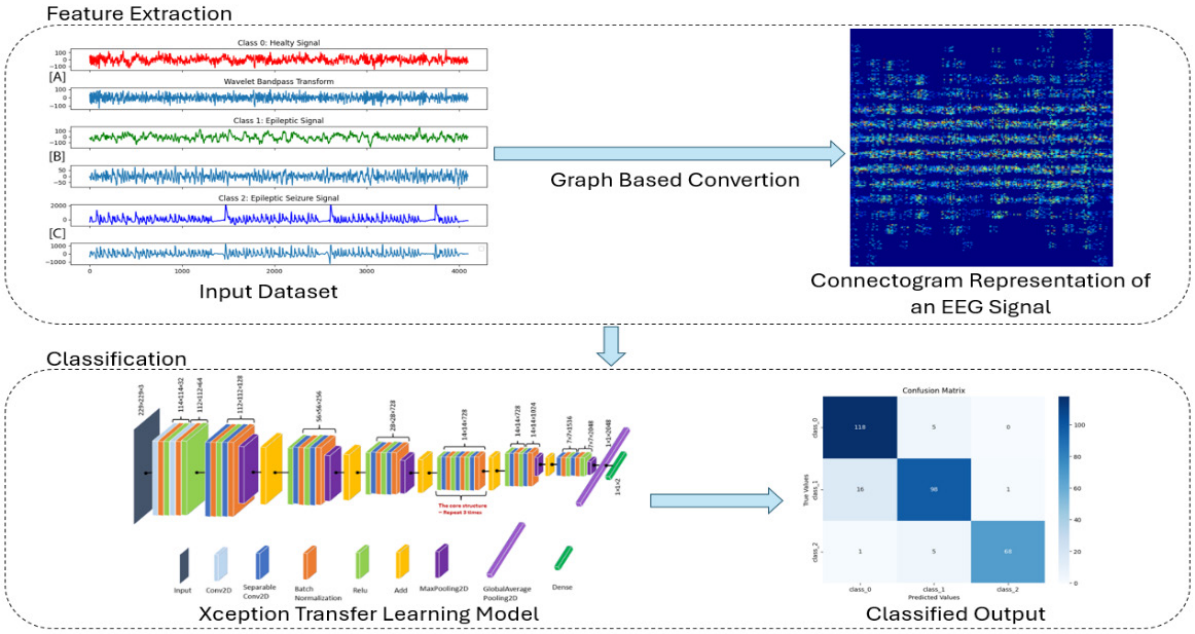
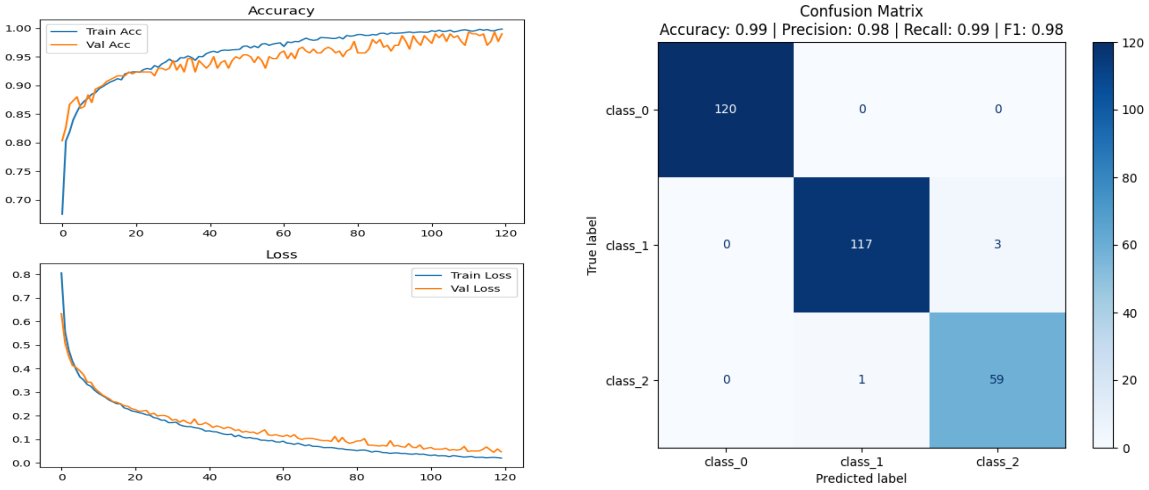


Figure 5: Connectogram-based feature extraction and classification of EEG signals [20].

Since the dataset used was not large enough for TL, data augmentation methods such as flipping, rotation, scaling, zooming, contrast, etc. were applied to avoid problems such as over-learning and insufficient generalization [18,19]. The connectogram images were adjusted to a 224x224 pixel resolution in RGB format to ensure an optimal input for the pre-trained deep learning models. The stages of extracting connectogram-based features of the EEG signals we used in this study and classifying them with an appropriate TL model are shown in Figure 5.

Results and Discussion

In this study, RGB layered images created with connectogram-based graphs were classified with the Xception TL model. The dataset used in the training was split into training/validation/test with a ratio of 80%/10%/10% respectively. The batch size and dropout were set to 16 and 0.25 respectively. During the training of the model, the validation performance was monitored, and overfitting was prevented with an early stopping mechanism.



(a) (b)

Figure 6: Classification performance graph (a) and confusion matrix values (b) on connectogram images of 3 classes obtained with Bonn dataset.

The overall accuracy of the model was 98.67% in the experiments on the test dataset consisting of 300 images, as shown in Figure 6, on data divided into three classes: healthy (class_0), epileptic interictal (class_1) and epileptic ictal seizure (class_2). This success shows that our proposed *connectogram* representation can discriminate between classes with high accuracy. Furthermore, the test results obtained for each class are shown in Table 2.

Table 2: Success values obtained from classification with Xception model.

Sütun1	precision	recall	f1-score	support
Healthy	1.00	1.00	1.00	120
Interictal	0.9915	0.9750	0.9831	120
Ictal	0.9516	0.9833	0.9672	60
Accuracy			0.9867	300
Macro Avg	0.9817	0.9861	0.9835	300
Weighted Avg	0.9873	0.9867	0.9870	300

When analyzed for each class, for class_0, the model predicted with 100% accuracy, which means that both precision, recall and F1-score values are 1.00. For class_1, precision is 99.15%, recall is 97.5% and F1-score is 98.3%. These results show that the model classifies the examples belonging to class_1 correctly for the most part and makes very few errors. On the other hand, for class_2, which has less data than the others, the model achieved 95.2% precision and 98.3% recall with an F1-score of 96.7%. This level of performance, especially for underrepresented classes such as class_2, shows that the model performs a balanced classification and has a high capacity to recognize minority classes.

The macro-averages (values calculated by taking the arithmetic mean across classes) are 98% for precision, 99% for recall and 98% for F1-score. The weighted average is similarly high, proving that the model performs consistently despite the imbalance between classes. These metrics support the generalization ability and class discrimination of the model. It can be said that the Xception-based architecture, especially the TL method, is successfully adapted in the training process and works with high accuracy on the test data.

Conclusion

In this study, we propose a new time-convexity representation for signals, which utilizes graph-based models of consecutive frames derived from segmenting the original EEG signal using fixed window and hop length parameters. The matrix representation, known as a connectogram, is composed of three layers, each created with varying sub-sampling rates, and results in an RGB image that can be used as input for classification models. From a visual perspective, the connectogram depicts the amplitude characteristics of the original EEG signal as vertical variations, with the correlation density between these variations represented by corresponding colors. As a result, frequency-related information is also integrated with the amplitude data within the connectogram.

When the results obtained in this study are examined, it is seen that connectogram images obtained from time series are suitable graph representations for DL models. In future studies, better performances can be obtained with different representation parameters by changing both window widths and jump parameters between signals.

Declaration of Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements: The authors received no financial support for this research.

Author Contributions: Serkan Aksu; Conceptualization, Methodology, Formal analysis, Supervision. Bayram Akgül; Software, Validation, Investigation, Resources. Fatih Dinç; Validation, Investigation, Resources.

REFERENCES

[1] Firpi, H., Goodman, E. D., and Echauz, J., "Epileptic seizure detection using genetically programmed artificial features", *IEEE Transactions On Biomedical Engineering*, 54 (2): 212–224 (2007).

[2] Adeli, H., Ghosh-Dastidar, S., and Dadmehr, N., "A wavelet-chaos methodology for analysis of EEGs and EEG subbands to detect seizure and epilepsy", *IEEE Transactions On Biomedical Engineering*, 54 (2): 205–211 (2007).

[3] Falsaperla, R., Vitaliti, G., Marino, S. D., Praticò, A. D., Mailo, J., Spatuzza, M., Cilio, M. R., Foti, R., and Ruggieri, M., "Graph theory in paediatric epilepsy: A systematic review", *Dialogues In Clinical Neuroscience*, 23 (1): 3–13 (2021).

[4] Zheng, M., Domanskyi, S., Piermarocchi, C., and Mias, G. I., "Visibility graph based temporal community detection with applications in biological time series", *Scientific Reports*, 11 (1): 5623 (2021).

[5] Wang, L., Long, X., Arends, J. B. A. M., and Aarts, R. M., "EEG analysis of seizure patterns using visibility graphs for detection of generalized seizures", *Journal Of Neuroscience Methods*, 290: 85–94 (2017).

[6] Chiang, S. and Haneef, Z., "Graph theory findings in the pathophysiology of temporal lobe epilepsy", *Clinical Neurophysiology*, 125 (7): 1295–1305 (2014).

[7] Bomela, W., Wang, S., Chou, C.-A., and Li, J.-S., "Real-time inference and detection of disruptive EEG networks for epileptic seizures", *Scientific Reports*, 10 (1): 8653 (2020).

[8] Jia, M., Liu, W., Duan, J., Chen, L., Chen, C. L. P., Wang, Q., and Zhou, Z., "Efficient graph convolutional networks for seizure prediction using scalp EEG", *Frontiers In Neuroscience*, 16: 967116 (2022).

[9] Smith, G. C. and Stacey, W. C., "Graph theory for EEG: can we learn to trust another black box?", *Brain*, 142 (12): 3663–3666 (2019).

[10] Kuang, Z., Liu, S., Zhao, J., Wang, L., and Li, Y., "Epilepsy EEG Seizure Prediction Based on the Combination of Graph Convolutional Neural Network Combined with Long-and Short-Term Memory Cell Network", *Applied Sciences*, 14 (24): 11569 (2024).

[11] Andrzejak, R. G., Lehnertz, K., Mormann, F., Rieke, C., David, P., and Elger, C. E., "Indications of nonlinear deterministic and finite-dimensional structures in time series of brain electrical activity: Dependence on recording region and brain state", *Physical Review E*, 64 (6): 061907 (2001).

[12] Türker, I. and Aksu, S., "Connectogram–A graph-based time dependent representation for sounds", *Applied Acoustics*, 191: 108660 (2022).

[13] Türker, İ. and Sulak, E. E., "A multilayer network analysis of hashtags in twitter via co-occurrence and semantic links", *International Journal Of Modern Physics B*, 32 (04): 1850029 (2018).

[14] Zou, Y., Donner, R. V, Marwan, N., Donges, J. F., and Kurths, J., "Complex network approaches to nonlinear time series analysis", *Physics Reports*, 787: 1–97 (2019).

[15] Sharma, S. and Guleria, K., "Deep learning models for image classification: comparison and applications", (2022).

[16] He, K., Zhang, X., Ren, S., and Sun, J., "Deep residual learning for image recognition", (2016).

[17] Chollet, F., "Xception: Deep learning with depthwise separable convolutions", (2017).

[18] Soekhoe, D., Van Der Putten, P., and Plaat, A., "On the impact of data set size in transfer learning using deep neural networks", (2016).

[19] Su, J., Yu, X., Wang, X., Wang, Z., and Chao, G., "Enhanced transfer learning with data augmentation", *Engineering Applications Of Artificial Intelligence*, 129: 107602 (2024).

[20] Mehmood, A., "Efficient anomaly detection in crowd videos using pre-trained 2D convolutional neural networks", *IEEE Access*, 9: 138283–138295 (2021).

DDOS SALDIRILARININ MAKİNE ÖĞRENMESİ YAKLAŞIMLARI İLE TESPİT EDİLMESİ

Seda Nur Şahinoğlu*
Fırat Üniversitesi, Fen Bilimleri Enstitüsü, Elazığ, Türkiye

Fatih Ertam
Fırat Üniversitesi, Adli Bilişim Mühendisliği, Elazığ, Türkiye
**(İletişimden sorumlu yazar) 241144104@firat.edu.tr*

ÖZET

İnternetin yaygın kullanımı ile birlikte siber saldırıların çeşitleri ve sayıları da artış göstermiştir. Bu saldırılardan birisi olan ve hizmetin aksamasını sağlayan dağıtılmış hizmet reddi (DDoS), yani çok sayıda istek göndererek belirli bir hedefe yapılan büyük bir saldırılar hala yaygın bir şekilde karşımıza siber tehdit olarak çıkmaktadır. Bu saldırılarda, mevcut tüm kaynakları kullanırlar ve hatta kuyruğu bloke ederler. Sonuç çoğunlukla tüm sistemin kesintiye uğraması ve hatta bloke edilmesidir. DDoS, günümüzde web'deki en yaygın saldırılardan biridir. Bu nedenle, hızlı bir tespit sistemi, otomatik engellemeyi veya bir saldırının bildirimini etkinleştirebilir. Bu çalışmada birden fazla sayıda SYN, ICMP, UDP, Slowloris, SlowHTTP ve DNS DDoS saldırıları bir laboratuvar ortamında yapılmış ve ağ akışları kayıt altına alınmıştır. Kayıt altında alınan ağ akışlarının özellikleri çıkarılarak makine öğrenmesi yaklaşımları Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Naive Bayes, Decision Tree ve LightGBM ile DDoS saldırısının türü belirlenmiştir. LightGBM algoritması ile %99 üzerinde başarım oranı ile elde edildiği görülmüştür.

Anahtar Kelimeler: DDoS, siber güvenlik, makine öğrenmesi.

Giriş

Dijital Hizmet Reddi (DDoS) saldırıları, günümüzde ağ altyapılarına yönelik en ciddi siber tehditlerden biri haline gelmiştir. Bu saldırılar, hedef sistemlerin hizmet veremez duruma gelmesine neden olarak hem bireysel hem de kurumsal ölçekte büyük zararlara yol açmaktadır[1]. Özellikle internetin yaygınlaşması ve Nesnelerin İnterneti (IoT) gibi teknolojilerin artan kullanımı, DDoS saldırılarının daha karmaşık ve tespit edilmesi zor biçimlerde gerçekleştirilmesine neden olmuştur [2]. Geleneksel güvenlik çözümleri, bu tür saldırıların dinamik yapısı karşısında yetersiz kalmakta; bu durum ise makine öğrenmesi temelli çözümlere olan ilgiyi artırmaktadır [3].

Makine öğrenmesi teknikleri, büyük boyutlu ağ trafiği verilerinde anormallikleri tespit etmede etkin bir çözüm sunmaktadır. Özellikle denetimli öğrenme yöntemleri, belirli saldırı türlerine karşı yüksek doğrulukla sonuçlar üretmektedir [4]. Yapılan bazı çalışmalarda, karar ağaçları, rastgele ormanlar (Random Forest), destek vektör makineleri (SVM) gibi

algoritmaların DDoS tespiti konusunda başarılı performanslar sergilediği gözlemlenmiştir [5]. Entropi tabanlı yaklaşımlar, SDN (Software Defined Networking) ortamlarında DDoS saldırılarının tespitinde hibrit yöntemlerle birlikte uygulanarak etkili sonuçlar vermektedir [6].

Bunun yanı sıra, ağ trafiğinden elde edilen PCAP dosyalarının çeşitli özellikleri çıkarılarak sayısal veriler elde edilmekte ve bu veriler üzerinde çeşitli sınıflandırma algoritmaları uygulanmaktadır [1]. K-En Yakın Komşu (KNN), Naive Bayes, LightGBM gibi modellerin farklı veri setleri üzerindeki başarıları birçok araştırmada karşılaştırılmıştır [2]. Bazı çalışmalarda, özellik seçimi ve veri ön işleme süreçlerinin model başarısı üzerinde önemli etkileri olduğu belirtilmiştir [3].

Bu çalışma kapsamında, farklı DDoS saldırı senaryoları oluşturularak saldırıya uğrayan makinelerde PCAP kayıtları elde edilmiş ve bu verilerden sayısal özellikler çıkarılmıştır. Elde edilen veri seti üzerinde; Random Forest, SVM, KNN, Naive Bayes, Decision Tree ve LightGBM algoritmaları eğitilmiş ve karşılaştırılmıştır. Amaç, DDoS saldırılarını en yüksek doğrulukla tespit edebilecek makine öğrenmesi yöntemini belirlemektir.

Materyal ve Metot

Bu çalışmada, farklı DDoS saldırı türlerinin ağ trafiği üzerindeki etkilerinin incelenmesi ve bu trafiğin makine öğrenmesi algoritmalarıyla analiz edilerek saldırıların tespit edilmesi amaçlanmıştır. Kali Linux üzerinde oluşturulan saldırılar Ubuntu işletim sistemiyle çalışan bir hedef sanal makineye yönlendirilmiş ve saldırı trafiği Wireshark ile kaydedilmiştir. Gerçekleştirilen DDoS saldırı türleri SYN Flood, ICMP Flood, UDP Flood, Slowloris, SlowHTTP ve DNS Flood şeklindedir. Her bir saldırı senaryosu için ayrı bir .pcap dosyası elde edilmiştir. Ayrıca, ağda herhangi bir saldırının gerçekleşmediği normal trafik verilerini içeren bir .pcap dosyası da referans olarak kullanılmıştır.

Toplanan trafik verileri daha sonra özellik çıkarımı (feature extraction) aşamasına tabi tutulmuştur. Özellik çıkarımı sonrası yalnızca sayısal veriler seçilmiş, eksik veya sonsuz değerler temizlenmiş ve uç değerler ayıklanarak veri seti makine öğrenmesi algoritmaları için hazır hale getirilmiştir. Çalışmanın temel amacı, bu veri seti üzerinde farklı makine öğrenmesi algoritmalarının performanslarını kıyaslamak ve DDoS saldırılarının en etkili şekilde tespit edilmesini sağlamaktır.

Özellik Seçimi ve Veri Ön İşleme

DDoS saldırılarının tespiti için etkili bir makine öğrenmesi modeli oluşturulmadan önce, veri setinin uygun şekilde işlenmesi ve anlamlı özelliklerin seçilmesi kritik öneme sahiptir. Bu çalışmada, öncelikle ağ trafiğinden elde edilen PCAP dosyalarından sayısal özellikler çıkarılmıştır. Özellik seçimi aşamasında, korelasyon analizi kullanılarak yüksek derecede ilişkili olmayan ve model performansına katkı sağlayan özellikler belirlenmiştir [7]. Eksik veriler, veri setinin bütünlüğünü korumak amacıyla uygun yöntemlerle doldurulmuş veya kaldırılmıştır [8]. Ayrıca, özelliklerin farklı ölçeklerde olması nedeniyle, Min-Max normalizasyonu uygulanarak tüm özellikler 0 ile 1 arasına ölçeklendirilmiştir [9]. Bu adımlar, modelin daha hızlı ve doğru öğrenmesini sağlamak için gereklidir.

Veri Seti Tanımı

Deneysel çalışmalarda kullanılan veri seti, Kali Linux sanal makinesi üzerinden Ubuntu sanal makinesine gerçekleştirilen farklı DDoS saldırıları sonucunda toplanmıştır. Gerçekleştirilen saldırı türleri arasında SYN, ICMP, UDP, Slowloris, SlowHTTP ve DNS saldırıları bulunmaktadır. Her bir saldırı türü için ayrı PCAP dosyaları Wireshark kullanılarak kaydedilmiş ve ayrıca saldırı içermeyen normal trafik de toplanmıştır. Bu veri seti, hem saldırı hem de normal trafik örneklerini içermesi nedeniyle, makine öğrenmesi modellerinin eğitimi ve test edilmesi için dengeli bir yapı sunmaktadır [10] .

Makine Öğrenmesi Modelleri

DDoS saldırılarının tespiti amacıyla literatürde yaygın olarak kullanılan ve farklı çalışma prensiplerine sahip çeşitli makine öğrenmesi algoritmaları değerlendirilmiştir. Bu çalışmada kullanılan modeller ve özellikleri aşağıda belirtilmiştir:

•Random Forest – RF: Birden fazla karar ağacının oluşturduğu bir topluluk yöntemidir. Her bir ağaç, rastgele seçilen özellikler ve örneklerle eğitilir, bu da modelin genelleme kabiliyetini artırır [7].

•Support Vector Machines – SVM: Veriyi yüksek boyutlu bir uzaya haritalayarak, farklı sınıfları en iyi ayıran hiper düzlemi bulmayı amaçlayan bir sınıflandırma yöntemidir. Özellikle lineer olarak ayrılabilir verilerde yüksek performans gösterir [11] .

•K-Nearest Neighbors – KNN: Bir örneğin sınıfını, en yakın komşularının sınıflarına dayanarak belirleyen bir yöntemdir. Basitliği ve yorumlanabilirliği nedeniyle tercih edilir [12].

•Naive Bayes – NB: Bayes teoremine dayanan ve özelliklerin birbirinden bağımsız olduğunu varsayan bir olasılık tabanlı sınıflandırma algoritmasıdır. Hızlı ve verimli olmasıyla bilinir [13].

•Decision Tree – DT: Veriyi, özelliklerine göre dallara ayırarak sınıflandıran hiyerarşik bir yapıdır. Kolay yorumlanabilir olması avantaj sağlar [10].

•Light Gradient Boosting Machine – LightGBM: Büyük veri setlerinde hızlı ve yüksek performanslı sonuçlar üreten, gradient boosting çerçevesine dayanan bir algoritmadır. Özellikle büyük ölçekli uygulamalarda tercih edilir.

Her bir model, belirlenen özellikler ve ön işleme adımlarıyla eğitilmiş ve test edilmiştir. Modellerin performansı, doğruluk, kesinlik, duyarlılık ve F1 skoru gibi metriklerle değerlendirilmiştir. Sonuçlar, LightGBM ve Random Forest modellerinin diğer modellere kıyasla daha yüksek doğruluk oranlarına sahip olduğunu göstermiştir.

Yöntem

Bu çalışmada, DDoS saldırılarını tespit etmek amacıyla çeşitli makine öğrenmesi algoritmaları değerlendirilmiştir. Python dili kullanılarak gerçekleştirilen bu analizde, CSE-CIC-IDS2018 ve CICDDoS2019 gibi güncel ve kapsamlı ağ trafiği veri setleri birleştirilmiş, sayısal özellikler çıkarılarak eksik ve aykırı değerler temizlenmiştir. Bu işlemlerin ardından veriler eğitim ve test kümelerine ayrılarak farklı sınıflandırma algoritmalarıyla analiz edilmiştir [14].

Tablo-1' de, yöntem sürecinin genel bir ÖZETini sunan bir akış yer almaktadır:

Tablo 1: Yöntem Aşamaları

No	Aşama Adı	Açıklama
1	Sanal Makinelerin Kurulumu	Kali Linux üzerinden Ubuntu'ya yönelik DDoS saldırıları gerçekleştirildi.
2	Saldırı Türlerinin Gerçekleştirilmesi	SYN, ICMP, UDP, Slowloris, SlowHTTP, DNS gibi saldırılar uygulandı.
3	Ağ Trafiğinin Dinlenmesi	Ubuntu'daki Wireshark ile saldırı ve normal trafikler dinlendi.
4	PCAP Dosyalarının Elde Edilmesi	Her saldırı türü için ayrı PCAP dosyaları oluşturuldu.
5	Özellik Çıkartımı	PCAP dosyalarından sayısal özellikler (feature) çıkarıldı.
6	Veri Temizleme ve Dönüştürme	Aykırı ve eksik veriler silindi, sonsuz değerler düzeltildi.
7	Özellik Normalizasyonu	Sayısal veriler ölçeklendirildi (standardizasyon/ normalizasyon).
8	Veri Setinin Ayrılması	Veriler eğitim (%60) ve test (%40) olarak ayrıldı.
9	Makine Öğrenmesi Modellerinin Uygulanması	RF, SVM, KNN, Naive Bayes, Decision Tree ve LightGBM modelleri eğitildi.
10	Performans Değerlendirmesi	Doğruluk oranı, hassasiyet, özgüllük, F1 skoru gibi metriklerle değerlendirme yapıldı.
11	Karışıklık Matrisi Analizi	Modellerin sınıflandırma başarıları matriks üzerinden incelendi.
12	Sonuçların Görselleştirilmesi	Matplotlib/Seaborn kullanılarak grafiksel ÖZETler üretildi.
13	Yorumlama ve Karşılaştırma	Modellerin güçlü/zayıf yönleri tartışıldı ve uygun model önerildi.

Veri Setine İlişkin Detaylar

Bu çalışmada kullanılan veri seti, yedi farklı DDoS saldırı türü ile saldırı içermeyen normal trafik verilerinden oluşturulmuştur. Veri seti, deneysel olarak gerçekleştirilen saldırı senaryoları sonucunda elde edilen ağ trafiği üzerinden oluşturulmuş ve her bir saldırı türü için ayrı ayrı PCAP dosyaları üretilmiştir. Bu ham veriler CICFlowMeter aracı kullanılarak sayısal değerlere dönüştürülmüş ve makine öğrenmesi algoritmaları ile kullanılabilir hale getirilmiştir.

Veri kümesinde toplam **6756 adet trafik akışı (flow)** yer almaktadır. Bu akışlar, aşağıda listelenen yedi sınıf altında gruplanmıştır:

Tablo-2'de, analizde kullanılan veri setindeki saldırı türlerinin sınıf dağılımı ve her bir sınıfa ait kayıt sayıları sunulmuştur.

Tablo 2: Sınıf Dağılımı

Sınıf No	Saldırı Türü	Kayıt Sayısı
1	SYN Flood	38
2	ICMP Flood	456
3	UDP Flood	1871
4	Normal Trafik	435
5	Slowloris	205
6	SlowHTTP	1511
7	DNS Flood	1240

Veri setinde yer alan özellikler arasında kaynak ve hedef IP adresleri, protokol tipi,

bağlantı süresi, paket boyutu, bayt transferi, zaman damgaları, akış yönü, TTL değeri ve bağlantı sayısı gibi ağ trafiğine özgü öznitelikler bulunmaktadır. Bu özellikler, saldırı davranışlarının sınıflandırılmasında önemli rol oynamaktadır.

Veri ön işleme aşamasında gereksiz sütunlar kaldırılmış, eksik değerler kontrol edilerek ayıklanmış ve veriler normalize edilerek makine öğrenmesi algoritmalarına uygun hale getirilmiştir.

Değerlendirme Ölçütleri

Model performansını değerlendirmek için doğruluk (accuracy), kesinlik (precision), geri çağırma (recall) ve F1-skora dayalı metrikler kullanılmıştır. Bu ölçütler, özellikle dengesiz veri kümelerinde modelin başarısını anlamada önemli rol oynamaktadır [15]. Doğruluk, tüm doğru tahminlerin toplam tahmin sayısına oranıdır ve genel başarıyı gösterir. Kesinlik, modelin olumlu olarak sınıflandırdığı örneklerden kaçının gerçekten olumlu olduğunu gösterirken; geri çağırma, gerçek olumlu örneklerin ne kadarının doğru tahmin edildiğini ölçer [9].

Bu çalışmada ayrıca, sınıflandırıcıların güvenilirliğini karşılaştırmak için ROC-AUC değerleri de hesaplanmıştır. Bu yaklaşım, özellikle saldırıların tespitinde yalancı olumlu oranlarının etkisini değerlendirmek açısından yararlıdır [13].

Model Performans Karşılaştırması

Veri kümesi üzerinde uygulanan algoritmalar şunlardır: Random Forest, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), Naive Bayes, Decision Tree ve LightGBM. Bu algoritmalar, eğitim ve test kümelerinde doğruluk oranları ve sınıflandırma raporları ile değerlendirilmiştir [16].

LightGBM, doğruluk açısından en iyi performansı göstermiştir. Bu modelin başarısı, yüksek hızda öğrenme sağlaması ve büyük veri kümeleriyle verimli çalışabilmesinden kaynaklanmaktadır [17]. Random Forest da benzer şekilde yüksek doğruluk göstermiş, ancak işlem süresi bakımından LightGBM kadar avantajlı olmamıştır [18]. Diğer yandan, Naive Bayes modeli veri setinin çok boyutlu yapısı ve özellik bağımsızlığı varsayımı nedeniyle en düşük başarıyı sergilemiştir [19].

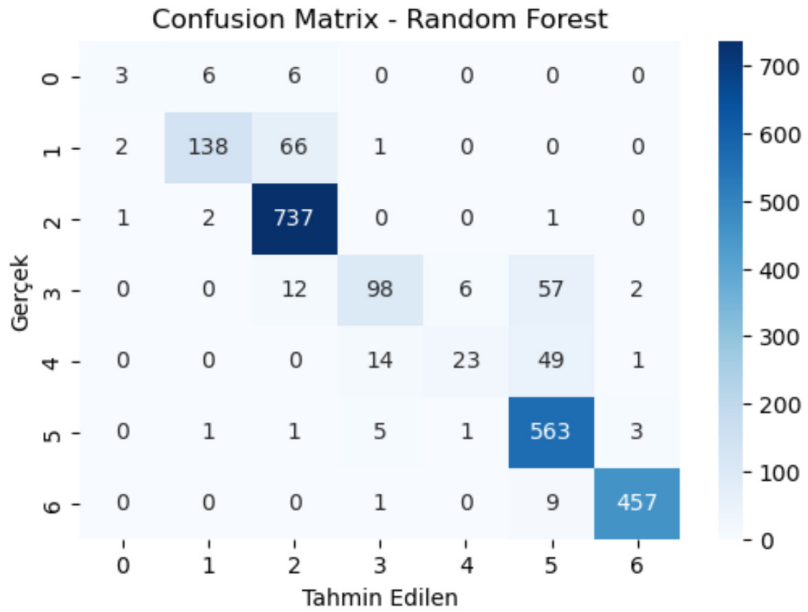
Karmaşıklık Matrisi Analizi

Karmaşıklık matrisleri, modellerin sınıflandırma performansını daha ayrıntılı analiz etmek amacıyla kullanılmıştır. Bu matrisler sayesinde, her modelin yanlış pozitif (FP) ve yanlış negatif (FN) tahmin oranları belirlenmiş, sınıflandırma hatalarının doğası anlaşılmıştır [20]. Örneğin, SVM ve LightGBM modelleri düşük FN değerleriyle dikkat çekerken, KNN modeli yüksek FP oranına sahip olmuştur. Bu bulgular, saldırı tespitinde hangi modelin daha etkili olduğunu ortaya koymak açısından kritiktir.

Çalışmada kullanılan veri kümesi, farklı DDoS saldırı türlerini içerecek şekilde etiketlenmiştir. Bu etiketleme doğrultusunda; 1 numaralı sınıf SYN Flood saldırılarını, 2 numaralı sınıf ICMP tabanlı saldırıları, 3 numaralı sınıf UDP Flood saldırılarını, 4 numaralı sınıf normal trafiği (herhangi bir saldırı içermeyen durumları), 5 numaralı sınıf Slowloris saldırılarını, 6 numaralı sınıf SlowHTTP saldırılarını ve 7 numaralı sınıf DNS tabanlı DDoS saldırılarını temsil etmektedir.

Tablo 3: Random Forest Algoritmasına ait performans metrikleri

DDos Tipi	Precision	Recall	F1-score
1	0,50	0,20	0,29
2	0,94	0,67	0,78
3	0,90	0,99	0,94
4	0,82	0,56	0,67
5	0,77	0,26	0,39
6	0,83	0,98	0,90
7	0,99	0,98	0,98



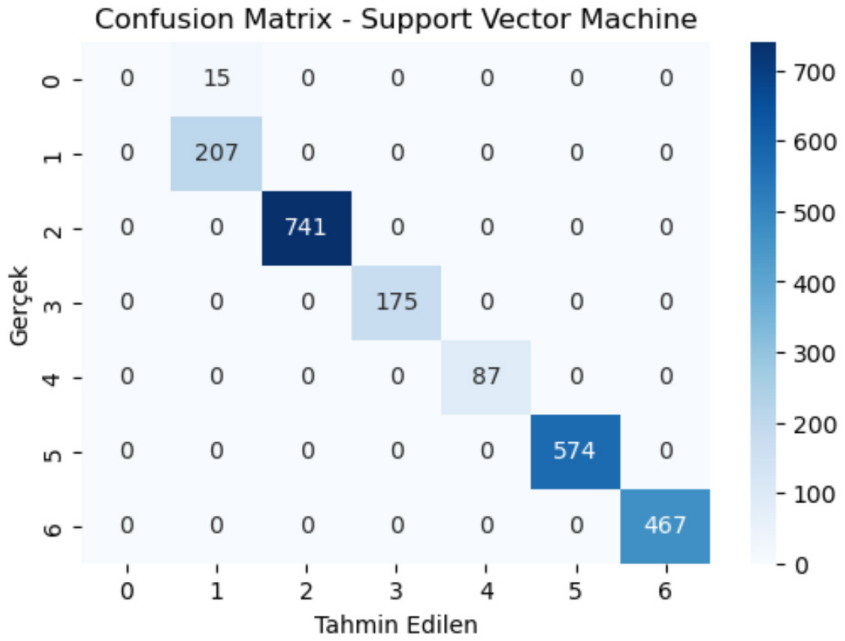
Şekil 1: Random Forest algoritmasına ait karmaşıklık matrisi

Şekil-1'de Random Forest algoritmasına ait karmaşıklık matrisi, her bir DDoS saldırı türünün doğru ve yanlış sınıflandırılma durumlarını görselleştirmektedir.

Tablo-3' de Random Forest algoritması için saldırı türlerine göre hesaplanan Precision, Recall ve F1-score metrikleri ile genel doğruluk (Accuracy) değerini göstermektedir.

Tablo 4: Support Vector Machine Algoritmasına ait performans metrikleri

DDos Tipi	Precision	Recall	F1-score
1	0,00	0,00	0,00
2	0,93	1,00	0,97
3	1,00	1,00	1,00
4	1,00	1,00	1,00
5	1,00	1,00	1,00
6	1,00	1,00	1,00
7	1,00	1,00	1,00



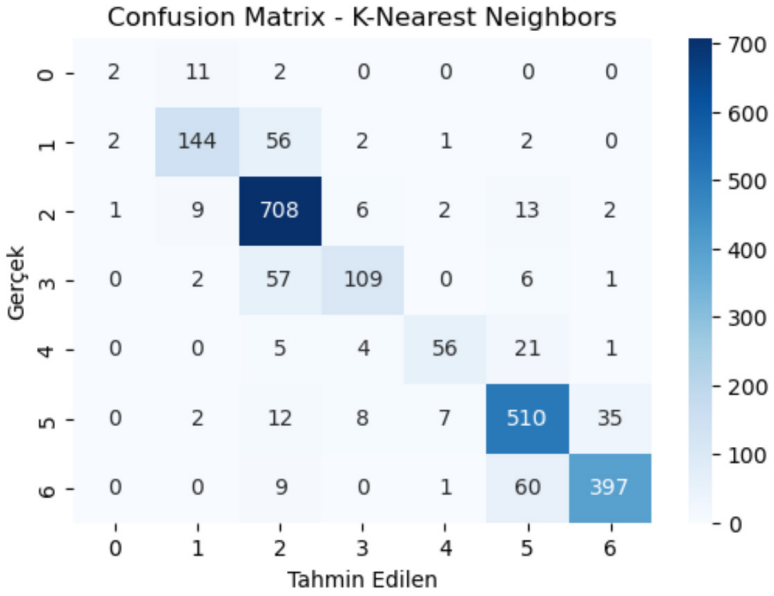
Şekil 2: Support Vector algoritmasına ait karmaşıklık matrisi

Şekil-2'de Support Vector Machine algoritmasına ait karmaşıklık matrisi, her bir DDoS saldırı türünün doğru ve yanlış sınıflandırılma durumlarını görselleştirmektedir.

Tablo-4' de Support Vector Machine algoritması için saldırı türlerine göre hesaplanan Precision, Recall ve F1-score metrikleri ile genel doğruluk (Accuracy) değerini göstermektedir.

Tablo 5: K-Nearest Neighbors Algoritmasına ait performans metrikleri

DDos Tipi	Precision	Recall	F1-score
1	0,40	0,13	0,20
2	0,86	0,70	0,77
3	0,83	0,96	0,89
4	0,84	0,62	0,72
5	0,84	0,64	0,73
6	0,83	0,89	0,86
7	0,91	0,85	0,88



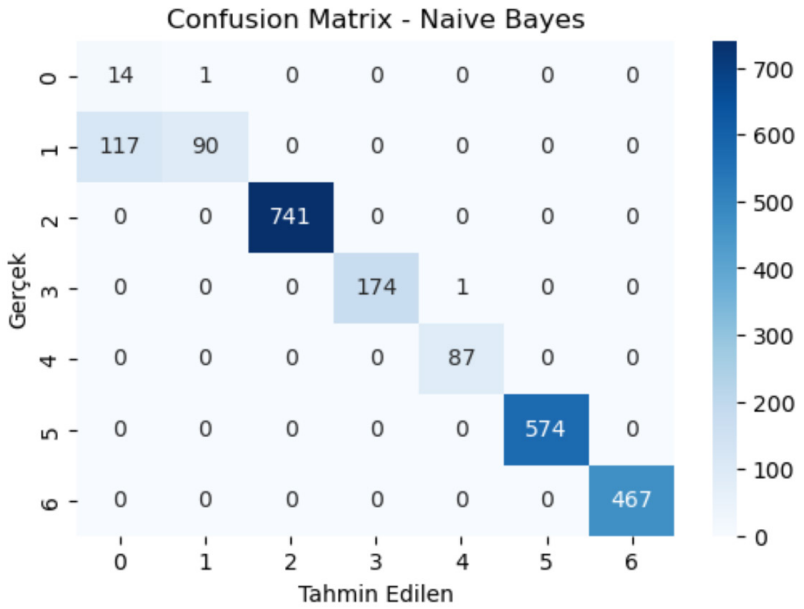
Şekil 3: K-Nearest Neigbors algoritmasına ait karmaşıklık matrisi

Şekil-3'de K-Nearest Neigbors algoritmasına ait karmaşıklık matrisi, her bir DDoS saldırı türünün doğru ve yanlış sınıflandırılma durumlarını görselleştirmektedir.

Tablo-5' de K-Nearest Neigbors algoritması için saldırı türlerine göre hesaplanan Precision, Recall ve F1-score metrikleri ile genel doğruluk (Accuracy) değerini göstermektedir.

Tablo 6: Naive Bayes Algoritmasına ait performans metrikleri

DDos Tipi	Precision	Recall	F1-score
1	0,11	0,93	0,19
2	0,99	0,43	0,60
3	1,00	1,00	1,00
4	1,00	0,99	1,00
5	0,99	1,00	0,99
6	1,00	1,00	1,00
7	1,00	1,00	1,00



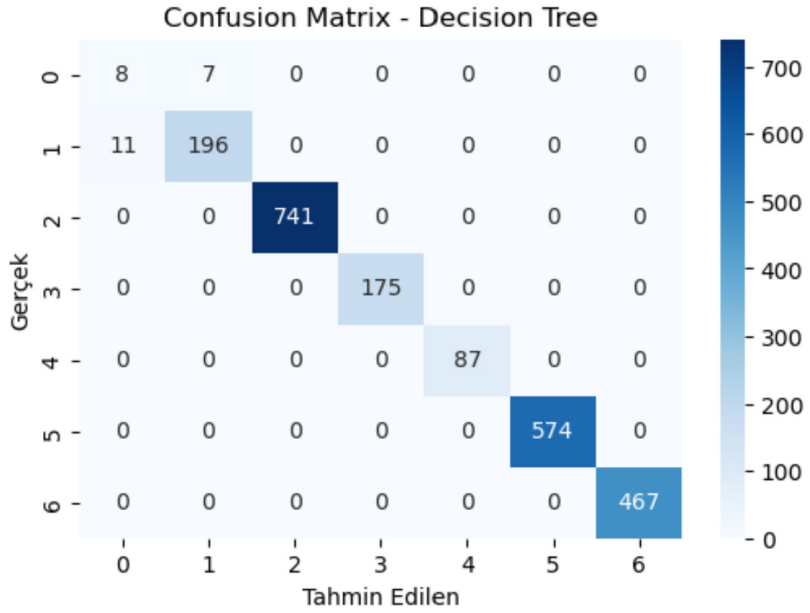
Şekil 4: Naive Bayes algoritmasına ait karmaşıklık matrisi

Şekil-4'de Naive Bayes algoritmasına ait karmaşıklık matrisi, her bir DDoS saldırı türünün doğru ve yanlış sınıflandırılma durumlarını görselleştirmektedir.

Tablo-6' de Naive Bayes algoritması için saldırı türlerine göre hesaplanan Precision, Recall ve F1-score metrikleri ile genel doğruluk (Accuracy) değerini göstermektedir.

Tablo 7: Decision Tree Algoritmasına ait performans metrikleri

DDos Tipi	Precision	Recall	F1-score
1	0,42	0,53	0,47
2	0,97	0,95	0,96
3	1,00	1,00	1,00
4	1,00	1,00	1,00
5	1,00	1,00	1,00
6	1,00	1,00	1,00
7	1,00	1,00	1,00



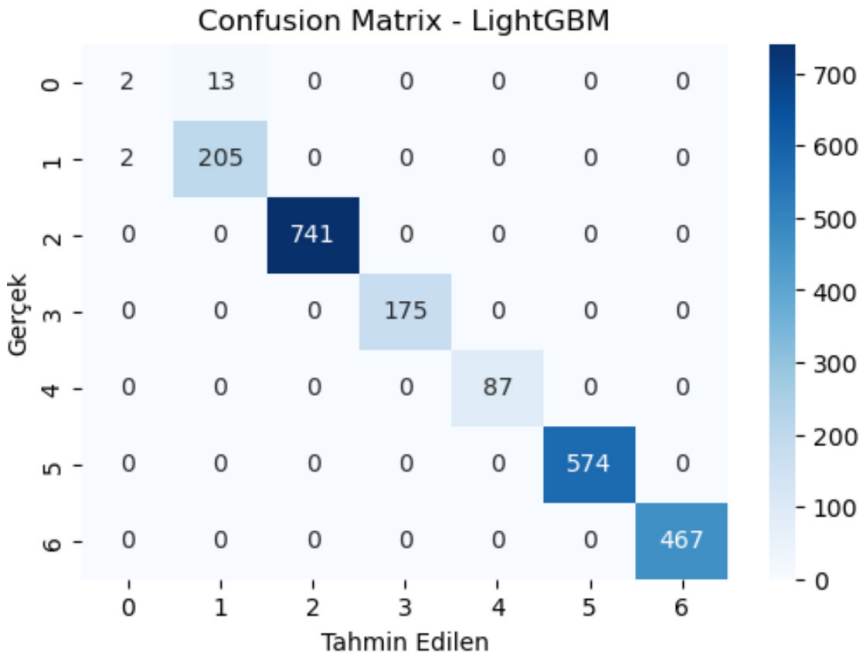
Şekil 5: Decision Tree algoritmasına ait karmaşıklık matrisi

Şekil-5'de Decision Tree algoritmasına ait karmaşıklık matrisi, her bir DDoS saldırı türünün doğru ve yanlış sınıflandırılma durumlarını görselleştirmektedir.

Tablo-7' de Decision Tree algoritması için saldırı türlerine göre hesaplanan Precision, Recall ve F1-score metrikleri ile genel doğruluk (Accuracy) değerini göstermektedir.

Tablo 8: LightGBM Algoritmasına ait performans metrikleri

DDos Tipi	Precision	Recall	F1-score
1	0,50	0,13	0,21
2	0,94	0,99	0,96
3	1,00	1,00	1,00
4	1,00	1,00	1,00
5	1,00	1,00	1,00
6	1,00	1,00	1,00
7	1,00	1,00	1,00



Şekil 6: LightGBM algoritmasına ait karmaşıklık matrisi

Şekil-6'de LightGBM algoritmasına ait karmaşıklık matrisi, her bir DDoS saldırı türünün doğru ve yanlış sınıflandırılma durumlarını görselleştirmektedir.

Tablo-8' de LightGBM algoritması için saldırı türlerine göre hesaplanan Precision, Recall ve F1-score metrikleri ile genel doğruluk (Accuracy) değerini göstermektedir.

Değerlendirme

Bu bölümde, çalışmada kullanılan makine öğrenmesi modellerinin başarı düzeyleri karşılaştırmalı olarak değerlendirilmektedir. Her bir modelin doğruluk (accuracy) oranlarının yanı sıra, dengesiz veri kümelerinde daha anlamlı sonuçlar veren kesinlik (precision), geri çağırma (recall) ve F1-skora dayalı performans metrikleri de incelenmiştir. Böylece, modellerin yalnızca genel doğruluk değil, aynı zamanda her bir saldırı türünü doğru şekilde tanımlama yetkinlikleri de ortaya konulmuştur. Ayrıca, karmaşıklık matrisleri yardımıyla yanlış sınıflandırma eğilimleri analiz edilerek modellerin güçlü ve zayıf yönleri karşılaştırılmıştır.

Uygulanan algoritmalar arasında Support Vector Machine (SVM) ve LightGBM modelleri, %99.33 doğruluk oranı ile en yüksek başarıyı göstermiştir. Bu modelleri sırasıyla Decision Tree (%99.20), Naive Bayes (%94.74) ve Random Forest (%89.09) izlemiştir. K-Nearest Neighbors algoritması ise %84.99 ile en düşük doğruluk oranına sahip model olarak belirlenmiştir. Bununla birlikte, model başarılarının sadece doğruluk oranı ile değerlendirilmesinin yanı sıra, Precision, Recall ve F1-score gibi detaylı metriklerle analiz edilmesi, özellikle dengesiz veri kümelerinde önemli farklılıkları ortaya koymuştur.

Tablo-9. Makine Öğrenmesi Modellerine Göre Sınıflandırma Başarı Oranları

	Model	Accuracy
0	Random Forest	0.890997
1	Support Vector Machine	0.993380
2	K-Nearest Neighbors	0.849956
3	Naive Bayes	0.947485
4	Decision Tree	0.992056
5	LightGBM	0.993380

Random Forest algoritması özelinde yapılan karmaşıklık matrisi ve performans metrikleri analizi, bu modelin bazı saldırı türlerinde (örneğin UDP ve DNS) oldukça başarılı olduğunu, ancak özellikle SYN ve Slowloris gibi saldırılarda hatalı sınıflandırmalara daha açık olduğunu göstermiştir. Bu durum, modelin tüm saldırı türleri için homojen bir başarı sağlayamadığını, bazı saldırıların ayırt edilmesinde zorluk yaşadığını ortaya koymuştur.

Genel olarak değerlendirildiğinde, makine öğrenmesi temelli yöntemlerin DDoS saldırılarının tespitinde güçlü araçlar sunduğu, ancak saldırı türlerinin çeşitliliği ve trafik yapılarının farklılığı nedeniyle model performanslarının değişkenlik gösterdiği anlaşılmaktadır. Bu nedenle, gelecekte yapılacak çalışmalarda, model başarısını artırmak adına daha dengeli veri kümeleri oluşturulması, derin öğrenme yöntemlerinin entegrasyonu ve özellik mühendisliği süreçlerinin daha derinlemesine ele alınması önerilmektedir.

Sonuçlar ve Tartışma

Bu çalışmada, farklı DDoS saldırı türlerinin tespitine yönelik olarak çeşitli makine öğrenmesi algoritmaları kullanılarak kapsamlı bir analiz gerçekleştirilmiştir. SYN, ICMP, UDP, DNS, Slowloris ve SlowHTTP gibi farklı protokollere dayalı saldırı türlerinin yanı sıra, saldırı gerçekleşmeyen normal trafik verisi de dikkate alınarak yedi farklı trafik türü sınıflandırılmıştır. Bu çeşitlilik, modellerin yalnızca saldırı tespiti değil, saldırı türünü ayırt etme yeteneklerini de değerlendirme imkânı sağlamıştır.

Gerçekleştirilen sınıflandırma deneylerinde, LightGBM ve Support Vector Machine algoritmaları %99.33 doğruluk oranıyla en yüksek başarıyı göstermiştir. Bu modelleri sırasıyla Decision Tree (%99.20), Naive Bayes (%94.74) ve Random Forest (%89.09) takip etmiştir. K-En Yakın Komşu algoritması ise %84.99 doğrulukla en düşük başarıyı göstermiştir. Bu sonuçlar, özellikle ağa dayalı saldırılarda yüksek ayırt ediciliğe sahip algoritmaların seçilmesinin, tespit başarısını doğrudan etkilediğini göstermektedir.

Random Forest algoritması için elde edilen karmaşıklık matrisi ve sınıflandırma metrikleri detaylı olarak analiz edildiğinde, modelin UDP ve DNS saldırılarını yüksek doğrulukla sınıflandırabildiği, ancak SYN ve Slowloris gibi bazı saldırı türlerinde ayırt ediciliğinin görece daha düşük olduğu gözlemlenmiştir. Bu durum, özellikle benzer trafik karakteristiğine sahip saldırı türlerinde sınıflandırma zorluğunun devam ettiğini göstermektedir. F1-score değerlerinin sınıf bazlı değişkenliği de bu sonucu desteklemektedir.

Tüm bulgular birlikte değerlendirildiğinde, makine öğrenmesi temelli modellerin DDoS saldırılarının tespitinde etkili ve güvenilir araçlar sunduğu, ancak model seçiminde saldırı türlerine özel performans farklılıklarının dikkate alınması gerektiği ortaya konmuştur. Bu kapsamda, LightGBM ve SVM gibi yüksek doğruluk sunan modellerin, gerçek zamanlı sistemlerde öncelikli olarak tercih edilebileceği sonucuna ulaşılmıştır.

REFERANSLAR

[1] Hacıbeyoglu, M., Arıcı, F., & Karaaltun, M. (2024). Intrusion detection system application with machine learning. Afyon Kocatepe University Journal of Sciences and Engineering, 24(5), 1165–1179.

[2] Özdemir, B., & Bilgisi, M. (2023). Siber tehdit istihbaratında yapay zeka ve makine öğrenmesi / Artificial intelligence and machine learning in cyber threat intelligence. The Journal of Defence and Security Research, 14(3), 136–143.

[3] Yalçın, N., Çakır, S., & Ünaldı, S. (2024). Çeşitli tekrarlayan sinir ağları kullanarak siber saldırı tespiti / Cyber-attack detection using various recurrent neural networks. Karaelmas Fen ve Mühendislik Dergisi, 14(3), 136–143.

[4] Varalakshmi, I., & Thenmozhi, M. (2025). Entropy based earlier detection and mitigation of DDoS attack using stochastic method in SDN_IoT. Measurement: Sensors, 101873.

[5] Yıldırım, A. (2024). Türkiye'nin siber güvenlik politikasının eylem planları üzerinden analizi. Organizasyon ve Yönetim Bilimleri Dergisi, 16(1), 23-47.

[6] Hassan, A. I., El Reheem, E. A., & Guirguis, S. K. (2024). An entropy and machine learning based approach for DDoS attacks detection in software defined networks. Scientific Reports, 14(1), 18159.

[7] Rudro, R. A. M., Sohan, M. F. A. A., Chaity, S. K., & Reya, R. I. (2023). Enhancing DDoS attack detection using machine learning: A framework with feature selection and comparative analysis of algorithms. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 14(3), 1185–1192.

[8] Raza, M. S., Sheikh, M. N. A., Hwang, I.-S., & Ab-Rahman, M. S. (2024). Feature- selection-based DDoS attack detection using AI algorithms. Telecom, 5(2), 333–346.

[9] Tymoshchuk, D., Yasniy, O., Mytnyk, M., Zagorodna, N., & Tymoshchuk, V. (2024). Detection and classification of DDoS flooding attacks by machine learning method. arXiv preprint arXiv:2412.18990.

[10] Tay, W.-W., Chong, S., & Chong, L.-Y. (2024). DDoS attack detection with machine learning. Journal of Informatics and Web Engineering, 3(3), 190–207.

[11] Rahman, M. A. (2025). Detection of distributed denial of service attacks based on machine learning algorithms. arXiv preprint arXiv:2502.00975.

[12] Borah, R., Sarmah, S., Choudhury, N., Mahanta, H., & Chodhury, A. (2023). DDoS attack detection using machine learning techniques. Indian Journal of Science and Technology, 16(SP2), 76–82.

[13] Hassan, A. I., El Reheem, E. A., & Guirguis, S. K. (2024). An entropy and machine learning based approach for DDoS attacks detection in software defined networks. Scientific Reports, 14(1), 18159.

[14] Mishra, A. (2024). Utilizing Random Forest for DDoS Attack Detection. International Journal on Recent and Innovation Trends in Computing and Communication, 12(2), 127–130.

[15] Suvra, D. K. (2025). An efficient real time DDoS detection model using machine learning algorithms. arXiv preprint arXiv:2501.14311.

[16] Ma, R., Wang, Q., Bu, X., & Chen, X. (2023). Real-time detection of DDoS attacks based on random forest in SDN. Applied Sciences, 13(13), 7872.

[17] Airlangga, G. (2024). Comparative analysis of machine learning models for intrusion detection in Internet of Things networks using the RT-IoT2022 dataset. MALCOM: Indonesian Journal of Machine Learning and Computer Science, 4(2), 656–662.

[18] Manjunath C R, Rathor K, Kulkarni N, Patil P. P., Patil M. S., & Singh J. (2022). Cloud Based DDoS Attack Detection Using Machine Learning Architectures: Understanding the Potential for Scientific Applications. International Journal of Intelligent Systems and Applications in Engineering, 10(2s), 268–271.

[19] Suresh, M., & Anitha, R. (2011). Evaluating machine learning algorithms for detecting DDoS attacks. In Advances in Network Security and Applications (pp. 441–452). Springer.

[20] Kayakuş, M., & Yiğit Açıkgöz, F. (2023). Twitter'da makine öğrenmesi yöntemleriyle sahte haber tespiti. Abant Sosyal Bilimler Dergisi, 23(2), 1017–1027.

DEEP LEARNING FOR CARDIOMYOPATHY DIAGNOSIS USING CARDIAC MRI: A COMPARATIVE ANALYSIS OF DROPOUT AND OPTIMIZATION TECHNIQUES ON CLASSIFICATION PERFORMANCE

Yeliz Alper

Zonguldak Bülent Ecevit University, Department of Biomedical Engineering, Zonguldak, Türkiye

Ceren Kaya*

Zonguldak Bülent Ecevit University, Department of Biomedical Engineering, Zonguldak, Türkiye

**(Corresponding author) ceren.kaya@beun.edu.tr, crnkaya@hotmail.com*

ABSTRACT

Cardiovascular diseases are one of the leading disease groups worldwide. Among heart-related diseases, cardiomyopathy is a class of disease that significantly affects heart muscle. The most well-known types are dilated cardiomyopathy (DCM) and hypertrophic cardiomyopathy (HCM). DCM is characterized by muscle weakness, while HCM is characterized by wall thickening. Symptoms of these diseases include arrhythmia, sudden fainting, and especially heart failure. If symptoms are severely intense and long-lasting, it is very important to consult a health professional. In their diagnosis, electrocardiography (ECG), echocardiography (ECHO), and cardiac magnetic resonance imaging (CMR) are commonly used. Computer-aided diagnosis systems are clinically of great importance for making a diagnosis of these diseases in a shorter time and preventing them. In this proposed study, classification of cardiomyopathy was performed using only raw CMR images based on model architectures such as ResNet18 and ResNet101 deep learning (CNN), without any segmentation process for heart. The effects of different optimization algorithms (Adam, AdamW, Adamax, SGD) and dropout regularization in various combinations on model performance were comprehensively examined and presented comparatively. As a result of the experiments, the best result was obtained with 94% accuracy and 0.13 loss using the combination of ResNet18 model + Adamax optimizer + 0.2% dropout. In this context, it is clearly seen that proper analysis of the dataset structure, visual preprocessing, data augmentation techniques for limited datasets, optimization algorithms, and the dropout technique significantly improve classification performance in deep learning models. Unlike most studies in the literature that focus on heart segmentation, this study focuses on disease classification by combining end-diastolic (ED) and end-systolic (ES) phases that contain clinically important information through deep learning algorithms.

Keywords: Cardiomyopathy, cardiac mri, deep learning, dropout, optimization algorithms.

Introduction

Cardiovascular diseases are leading cause of death worldwide, and 17.3 million deaths are caused by it every year. The number is expected to grow up to more than 23.6 million by 2030 [1]. Dilated cardiomyopathy (DCM) and hypertrophic cardiomyopathy (HCM) are common chronic and life-threatening cardiomyopathies encountered in clinical practice [2, 3]. Early diagnosis of these diseases can significantly increase chances of correct treatment and survival. However, accurate and rapid diagnosis of DCM or HCM remains challenging due to the high complexity of cardiac structure and cardiac Dynamics [4]. DCM, among the mentioned diseases, is characterized by the enlargement of heart muscle and a decrease in its contractile strength. As a result, heart is unable to pump enough blood, leading to heart failure. HCM, on the other hand, is characterized by an abnormal thickening of heart muscle, particularly in left ventricle. This thickening can restrict blood flow and cause cardiac arrhythmias. Symptoms common to both diseases include chest pain, shortness of breath, irregular heartbeats, and sudden fainting. Echocardiography (ECHO), electrocardiography (ECG), and cardiac magnetic resonance (CMR) play an important role in the diagnosis of mentioned diseases. CMR imaging has gained widespread acceptance in the assessment of cardiac structure and function due to its noninvasiveness and versatility [5], and the quantitative measurements offered by CMR images play a significant role in understanding cardiac systolic and diastolic dysfunction [6, 7]. For this reason, CMR has been widely used in the clinical diagnosis of HCM [8, 9] and DCM [10].

Within the scope of this study, a classification study of cardiomyopathy disease was conducted using deep learning algorithms (CNN) on a Cardiac MRI dataset. In addition, the effects of Dropout (Regularization) technique and various optimization methods on deep learning models were examined. This study was carried out using Python programming language and Pycharm interface. The computer used had the following specifications: Lenovo Legion 7OQ1VDD, 16 GB RAM, Intel(R) Core(TM) i5-7300HQ CPU, and Windows 10 operating system.

Materials and Methods

In this section, information is provided about related dataset used and preprocessing steps applied to the dataset.

The dataset used in this study was established as part of "Multi-Centre, Multi-Vendor, and Multi-Disease Cardiac Segmentation (M&Ms) Challenge," organized within "Statistical Atlases and Computational Modelling of the Heart (STACOM)" workshop at MICCAI 2020 Conference [11]. The dataset comprises cardiac magnetic resonance (CMR) images of 375 participants scanned using four different scanners (Siemens, Philips, General Electric (GE), and Canon) in collaboration with clinical partners from six hospitals located in Spain, Canada, and Germany. All images were annotated consistently across centres using a standardized contouring standard operating procedure (SOP) and made publicly available. To the best of our knowledge, this dataset represents one of the most diverse open-access CMR data resources for developing and evaluating generalizable deep learning models in cardiac image segmentation [11, 12]. Due to the dataset update, there are currently a total of 345 patient images belonging to 8 different disease classes and one normal class.

The dataset also includes an Excel spreadsheet containing individuals' end-diastolic (ED) and end-systolic (ES) time information, as well as age, gender, height, and weight data.

Dataset Preparation

Within the scope of this study, modifications were made to the referenced M&Ms Cardiac MRI dataset for the sole purpose of classifying cardiomyopathy. A new dataset was created consisting of images from 271 subjects belonging to DCM, HCM, and Normal classes, along with ED and ES time information provided in the accompanying Excel spreadsheet. Using developed code, original 4-dimensional (4D) cardiac MRI data were converted into 2-dimensional (2D) images by extracting the middle slice based on patient-specific ED and ES time frames. Through this transformation, central slices corresponding to clinically critical ED and ES phases were selected from 4D volumetric images, thereby reducing data dimensionality while preserving essential structural and functional information required for analysis. Subsequently, resulting images were normalized and resized to a fixed dimension of 256×256 pixels. Additionally, data in (.nii) format were converted to (.jpeg) format and saved.

In order to improve the performance of deep learning models used for cardiomyopathy disease classification, detailed preprocessing and filtering were conducted on the images obtained from conversion process. Within this scope, images were sequentially evaluated based on brightness-contrast analysis and potential imaging artifacts. Initially, brightness and contrast of the images were assessed by calculating pixel intensity means and standard deviations. Images identified as too dark (mean < 30), too bright (mean > 220), or with very low contrast (std < 10) were excluded from the study. Subsequently, possible imaging artifacts or motion-related distortions were evaluated by analyzing edge intensities using a Laplacian filter. Images exhibiting variance above the defined threshold (8000) were considered to contain artifacts and thus excluded. Threshold values used in the designed code structure were determined based on literature and preliminary experiments [13]. Images meeting all specified evaluation criteria were deemed suitable for deep learning-based classification tasks and were selected for use in subsequent model training phase.

Example images illustrating evaluated image preprocessing steps are presented in Figure 1 and Figure 2.

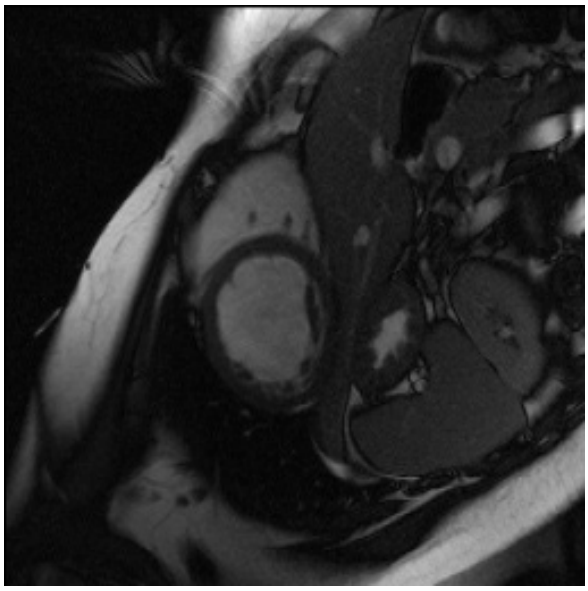


Figure 1: Accepted Image (DCM)

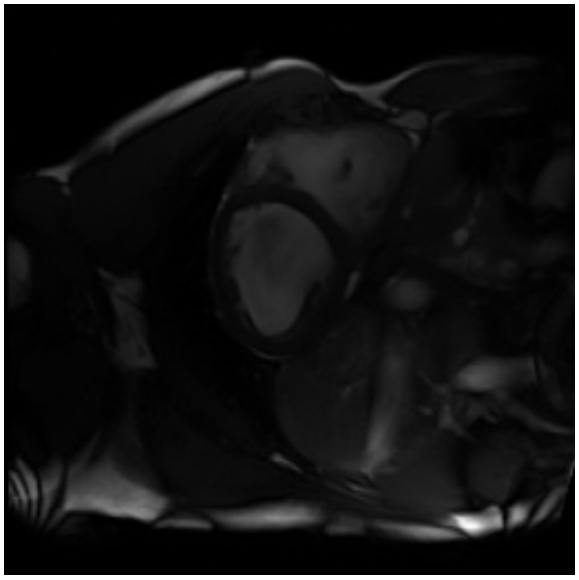


Figure 2: Rejected Image (DCM)

As a preparation for the next step, which is the model training process, images obtained for ED and ES phases of each patient in the dataset were concatenated side-by-side to create a single input image. An example of the concatenated image is presented in Figure 3. Then, grayscale images were converted into RGB format to meet the input requirements of models and resized to a fixed dimension of 224×448 pixels. The images were normalized with constant values (0.5, 0.5, 0.5). To maintain class distribution, the dataset was divided into training (70%), validation (15%), and test (15%) sets using the Stratified Split method. Accordingly, the number of images in the training set was determined as 28 for the Normal class, 27 for the HCM class, and 26 for the DCM class. The distribution of the created dataset is presented in Table 1. During training, data augmentation was applied only to the Normal class, including horizontal flipping, rotation up to 10 degrees, and brightness-contrast variations. For other classes, only fixed transformations were used.

Table 1: Dataset Distribution

Disease Label	Training Set	Validation Set	Test Set	Total
DCM	25	6	6	37
HCM	27	6	6	39
Normal	28	6	6	40
Total	80	18	18	116

ResNet18 and ResNet101 [14] were used as deep learning models. These models were trained using Adam, AdamW, Adamax, and Stochastic Gradient Descent (SGD) optimization algorithms [15,16]. Optimal model parameters were determined through hyperparameter optimization using Optuna [17]; accordingly, the number of epochs was set to 10, batch size to 8, learning rate to 5e-4, weight decay to 1e-5, and dropout [18] rate to 0.2. To evaluate the performance and robustness of models, Stratified 5-Fold Cross Validation was used. Following cross-validation, the final model was trained using the best parameters identified by combining the training and validation sets, and subsequently evaluated on an independent test set. Various performance metrics such as accuracy, precision, sensitivity, confusion matrix, ROC curves, Matthews correlation coefficient (MCC) and Co-

hen's Kappa were utilized for model evaluation [19]. The random seed was fixed across all experiments to ensure reproducibility.

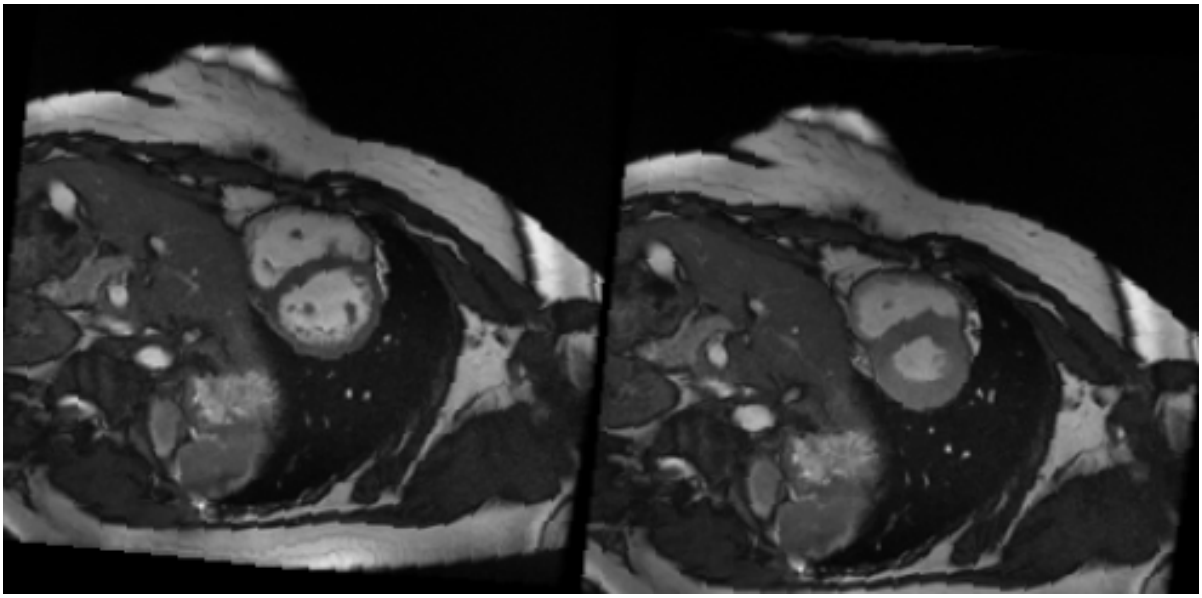


Figure 3: Example of Concatenated ED and ES Phase Images (DCM)

Results

In this section, all the results obtained are presented. In addition, visualizations related to the best-performing combination and comparisons regarding the effect of dropout are also included.

In this study, deep learning models ResNet18 and ResNet101 were compared for the classification of cardiomyopathy using different optimization algorithms (Adam, AdamW, Adamax, and SGD) along with the dropout regularization technique. The results were evaluated using various performance metrics, including accuracy, precision, sensitivity, confusion matrix, ROC curves, MCC, Cohen's Kappa coefficient, and F1-score.

The performance comparison of models with and without dropout, as well as results related to the applied optimization techniques, are summarized in Table 2.

Table 2: Performance Metrics of All Models and Optimization Algorithms

Model	Dropout	Optimizer	Overall Accuracy	Precision	Sensitivity	Cohen's Kappa	F1-Score	MCC	Loss
ResNet18	Yes	AdamW	0.94	0.95	0.94	0.91	0.94	0.92	0.18
ResNet18	Yes	Adam	0.94	0.95	0.94	0.91	0.94	0.92	0.19
ResNet18	**Yes**	**Adamax**	**0.94**	**0.95**	**0.94**	**0.91**	**0.94**	**0.92**	**0.13**
ResNet18	Yes	SGD	0.78	0.87	0.78	0.66	0.78	0.71	0.35
ResNet18	No	AdamW	0.89	0.92	0.89	0.83	0.89	0.84	0.24
ResNet18	No	Adam	0.78	0.83	0.78	0.66	0.77	0.70	0.69
ResNet18	**No**	**Adamax**	**0.89**	**0.92**	**0.89**	**0.83**	**0.89**	**0.84**	**0.30**
ResNet18	No	SGD	0.83	0.89	0.83	0.75	0.84	0.77	0.38
ResNet101	Yes	AdamW	0.83	0.85	0.83	0.75	0.83	0.76	0.90
ResNet101	Yes	Adam	0.67	0.68	0.67	0.50	0.67	0.50	0.83
ResNet101	Yes	Adamax	0.89	0.92	0.89	0.83	0.89	0.84	0.44
ResNet101	Yes	SGD	0.83	0.85	0.83	0.75	0.84	0.75	0.47
ResNet101	No	AdamW	0.78	0.78	0.78	0.67	0.78	0.67	0.60
ResNet101	No	Adam	0.72	0.80	0.72	0.58	0.72	0.60	0.84
ResNet101	No	Adamax	0.78	0.87	0.78	0.67	0.78	0.71	0.52
ResNet101	No	SGD	0.89	0.92	0.89	0.83	0.89	0.84	0.34

Performance Analysis of the Effect of Dropout in ResNet18 Model

In this study, the best results obtained using ResNet18 model with dropout were achieved with Adamax optimizer. The performance results of ResNet18 model using Adamax optimizer with dropout are presented in Figure 4 and Figure 5. Class-wise classification report results are given in Table 3.

Additionally, the performance results for ResNet18 model with all hyperparameters and optimizer kept the same, but without dropout, are shown in Figure 6 and Figure 7. Class-wise classification report results without dropout are given in Table 4.

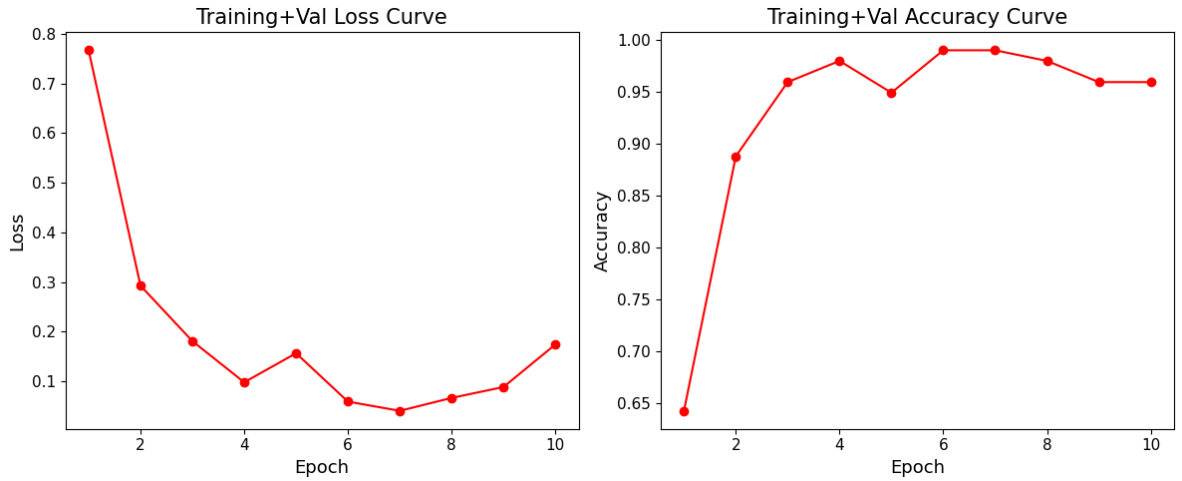


Figure 4: (a) Loss Curve of Final Model Training, (b) Accuracy Curve of Final Model Training

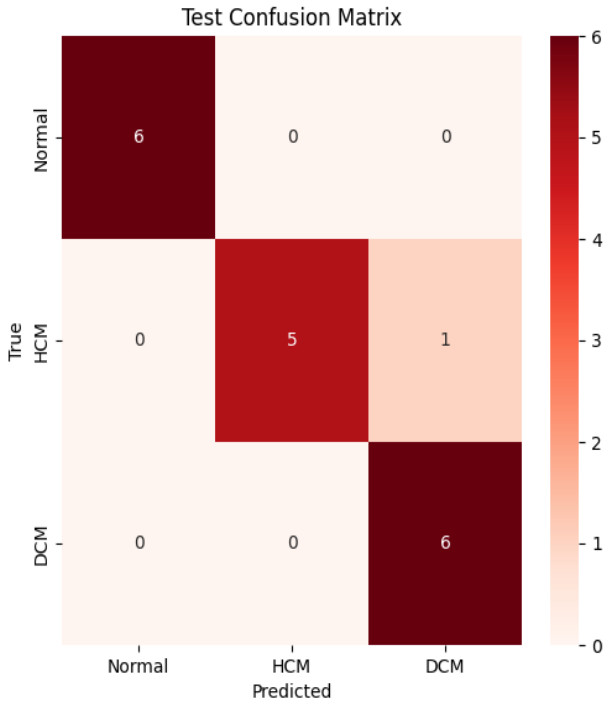


Figure 5: Test Confusion Matrix

Table 3: Class-wise Classification Report

Disease Label	Precision	Sensitivity	F1- Score	Support
Normal	1.00	1.00	1.00	6
HCM	1.00	0.83	0.91	6
DCM	0.86	1.00	0.92	6

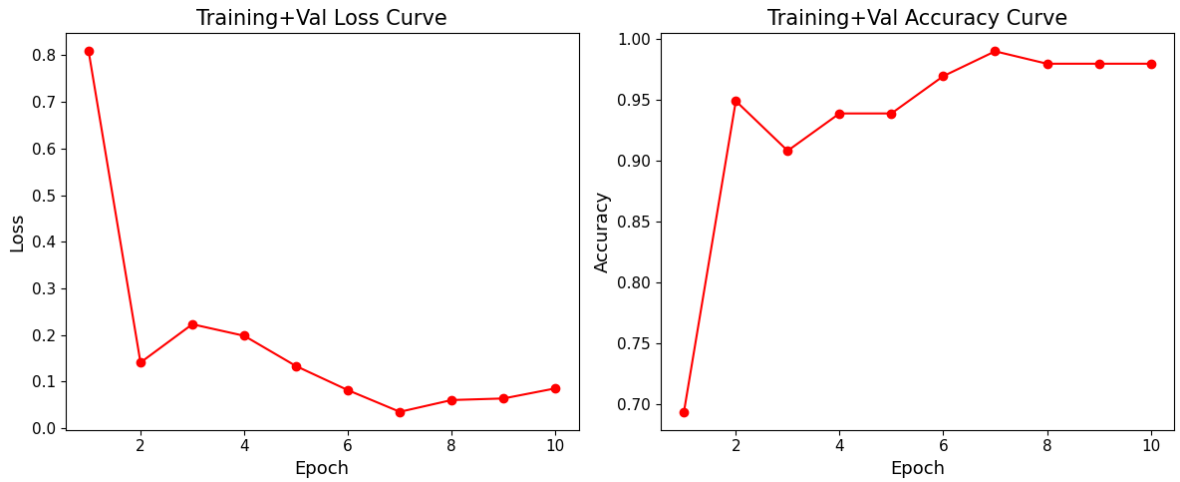


Figure 6: (a) Loss Curve without Dropout, (b) Accuracy Curve without Dropout

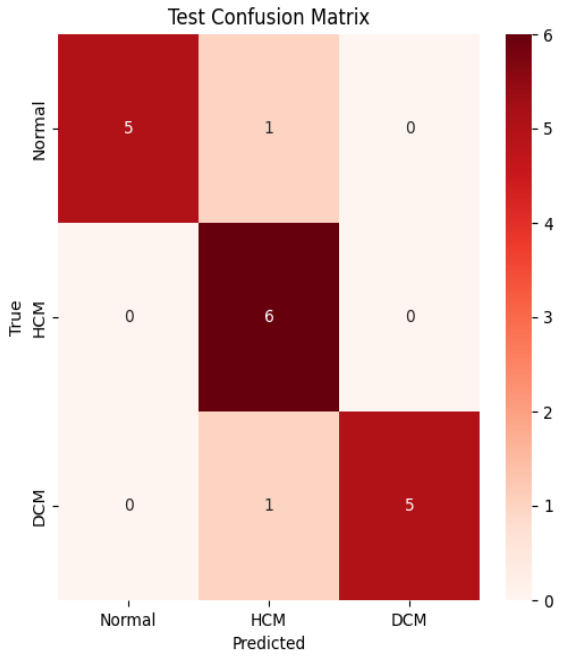


Figure 7: Test Confusion Matrix without Dropout

Table 4: Class-wise Classification Report without Dropout

Disease Label	Precision	Sensitivity	F1- Score	Support
Normal	1.00	0.83	0.91	6
HCM	0.75	1.00	0.86	6
DCM	1.00	0.83	0.91	6

Discussion

In this section, the effects of data preparation techniques, selected hyperparameters, dropout usage, optimization algorithms, and model architectures on model performance are discussed, and obtained results are evaluated within the scope of existing literature.

Data Preparation and the Effect of Selected Hyperparameters

In this study, M&Ms Cardiac MRI dataset was used to classify cardiomyopathy disease. In order to ensure the quality and consistency of dataset, some threshold values were determined and low-quality images were eliminated. The threshold values were determined based on the technical data of MRI device and literature research [13]. At first, performance analysis was done without filtering the images, and then the analysis was repeated after applying threshold-based filtering. In this context, the results after filtering were higher than those without filtering. Because it was found that low-quality images negatively affected the model's learning process and caused the training to progress unevenly. For this reason, image filtering was applied in the study.

Then, during the model training process, ED and ES phase information was combined side-by-side into a single input image. Because when ED and ES phases were given as two separate channels, it was observed that there was data loss during the Cross Validation

part. Therefore, it was determined that combining ED and ES phases side-by-side was more beneficial for model training performance.

Also, to provide balanced evaluation in multi-class classification problems, Stratified k-Fold Cross Validation was used instead of classic k-Fold, to preserve the class distribution in each fold. Both cases were evaluated, and the highest accuracy was observed with Stratified k-Fold Cross Validation. It is also stated in the literature that this method gives more accurate and higher results in imbalanced datasets [20].

Since the data used in this study was in grayscale format, it was converted to RGB format in line with the input requirements of model architectures using designed code structure. First, normalization was applied according to the normalization values required by models (mean: 0.485, 0.456, 0.406 / std: 0.229, 0.224, 0.225), but these values caused artificial contrast and brightness differences on the images converted from grayscale, and this caused a decrease in overall model accuracy. Therefore, fixed normalization values (mean=(0.5,0.5,0.5), std=(0.5,0.5,0.5)) were applied to all channels. In this way, the image structure was preserved and the overall accuracy of models increased. In the literature, it is also stated that customizing normalization values according to the structure of medical data affects model performance [21].

In the study, hyperparameter values were first determined using Optuna hyperparameter optimization method. In this way, efficient scanning was made over a wide range. Then, all value ranges of the hyperparameters were also tested experimentally, and the combinations obtained from Optuna were confirmed. In the literature, it is also stated that hyperparameters affect model performance and that automatic optimization methods provide an advantage in terms of computational cost [17].

Data augmentation was applied only to the normal class, because when it was applied to all classes, a significant decrease in model performance was observed. To deal with this, a strategy was applied where only basic transformation was used for diseased classes, and data augmentation was applied to normal class. In this way, data diversity was increased for normal class and the generalization ability of model was improved. For DCM and HCM classes, only basic transformation was used to prevent distortion of pathological patterns. The values selected in the data augmentation part were not chosen to be high, because in the literature it is stated that high values can distort the anatomical structure and mislead the model [22].

The Effect of Dropout Regularization

In this study, all dropout values were tested at first. It was observed that model performance was lower at dropout value 0.5 compared to 0.2. The reason for this is thought to be that due to model architecture and limited dataset size, the models could not learn sufficiently and caused information loss. For this reason, the dropout rate was chosen as 0.2 in this study. In the literature, it is also stated that dropout values of 0.2 or 0.3 in classification studies involving medical images give more efficient and stable results in terms of model performance [18, 23]. For this reason, when a comparison was made by keeping all hyperparameters constant and using Adamax optimization algorithm in ResNet18 and

ResNet101 models, 94% accuracy and 0.13 loss were obtained when dropout was applied in the ResNet18 model. Similarly, when dropout was not applied, 89% accuracy and 0.30 loss were obtained. When the same strategy was applied in the ResNet101 model, 89% accuracy was obtained with dropout effect, and 78% accuracy was obtained without dropout effect. (See Table 2 and Figure 4–7). When the whole table is examined within the scope of this study, it is confirmed that general dropout effect is consistent with the literature and has a positive effect on model performance.

The Effect of Optimization Algorithms

The results obtained in this study revealed that performance of optimization algorithms differs depending on the use of dropout and model architecture. For example, for ResNet18 model, when dropout was applied, higher accuracy and loss values were observed in adaptive optimization algorithms (Adam, AdamW, and Adamax). SGD, on the other hand, remained much lower. At the same time, in ResNet101 model, the opposite situation was observed. When dropout was not applied, SGD gave a much higher accuracy rate. The reason for this is thought to be that in models with fewer layers such as ResNet18, information loss that may occur due to the dropout effect can be better compensated by adaptive optimizers, and in model architectures with more layers such as ResNet101, since they already have a regularization effect, adaptive optimizers may learn in an unbalanced way due to their flexible nature. For this reason, SGD, which provides more controlled learning due to its structure, is thought to have achieved higher accuracy. At the same time, when dropout was not used in ResNet18 model and when dropout was applied in ResNet101 model, the reason why Adamax and AdamW optimizers generally had higher accuracy is thought to be that they are more balanced and stable due to their structure in medical image classification. This idea is also confirmed in the literature [15]. In the light of all this information, it is seen that conclusions made are also parallel with the literature. In the literature, it has been stated that the performance of optimization algorithms varies depending on the dropout method used, model architecture, and the scope of dataset [16, 24].

The Effect of Model Architecture

In this study, the models used are ResNet18 and ResNet101 architectures. When the performances of model architectures are compared in general, it is observed that ResNet18 architecture reaches higher accuracy values. The main reason for this situation is that dataset used in the study contains few samples. For this reason, ResNet18 model, which has fewer layers and a simple structure, understands the dataset better with its number of samples. On the other hand, deep architectures such as ResNet101, which contain many layers, are generally used with datasets that contain more samples. At this stage, it is thought that datasets with a small number of samples have difficulty in providing meaningful learning in deep architectures, and therefore, since some layers cannot be fed with sufficient samples, and excessive number of layers are continuously fed with same examples, overfitting occurs and the overall performance decreases. In the literature, it is also stated that small datasets are more successful in simpler model architectures [25].

Conclusion

In this study, a classification method is proposed without using segmentation in order to classify cardiomyopathy diseases using Cardiac MRI dataset with deep learning algorithms. In this study, it is focused on performing classification directly on raw MRI images by combining the images of ED and ES phases, which are clinically important in the classification of cardiomyopathy disease. Unlike this study, in the literature, there are generally studies focused on segmentation of heart structures. This study significantly reduces dependency on algorithms in terms of not requiring any segmentation or annotation. In this way, it stands out in terms of computational cost, practicality, and time for real clinical scenarios.

A comparative evaluation was made in terms of performance of four different optimization algorithms (Adam, AdamW, Adamax and SGD) and regularization techniques (Dropout) on different model architectures (ResNet18, ResNet101). The best result obtained in the study was 94% accuracy and 0.13 loss in ResNet18 model with Adamax optimizer and 0.2 dropout rate. It is understood from this that in classification studies with low data numbers, the performance of models used is significantly related to dropout and optimization algorithms. The main limitation in this study is the small number of dataset used. For future studies, it is recommended to increase the number of samples in dataset to be used, to apply segmentation, and to study hybrid models that can combine both CNN-based and demographic and biological data. Despite the mentioned limitations, a significant contribution has been made to the literature with high classification accuracy without the need for segmentation, through proper regularization, model architecture, and optimizer selection.

Declaration of Interest: The authors declare that there is no conflict of interest.

Author Contributions: The authors contributed equally to this work.

REFERENCES

[1] L.V. Romaguera et al., "Left Ventricle Segmentation In Cardiac MRI Images Using Fully Convolutional Neural Networks", in Proceedings of SPIE Medical Imaging, International Society for Optics and Photonics, 10134, 760–770, 2017.

[2] C. Xia et al., "A Multi-Modality Network For Cardiomyopathy Death Risk Prediction With CMR Images and Clinical Information", 22nd International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 577–585, October 13-17, 2019, Shenzhen, China.

[3] D.S.B. Sundaram et al., "Natural Language Processing Based Machine Learning Model Using Cardiac MRI Reports To Identify Hypertrophic Cardiomyopathy Patients," Frontiers in Biomedical Devices, American Society of Mechanical Engineers, 84812, V001T03A005, 2021.

[4] D. Alis et al., "The Influence of Cardiac Motion On Radiomics Features: Radiomics Features of Nonenhanced CMR Cine Images Greatly Vary Through The Cardiac Cycle", Eur. Radiol., 31, 2706–2715, 2021.

[5] T.D. Karamitsos et al., "The Role of Cardiovascular Magnetic Resonance Imaging In Heart Failure", J. Am. Coll. Cardiol., 54, 1407–1424, 2009.

[6] M. Loecher, L.E. Perotti, and D.B. Ennis, "Using Synthetic Data Generation To Train A Cardiac Motion Tag Tracking Neural Network", Med. Image Anal., 74, 102223, 2021.

[7] A.S. Antonopoulos et al., "Machine Learning of Native T1 Mapping Radiomics For Classification of Hypertrophic Cardiomyopathy Phenotypes", Sci. Rep., 11, 1–11, 2021.

[8] M. Śpiewak et al., "Impact of Cardiac Magnetic Resonance On The Diagnosis of Hypertrophic Cardiomyopathy—A 10-Year Experience With Over 1000 Patients", Eur. Radiol., 31, 1194–1205, 2021.

[9] N.Z. Farahani et al., "Explanatory analysis of a machine learning model to identify hypertrophic cardiomyopathy patients from EHR using diagnostic codes," IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 1932–1937, December 16-19, 2020, Seoul, South Korea.

[10] A.J. Jacob et al., "Deep-Learning-Based Disease Classification In Patients Undergoing Cine Cardiac MRI," J. Magn. Reson. Imaging, 61, 4, 1635–1647, 2025.

[11] V.M. Campello et al., "Multi-Centre, Multi-Vendor and Multi-Disease Cardiac Segmentation: The M&Ms Challenge", IEEE Trans. Med. Imaging, 40, 12, 3543-3554, 2021.

[12] C. Martín-Isla et al., "Deep Learning Segmentation of The Right Ventricle In Cardiac MRI: The M&Ms Challenge", IEEE J. Biomed. Health Inform., 27, 7, 3302-3313, 2023.

[13] Y. Wang et al., "A Fast and Robust Framework For Automatic MRI Quality Assessment Using Image Processing Techniques", Med. Phys., 46, 7, 3122–3133, 2019.

[14] M. Shafiq, and Z. Gu, "Deep Residual Learning For Image Recognition: A Survey", Appl. Sci., 12, 18, 8972, 2022.

[15] S.H. Haji, and A.M. Abdulazeez, "Comparison of Optimization Techniques Based On Gradient Descent Algorithm: A Review", PalArch's J. Archaeol. Egypt/Egyptol., 18, 4, 2715–2743, 2021.

[16] H. Kohli, J. Agarwal, and M. Kumar, "An Improved Method For Text Detection Using Adam Optimization Algorithm", Global Transitions Proc., 3, 1, 230–234, 2022.

[17] P. Srinivas, and R. Katarya, "hyOPTXg: OPTUNA Hyper-Parameter Optimization Framework For Predicting Cardiovascular Disease Using Xgboost", Biomed. Signal Process. Control, 73, 103456, 2022.

[18] N. Srivastava et al., "Dropout: A Simple Way To Prevent Neural Networks From Overfitting", J. Mach. Learn. Res., 15, 1, 1929–1958, 2014.

[19] M. Sokolova, and G. Lapalme, "A Systematic Analysis of Performance Measures For Classification Tasks", Inf. Process. Manag., 45, 4, 427–437, 2009.

[20] S. Widodo, H. Brawijaya, and S. Samudi, "Stratified k-Fold Cross Validation Optimization On Machine Learning For Prediction", Sinkron: J. Dan Penelit. Tek. Informatika, 6, 4, 2407–2414, 2022.

[21] G. Tarroni et al., "Learning-Based Quality Control For Cardiac MR Images", IEEE Trans. Med. Imaging, 38, 5, 1127–1136, 2019.

[22] M. Buda, A. Maki, and M.A. Mazurowski, "A Systematic Study of The Class Imbalance Problem In Convolutional Neural Networks", Neural Netw., 106, 249–259, 2018.

[23] M. Ghafoorian et al., "Location Sensitive Deep Convolutional Neural Networks For Segmentation of White Matter Hyperintensities", Sci. Rep., 7, 1, 2017.

[24] L. Bottou, "Stochastic Gradient Descent Tricks," Neural Networks: Tricks of The Trade, 2nd ed. Berlin: Springer-Verlag, 421–436, 2012.

[25] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How Transferable Are Features In Deep Neural Networks?", Adv. Neural Inf. Process. Syst., 27, 3320–3328, 2014.

DEEP LEARNING MODELS FOR CLASSIFICATION, DIAGNOSIS, AND TREATMENT OF ROTATOR CUFF TEARS: A COMPREHENSIVE EXAMINATION

Kemal Polat
Bolu Abant Izzet Baysal University, Faculty of Engineering, Department of Electrical and Electronics Engineering, Bolu, Türkiye

Büşra Kaya Olgaç*
Bolu Abant Izzet Baysal University, Faculty of Engineering, Department of Electrical and Electronics Engineering, Bolu, Türkiye
*(Corresponding author) busra6578@gmail.com

ABSTRACT

Orthopedics and Traumatology is a branch of medicine that deals with preventing, diagnosing and treating problems related to the skeletal system such as bones, muscles, joints, tendons and ligaments. This field has significant potential in terms of Deep Learning (DL) and its applications due to its openness to technological developments. The use of deep learning models in the classification, diagnosis and treatment of rotator cuff (RC) tears is very important for the accurate evaluation of scientific results. In this study, it is aimed to propose new approaches by examining the classification, diagnosis and treatment studies of RC tears based on medical image analysis of DL models. For this purpose, deep learning models used in the classification, diagnosis and treatment of RC tears were examined. In these examinations, it was seen that the DL models used were focused on the Supraspinatus muscle tear, which is a RC tear. In this study, where DL models made on RC tears were examined, it was tried to reveal what to do to detect the tear location and determine the degree of tear of supraspinatus, infraspinatus, teres minor and subscapularis muscles with the DL model.

Keywords: deep learning, rotator cuff muscles, supraspinatus, infraspinatus, teres minor, subscapularis.

Introduction

DL plays an important role in the rapid technological developments in the field of medicine. DL Models are a sub-field of Machine Learning(ML) that uses neural network algorithms to correctly evaluate the results obtained [1]. DL is used in the field of medicine in the decision-making and classification phase through imaging systems [2]. Orthopedics and Traumatology are among the sub-branches of medicine and are one of the areas where rapid decisions must be made in emergency interventions. Incorrect diagnoses lead to incorrect diagnoses, which can lead to defects that limit or prevent mobility [3].

The diversity of studies conducted in the field of Orthopedics and Traumatology, and the fact that it is open to technological developments for fast and accurate interventions, have become an area where successful studies can be conducted in terms of DL and its Applications. Especially for medical interventions that require urgent decision-making, DL Applications, which are expected to shorten and accelerate the decision-making process of healthcare professionals and doctors, will greatly benefit the field of Orthopedics and Traumatology. The Orthopedics and Traumatology department examines RC tears as a subject. The classification of RC tears, DL models to be used in their diagnosis and treatment are very important for the correct evaluation of scientific results.

There have been and continue to be many studies on Deep Learning and its applications in the field of Orthopedics and Traumatology. DL Models have been used in studies on image quality to determine accurate diagnosis and treatment methods in imaging traumatic diseases [4], in distinguishing similar fractures that are difficult to distinguish from each other [5], in medical imaging methods in spine care [6], in determining the diagnosis and treatment method of scoliosis [7,8], in bone tumor classification [9, 10, 11] and in surgical planning [12]. These studies are some of the applications of DL models in the field of Orthopedics and Traumatology.

Orthopedics and Traumatology; as a field where immediate interventions are required as a result of diseases and serious accidents affecting people's daily lives, this study has focused on RC tears, which is one of the health problems in this field. The RC consists of the supraspinatus, infraspinatus, teres minor and subscapularis muscles and provides movement and stability of the shoulder joint by wrapping the humeral head [13]. The working mode and traction directions of the RC tendons [14] are shown in Figure 1.

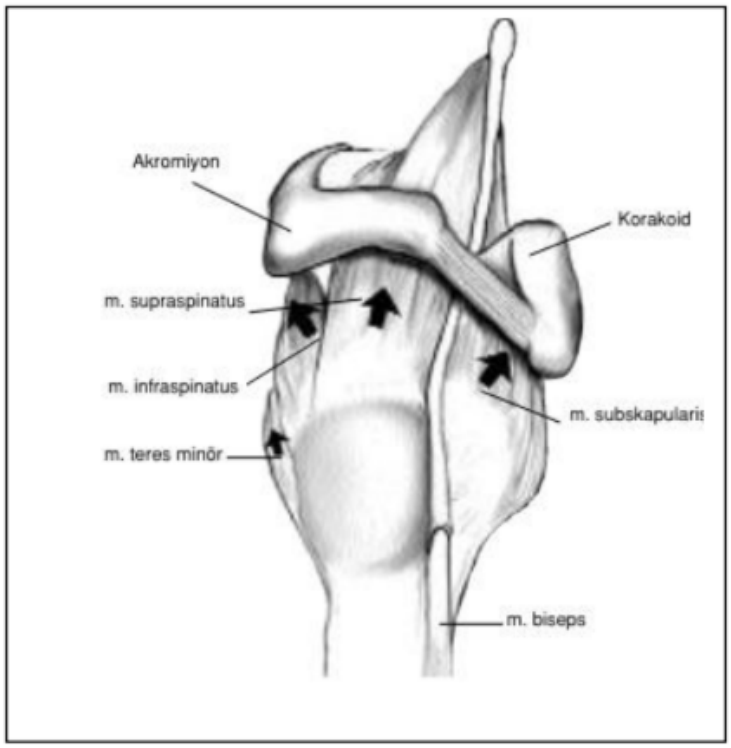


Figure 1: The working mode and traction directions of the RC tendons

One of the diseases commonly seen in orthopedic surgery is RC tear, which is the cause of lipid infiltration and changes that cause damage to the supraspinatus, infraspinatus, subscapularis and teres minor muscles[15]. The generally accepted theories for the formation of RC tears are that the tear occurs after acute trauma or on the basis of chronic repetitive trauma [16]. RC tears have been understood more clearly with the developments in technology. Technological devices in the field of imaging provide useful information on this subject. Visual assessment using magnetic resonance imaging (MRI) plays an important role in organizing the operative solution, detecting RC tears, and estimating their degree[15]. DL is effectively used in the classification, diagnosis and prognosis stages of the images obtained.

DL models are quite effective for the preliminary classification of clinical diagnoses. In the health sector, due to reasons such as limited vision of the human eye and intensive work, many problems may arise in the diagnosis, treatment and classification of the disease [17]. For these reasons, the development of DL models in the field of RC tears has focused on diagnosis and classification. An example of MR images belonging to two classes, (up) RC tear and (down) normal, is shown in Figure 2 [30].

Figure 2: MR images belonging to two classes, (up) RC tear and (down) normal

This article presents new approaches by examining DL models in determining the classification, diagnosis and treatment methods of RC tears.

Related Studies

Taghizadeh et al. [18] developed a convolutional neural network (CNN) that can automatically measure and characterize the level of RC muscle degeneration from shoulder CT images, including atrophy and fat infiltration. Radiologists manually determined the premorbid (healthy) boundaries of the RC muscle from CT images. They used the CNN

architecture U-Net to estimate these premorbid muscles. The automatically estimated premorbid boundaries were then used to measure the rate of muscle atrophy, fat infiltration, secondary bone formation, and overall muscle degeneration. With the U-Net architecture used, they achieved 87% prediction in muscle atrophy, 91% in fat infiltration, and 91% in overall muscle degeneration. The architecture is shown in detail in Figure 3 [18].

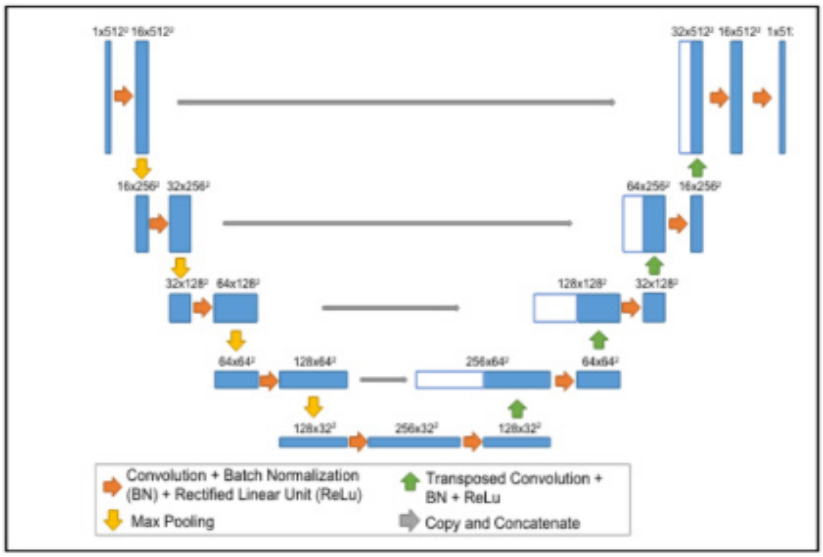


Figure 3: The U-Net architecture used to predict these premorbid muscles

Dana et al. [19] developed a DL model to detect RC tears on shoulder MRI and classify them as no tear, partial or complete tear. They developed a DL algorithm using four data sets from MRI findings as input: fluid-sensitive sequences in three planes and a sagittal oblique T1-weighted sequence.The architecture is composed of four parallel 3D ResNet50 CNN architectures that are trained with transfer learning and then reorganized for the targeted domain. After averaging the class predictions of four different models, the class with the highest probability was determined to be the RC tear type. They found overall AUC (area under the RoC curve) values of 0.93, 0.89, and 0.90 for partial thickness infraspinatus, supraspinatus, and subscapularis tears. They found AUC values of 0.98, 0.99, and 0.95 for full-thickness infraspinatus, supraspinatus, and subscapularis tendon tears.

Kyungsu et al. [20] developed an end-to-end fully convolutional neural network called pRe-trained (pre-trained model trained on a specific task or dataset) Segmentation Model Adopting Classification Architecture (SMART-CA) integrated on a novel positive loss function (IPLF) to accurately detect the locations of RC tears in an orthopedic examination using ultrasound imaging. The precision, recall and Dice coefficient values of SMART-CA in segmenting RC tears from ultrasound images with severe speckle noise were calculated as 0.337% (+22.5%), 0.860% (+15.8%) and 0.484% (+28.5%), respectively.

Giovanna et al. [21] developed a CNN structure that can select a given shoulder sagittal MR image (Y-view) and automatically segment the rotator cuff muscles in a Y-view. They used the U-Net architecture. They created two models, Model-A for selecting Y posture and Model-B for muscle separation. They observed that the first 3 produced accurate RC muscle segmentations with an accuracy of 98% for selecting a suitable Y-view for Model A, and an average Dice score of >0.93 for Model B.

Thao et al. [22] developed a perspective for automatically categorizing cuff tears, visualizing the tear area using ultrasound images and CNNs.

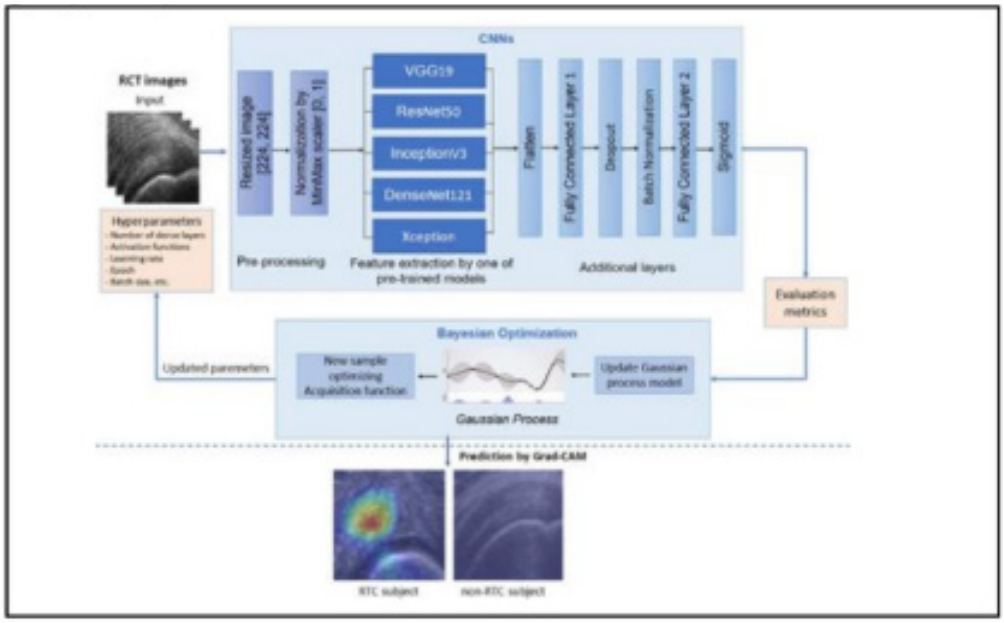


Figure 4: CNN architecture used to automatically categorize headline tears

To this end, they developed transfer learning with five fully trained models: VGG19, InceptionV3, Xception, ResNet50, and DenseNet121. They used it to test and train CNN models with five-fold cross-validation. The model they used is shown in Figure 4. They claimed that among the five fully trained models, DenseNet121 is a better model than the other models with 88.2% accuracy, 93.8% sensitivity, 83.6% specificity and 0.832 AUC value.

Eungjune et al. [23] studied a 3D CNN-based method for automatic and accurate diagnosis of RC tears. In this study, 3D CNN aims to automatically diagnose the presence or absence of RC tears, categorize the tear size, and visualize the tear location in 3D. They used the Voxception-ResNet (VRN) neural network model to complete the training of the 3D CNN. They noted that the sensitivity value was 0.94 and the specificity value was 0.90. They explained that the proposed method showed higher sensitivity (0.94) and specificity (0.90) than clinical experts.

Eiko et al. [24] aimed to evaluate the diagnostic performance of CNN to detect and classify RC tears in shoulder radiographs and compare it with the diagnostic performance of orthopedic surgeons. They recorded 92% sensitivity, 69% specificity, 86% accuracy and 0.88 area under the receiver operating curve for the detection of rotator cuff tears and claimed that CNN exhibited a high performance compared to orthopedic surgeons.

According to Yuichi et al. [25], the method they used consists of determining the supraspinatus muscle location using the scapula classification result and classifying the supraspinatus muscle according to this determination. They used U-Net in the classification of scapula and supraspinatus muscles. With this method, they observed that the average Dice = 0.863 in sagittal-oblique images of the supraspinatus muscle classification. They claimed that this value was lower than the Dice value (0.91) obtained with the traditional method.

Deming et al. [26] developed a deep learning-based model to automatically diagnose supraspinatus tears using shoulder MRI. They trained two advanced CNNs based on Xception and optimized them to detect supraspinatus tears. They claim that it is consistent with clinical comparisons.

Jiufa et al. [27] investigated the usefulness of DL method to detect supraspinatus tendon injury in shoulder MRI. They developed this research by using U-Net networks to segment and isolate supraspinatus tendon tear and then a swin transformer architecture to determine the presence or absence of supraspinatus tendon tear. They also evaluated Densnet101 and Resnet50 as classifiers in this study. In their study, they claimed that the ST tear detection system has a diagnostic performance equal to that of experienced clinical radiologists.

Jason et al. [28] examined the feasibility of DL for spontaneous detection of supraspinatus tears on MRI. For this study, images were collected of complete or partial tears and intact supraspinatus tendons. They developed a pipeline-like pipeline consisting of three stages with ResNet, U-Net, and CNN classifier neural network models.They found the test sensitivity and specificity as 85% and 85%, AUROC 0.943, and Dice segmentation accuracy as 0.814, respectively. From these evaluations, they concluded that deep learning is a useful method for detecting supraspinatus tears on MRI.

If we summarize the studies done in this field in a single table 1 and table 2 ,

Table 1: Results of the researched studies

	Fat Infiltration	**Muscle Atrophy**	**Muscle Degeneration**
Taghizadeh et al.			
U-Net	0,91	0,87	0,91
	Infraspinatus AUC	**Supraspinatus AUC**	**Subscapularis AUC**
Dana et al.			
3D ResNet50 (full-thickness)	0,98	0,99	0,95
3D ResNet50 (partial thickness)	0,93	0,89	0.90
	The Precision	**Recall**	**Dice Coefficient**
Kyungsu et al.			
SMART-CA	0,33	0,86	0,48
	Accuracy		
Giovanna et al.			
U-Net (Model-A)	0,98	-	-
U-Net (Model-B)	0,93	-	-

Table 2: Continuation of the researched study results

	Accuracy	**Sensitivity**	**Specificity**	**AUC**
Thao et al.				
DenseNet121	0,88	0,93	0,83	0,83
	Accuracy	**Sensitivity**	**Specificity**	**AUC**
Eungjune et al.				
Voxception-ResNet		0,94	0,90	-
	Accuracy	**Sensitivity**	**Specificity**	**AUC**
Eiko et al.				
CNN	0,86	0,92	0,69	0,88
	Accuracy	**Sensitivity**	**Specificity**	**AUC**
Jason et al.				
ResNet, U-Net,CNN	0,81	0,85	0,85	0,94
	Dice Coefficien (traditional method)	**Dice Coefficien**		
Yuichi et al.				
U-Net	0,91	0,86	-	-

Conclusion

As a result of this study, it is observed that RC tears significantly affect the daily lives of patients before and after diagnosis. Orthopedic surgeons play a role in determining the diagnosis and treatment of this disease. After the findings obtained from MRI are evaluated by orthopedic surgeons, the method of intervention is determined. At this stage, rapid and automatic determination of RC muscle tears will contribute to the determination of the surgical method. The necessary acceleration in classification, diagnosis and prognosis can be achieved by using DL models. The studies conducted for this purpose do not point to a single common point, but they are diverse. This gap in the literature can be filled by gathering all these diversities, even partially, under a single roof. The planned study aims to classify the tears found in the RC [29], which is formed by the supraspinatus, infraspinatus, teres minor and subscapularis muscles. It has been observed in the studies that CNNs are successful especially in the field of image processing and classification. For the planned study, different CNN architectures such as AlexNet, VGG, ResNet can achieve success in the muscle classification stage. It has been observed in the studies that 3D-CNN is used in three-dimensional data sets [23]. 3D-CNNs have been observed to be especially effective in detecting detailed anatomical structures and pathological changes [23]. The location of the tear can be detected using the Voxception-ResNet (VRN) structure.

Declaration of Interest: The authors declares that there is no financial or personal conflict of interest regarding this study. No interest group or commercial organization has supported the conduct of the study, analysis of the data, and writing of the article.

Author Contributions: The authors confirms that there is no conflict of interest in the study.

REFERENCES

[1] C. Chakraborty, M. Bhattacharya, S. Pal, S.S. Lee, "From Machine Learning to Deep Learning: An Advances of The Recent Data-Driven Paradigm Shift in Medicine and Healthcare", Current Research Biotechnology, 2023.

[2] I. Castiglioni, R. Leonardo, M. Codari, L. Giovanni, C. Salvatore, M. Interlenghi, F. Gallivanone, A. Cozzi, N.C. D'Amico, F. Sardanelli, "AI Applications to Medical Images: From Machine Learning to Deep Learning", Physica Medica, 83, 9-24, 2021.

[3] A.M. Keating, "Orthopedic Emergencies: The Common and the Critical", Physician Assistant Clinics, 9, 1, 81-90, 2024.

[4] G. Feuerriegel, K. Weiss, V.A. Tu, Y. Leonhardt, J. Neumann, F. Gassert, Y. Haas, M. Wurm, M. Makowski, K. Woertler, D. Karampinos, A. Gersing, "Deep-Learning-Based Image Quality Enhancement of CT-Like MR Imaging in Patients with Suspected Traumatic Shoulder Injury", European Journal of Radiology, 170, 111246, 2024.

[5] J. Schilcher, A. Nilsson, O. Andlid, A. Eklund, "Fusion of Electronic Health Records and Radiographic Images for a Multimodal Deep Learning Prediction Model of Atypical Femur Fractures", Computer in Biology and Medicine, 168, 107704, 2024.

[6] C. Constant, C. Aubin, H. Kremers, D. Garcia, C. Wyles, P. Rouzrokh, A. Larson, "The Use of Deep Learning in Medical Imaging to Improve Spine Care: A Scoping Review of Current Literature and Clinical Applications", North American Spine Society Journal, 15, 100236, 2023.

[7] S. Goral, U. Kose, "Skolyoz için Kapsül Ağları Tabanlı Otomatik Ölçüm Sistemi", Düzce Üniversitesi Bilim ve Teknoloji Dergisi, 9, 5, 2087-2101, 2021.

[8] T. Kokabu, S. Kanai, N. Kawakami, K. Uno, T. Kotani, T. Suzuki, H. Tachi, Y. Abe, N. Iwasaki, H. Sudo, "An Algorithm for Using Deep Learning Convolutional Neural Networks with Three Dimensional Depth Sensor Imaging in Scoliosis Detection", Spine Journal, 21, 6, 980-987, 2021.

[9] C. Pan, L. Lian, J. Chen, R. Huang, "FemurTumorNet: Bone Tumor Classification in the Proximal Femur Using DenseNet Model Based on Radiographs", Journal of Bone Oncology, 42, 100504, 2023.

[10] F. Eweje, B. Bao, J. Wu, D. Dalal, W. Liao, Y. He, Y. Luo, S. Lu, P. Zhang, X. Peng, R. Sebro, H. Bai, L. States, "Deep Learning for Classification of Bone Lesions on Routine MRI", eBioMedicine, 68, 103402, 2021.

[11] Y. He, I. Pan, B. Bao, K. Halsey, M. Chang, H. Liu, S. Peng, R. Sebro, J.Guan, T. Yi, A. Delworth, F. Eweje, L. States, P. Zhang, Z. Zhang, J. Wu, X. Peng, H. Bai, "Deep Learning-Based Classification of Primary Bone Tumors on Radiographs: A Preliminary Study", eBioMedicine, 62, 103121, 2020.

[12] Y. Qu, X. Li, Z. Yan, L. Zhao, L. Zhang, C. Liu, S. Xie, K. Li, D. Metaxas, W. Wu, Y. Hao, K. Dai, S. Zhang, X. Tao, S. Ai, "Surgical Planning of Pelvic Tumor Using Multi-View CNN With Relation-Context Representation Learning", Medical Image Analysis, 69, 101954, 2021.

[13] S. Özkan, C. Adanaş, "Omuz İmpingement Sendromu ve Rotator Cuff Cerrahisinde Orta Dönem Sonuçaları", Van Sağlık Bilimleri Dergisi, 11, 11-14, 2018.

[14] M. Yel, M. Arazi, "Rotator Manşet Yırtıklarının Sınıflandırılması", Acta. Orthopaedica et Traumatologica Turcica, 37, 77-82, 2003.

[15] J. Young, R. Kyunghan, Y. Sungmi, N. Rum, Y. Sunhyun, H.S. Park, J.C. Yoo, E. Park, K. Cho, B.H. Cho, I.Y. Kim, "Development of An Automatic Muscle Atrophy Measuring Algorithm to Calculate the Ratio of Supraspinatus in Supraspinous Fossa Using Deep Learning", Computer Methods in Programs Biomedicine, 182, 105063, 2019.

[16] C.L. Getz, J.E. Buzzell, S.G. Krishnan, "Shoulder Inst ability and Rotator Cuff Tears", JAAOS, 299–315, 2011.

[17] S. Yang, B. Yin, W. Cao, C. Feng, G. Fan, S. He, "Diagnostic Accuracy of Deep Learning in Orthopaedic Fractures: A Systematic Review and Meta- Analysis", Clinical Radiology, 75, 9, 713-728, 2020.

[18] E.Taghizadeh, O. Truffer, F. Becce, S. Eminian, S, Gidoin, A. Terrier, A. Farron, P. Büchler, "Deep Learning for The Rapid Automatic Quantification and Characterization of Rotator Cuff Muscle

Degeneration From Shoulder CT Datasets", European Radiology, 31, 181-190, 2021.

[19] D. Lin, M. Schwier, B. Geiger, E. Raithel, H. Busch, J. Fritz, M. Kline, M. Brooks, K. Dunham, M. Shukla, E. Alaia, M. Samim, W. Joshi, W. Walter, J. Ellermann, H. Ilaslan, D. Rubin, C. Winalski, M. Recht, "Deep Learning Diagnosis and Classification of Rotator Cuff Tears on Shoulder MRI", Investigative Radiol, 58, 405-412, 2023.

[20] K. Lee, J.Y. Kim, M.H. Lee, C.H. Choi, J.Y. Hwang, "Imbalanced Loss- Integrated Deep-Learning-Based Ultrasound Image Analysis for Diagnosis of Rotator-Cuff Tear", Sensors, 21, 2021.

[21] G. Medina, C.G. Buckless, E. Thomasson, L.S. Oh, M. Torriani, "Deep Learning Method for Segmentation of Rotator Cuff Muscles on MR Images", Skeletal Radiology, 50, 683-692, 2021.

[22] T.T.Ho, G.T. Kim, T. Kim, S. Choi, E.Park, "Classification of Rotator Cuff Tears in Ultrasound Images Using Deep Learning Models", Medical and Biological. Engineering Computing, 60, 1269-1278, 2022.

[23] E. Shim, J.Y. Kim, J.P. Yoon, S.Y. Ki, T. Lho, Y. Kim, S.W. Chung, "Automated Rotator Cuff Tear Classification Using 3D Convolutional Neural Network", Scientific Reports, 10, 2020.

[24] E. Hashimoto, S. Maki, N. Ochiai, S. Ise, K. Inagaki, Y. Hiraoka, F. Hattori, S. Ohtori, "Automated Detection and Classification of The Rotator cuff Tear on Plain Shoulder Radiograph Using Deep Learning", Journal of Shoulder and Elbow Surgery, 2024.

[25] Y. Wakamatsu, N. Kamiya, X. Zhou, H. Kato, T. Hara, H. Fujita, "Automatic Segmentation of Supraspinatus Muscle via Bone-Based Localization in Torso Computed Tomography Images Using U-Net", IEEE Access, 9, 2021.

[26] D. Guo, X. Liu, D. Wang, X. Tang, Y. Qin, "Development and Clinical Validation of Deep Learning for Auto-Diagnosis of Supraspinatus Tears", Journal of Orthopaedic Surgery and Research, 18, 2023.

[27] J. Cui, X. Xia, J. Wang, X. Li, M. Huang, S. Miao, D. Hao, J. Li, "Fully Automated Approach for Diagnosis of Supraspinatus Tendon Tear on Shoulder MRI by Using Deep Learning", Academic Radiology, 31, 994- 1002, 2024.

[28] J. Yao, L. Chepelev, Y. Nisha, P. Sathiadoss, F.J. Rybicki, A.M. Sheikh, "Evaluation of a Deep Learning Method for the Automated Detection of Supraspinatus Tears on MRI", Skeletal Radiology, 51, 1765-1775, 2022.

[29] S. Akpınar, G. Özkoç, N. Cesur, "Rotator Manflet Anatomisi, Biyomekaniği ve Fizyopatolojisi", Acta Orthopaedica et Traumatologica Turcica, 37, 4-12, 2003.

[30] M.A. Esfandiari, M.F. Tafti, N.J. Dabanloo, F. Yousefirizi, "Detection of The Rotator Cuff Tears Using a Novel Convolutional Neural Network From Magnetic Resonance Image (MRI)", Heliyon, 9, 2023.

DEEPFAKE VIDEO DETECTION: A COMPARATIVE STUDY OF RESNEXT50 AND LSTM ARCHITECTURES

Leili Eslam JamalGolzari*
Gazi University, Graduate School of Natural and Applied Sciences, Computer Engineering, Ankara, Turkey

Aydin Cetin
Gazi University, Faculty of Technology, Computer Engineering, Ankara, Turkey
**(Corresponding author) leili.eslamjamalgolzari@gazi.edu.tr*

ABSTRACT

Recently, the advancement of deepfake technology has hit a point where distinguishing between real and fake content has reached a level it never had been to. In the past, fake content could be easily distinguished by normal people by naked eye, but nowadays doing so is not as easy as it used to be. Creating near to real deepfake content and using it unethically in the versus of politics or social life has brought up consequences in real world which cannot be compensated. Trust and morality have lost their places in people's lives. Therefore, actions should be taken to ease the way for distinguishing what is real and fake. In this regard, with advancement of deepfake technologies, ways to detect them advanced either, and different approaches has been introduced. In this paper, deepfake detection for videos has been implemented on Deep Fake Detection Challenge (DFDC) dataset, using two methods, a Convolutional Neural Network based ResNeXt50 method, alongside Long Short Term Memory (LSTM) method. The accuracy and results of both methods have been found and evaluated with each other. As an outcome, our experiments show that ResNeXt50 model outperforms LSTM model so is a stronger candidate for deployment in most realworld deepfake detection tasks.

Keywords: Deepfake detection, CNNs, RNNs, ResNeXt50, LSTM.

Introduction

The term "Deepfake" is a combination of "deep learning" and "fake". Deepfakes are realistic fake media _videos, images and audio_ which are generated using advanced artificial intelligence techniques, in particular by deep learning models such as Generative Adversarial Networks (GANs) and autoencoders, in a way that it is hard to distinguish whether they are real or fake. Photo manipulation has been existed since the nineteenth century and later, it was applied on the videos. The origin of media manipulation goes back to the 1990s when Adobe Photoshop was used in order to manipulate images. Hence, deepfake technology development was introduced to academia in 1990s by exploring the application of Artificial Intelligence in image processing [1]. In 2014 Generative Adversarial Networks (GAN) was introduced [2]. This can be called the start of creating

deepfake we use these days. The term deepfake was first developed at 2017 by a Reddit user. He created a community to share the deepfake-based content with members of his community. This content was basically faces of celebrities swapped into bodies of actors of videos mostly with pornographic content and videos with face of actor Nicolas Cage swapped into different movies. In 2018, the first desktop application called FakeApp which allowed users to swap faces with each other was launched and deepfake officially made it to the industry.

Deepfake content has grown exponentially, over the years since its emergence. Over the years by development in artificial intelligence and deep learning and generative adversarial networks (GANs) and it has gone through different phases.

1.Before 2017, in the phase of early development, GANs were introduced in 2014 and made it easier for generating realistic images and videos. The technology of deepfake was about face swapping and style transfer. Deepfake was mostly used between academics, and researchers were experimenting new deep learning models such as autoencoders.

2.The second phase of deepfake content is the popularity era which was from 2017 to 2019. In this era non-expert people could utilize deepfake applications on their devices and create deepfake content. In this era the first deepfake detection model was introduced.

3.Third phase is the deepfake content explosion which is between 2020 and 2022. In this era more powerful AI models were introduced which resulted in an increase in the number of deepfake content. Consequently, political misinformation and financial fraud also increased. Governments and social platforms started taking the danger of deepfake threats seriously that resulted in setting some regulations and bans.

4.The fourth phase is the era of advanced AI and realism which is from 2023 to 2025. In this era AI models such as DALLE and Stable Diffusion and Make-A-Video resulted in creating more complicated deepfake content. With increase in the quality of deepfake content and development of text-to-video and deepfake applications, number of cybercrimes and scams also increased. Therefore, as a counter measure, more deepfake detection techniques were presented.

5.Fifth era is the future trends of deepfake. The quality of deepfake content is increasing on a daily basis and it is getting more realistic. Tends in detection techniques may change to behavioral analysis and multimodal AI. Regulations and bans may help stop people who try to take advantage of the deepfake technology. Besides, ethical concerns and legal databases will develop even more.

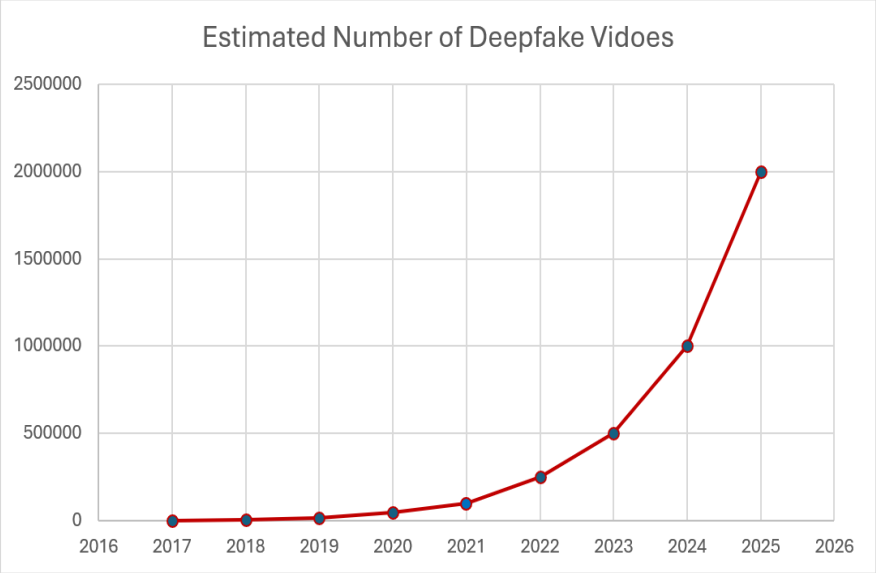


Figure 1: The growth of Deepfake Content

Deepfake are being used as such [3]:

• Entertainment and Media: to create realistic visual effects in movies and video games and in Applications like Synthesia in order to generate personalized avatars for corporate training videos.

• Customer Support: AI-powered voice cloning is applied for automated customer service systems

• Misinformation and Disinformation: These days deepfake is also applied in spreading false information, manipulating public opinion and interfering with election

• Fraud and Cybersecurity: Scams involving deepfake audio or video impersonations have resulted in financial losses exceeding billions of dollars globally. For instance, a deepfake voice scam in 2019 led to a fraudulent transfer of €220,000.

Recently, due to advancement in Artificial Intelligence, deepfakes has become more accessible and sophisticated. Everyone with no special technical ability can have access to open-source tools like DeepFaceLab and mobile applications like Zao and create a top-notch deepfake media. Although this kind of technological development is admirable and could be said that human being is living the golden era of technology, the consequences of this advancements should not be neglected. These days fraudulent schemes using deepfake are rising and the estimation for this loss is $40 billion by the year 2028. Also, disinformation campaigns using deepfake create trust issues in media and it is getting hard day by day what to believe [4-5]. The speed of synthetic media technologies evolution is so high that even techniques such as image forensics and AI-based detection algorithms cannot catch up with that. However, governments, researchers and private sector are in a daily combat trying to develop more effective tools to stop malicious deepfake usage [6]. On the other side, deepfake can be quite useful and have legitimate applications in areas such as education entertainment and accessibility for example creating content for visually impaired people [3].

Deepfake Detection

There are different approaches in detecting deepfake but among all the available approaches, deep learning-based approaches can be considered the most effective ones. Deep learning-based methods such as CNN, RNN, DNN and LSTM. Based on the architecture and techniques used in deep learning approaches, the below mentioned models for detecting deepfake can be referred:

• Convolutional Neural Networks (CNNs): CNNs can be applied in order to detect spatial inconsistency in manipulated videos and images. Traditional CNNs necessitate having additional modules such as multi-attentional architecture or ensemble methods for improving the performance. RNN models are subdivision of CNNs [7, 8].

• Transformer Models: Vision Transformers utilize self-attention mechanisms for detecting subtle artifacts in deepfake content [9].

• Temporal Analysis: models which can analyze temporal inconsistencies in the video frameworks for example things like unnatural blinking method, or irregular facial features or unnatural facial motion, shifts in lighting from one frame to another have proven to be of a great use in deepfake detection [10].

• Biological Signal Analysis: Detecting deepfake can occur utilizing monitoring physiological cues such as pulse signals that can be imbedded in the videos or even eye blinking and shifts in skin tone. Detecting any irregular pattern can be a cue for detecting deepfake [10].

• Alongside deep learning based approaches to detect deepfake, Classical Machine Learning Approaches is another solution. The way these models work is based on the hand-crafted features like text analysis or statistical patterns. These methods are not as computationally intensive as deep learning based approaches [8].

• Additional to the two previous deepfake detection approaches, there are also Hybrid and Fusion methods as below:

• Multimodal Fusion: Combining the visual, textual, audio modalities can increase the accuracy of detection by leveraging the complimentary information [9].

• Adversarial Training: Training the models with adversarial cases can ameliorate the robustness of the model against advanced deepfake detection techniques [9].

Blockchain-based Techniques: It has been proven that using blockchain in verifying the authenticity of the media can be done by tracking down its origin and modifications [8].

Materials and method

Despite the eye-catching advancements in the deepfake detection area, there are a few challenges which are unresolved:

1. Generalization Across Datasets: Models that are trained on particular datasets, usually cannot succeed in generalizing the unseen manipulations because of differences in generation methods [7,10].

2. Real-Time Detection: To guarantee the scalability and efficiency of real-time applications can be cumbersome, specifically on devices with limited accessibility to the resources like mobile platforms [11].

3. Evolving Techniques: The fast speed of progress of Generative Adversarial Networks (GANs) improves the deepfake quality, which makes the detection even harder day by day.

With the improvement of technology, new trends are emerging and some of them are as follows [9, 11]:

• Lightweight Architecture: By improving the current models of edge computing and mobile applications, their accessibility would be high and at the meantime their computational constraints would be addressed.

• Zero-Shot and few-Shot Learning: The goal of these approaches is to ameliorate the model robustness in case they are used in unseen datasets or new manipulation methods without requiring more retraining.

• Temporal-Spatial Features: In order to level up the interpretability and performance of deepfake detection systems, temporal logic can be integrated with spatial features.

Table 1: Comparison of Deepfake Detection Methods

Approach	Key Techniques	Strengths	Weaknesses
Deep Learning-Based	CNN, RNN, DNN, LSTM, Transformers, Temporal Analyses	High accuracy, Learns complex patterns	Requires large datasets, , computationally expensive
Classical Machine Learning	Hand-Crafted Features, Statistical Models	Less computationally expensive	Lower accuracy, Struggles with new manipulation methods
Hybrid/Fusion Methods	Multimodal Fusion, Adversarial Training	Can leverage multiple detection cues	Complexity in models design and training
Blockchain-Based	Media Authentication, Provenance Tracking	Reliable tracking of media alteration	High implementation cost, not a direct detection method

ResNeXt50(Residual Networks with Aggregated Transformations) is a CNN based network which can be used in deepfake detections tasks. LSTM (LongSHortTermMemory) is a RNN based network which can be used in deepfake detection tasks either. In this paper DFDC (DeepFakeDetectionDataset) has been used as dataset. Both ResNext50 and LSTM networks have been trained using 20 epochs to detect the deepfake videos. The performance of the provided solution for detecting deepfake videos have been evaluated at the end.

ResNeXt50

Resnet was introduced in 2015 for addressing the vanishing gradient problem using skip connections (residual blocks) and revolutionized the deep learning world [12]. Networks are trained by these blocks using hundreds of layers through allowing gradients to bypass nonlinear layers, stopping the information flow [13].

Residual mapping F(x) = H(x)-x (1)

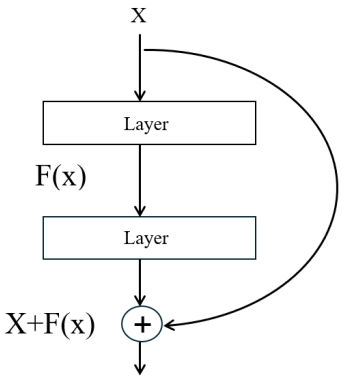


Figure 2: ResNet Architecture

Resnet can be utilized in medical imaging, agriculture and object detection. The challenge in applying Resnet is overfitting. For deep ResNet aggressive augmentation and label smoothing is required [14].

Residual Networks with Aggregated Transformations (ResNeXt50) was introduced in 2016 as an evolution of ResNet50 that could enhance the feature diversity using the cardinality, a dimension that groups convolutional transformations in a network layer. Adding cardinality enables the network to have more diverse and better feature transformation without changing the structure. ResNeXt50 blends the layer stacking approach of VGG/ ResNet with split transform-merge approach of Inception. The main goal here is to enhance the model performance without increasing the number of layers (adding more computation).

Main features of ResNeXt50 is having grouped convolutions by which can split the filters into 32 groups which helps in having better feature extraction with less parameters. ResNeXt50 has the ability to squeeze and excitation (SE) integration which helps it in calibrating features related to channels response using average pooling and adaptive reweighting that can boost the precision 2-3% in tasks related to medical imaging. Another specification of ResNeXt50 is having modular design that can combine identity and convolutional blocks in order to have flexible depth scaling (50-101 layers) [15]. ResNeXt structure is mainly made up of five sections including convolutional blocks and identity blocks. First section is convolutional blocks consisting of three convolutional layers, second section is identity block, consisting of three transformation stages. The SE block which is used as unit of calculation converting inputs into feature maps which is compatible with different types of convolutional neural network architectures and residual networks. In order to enhance the computational load, SE block is placed before performing the sum which elevates the accuracy of ResNeXt50 comparing to ResNet50 model. The suffix in the name ResNeXt is the indicator of total number of layers. ResNeXt50 architecture enables it to manage complex tasks applicating its modular components cardinality dimensions [16, 18].

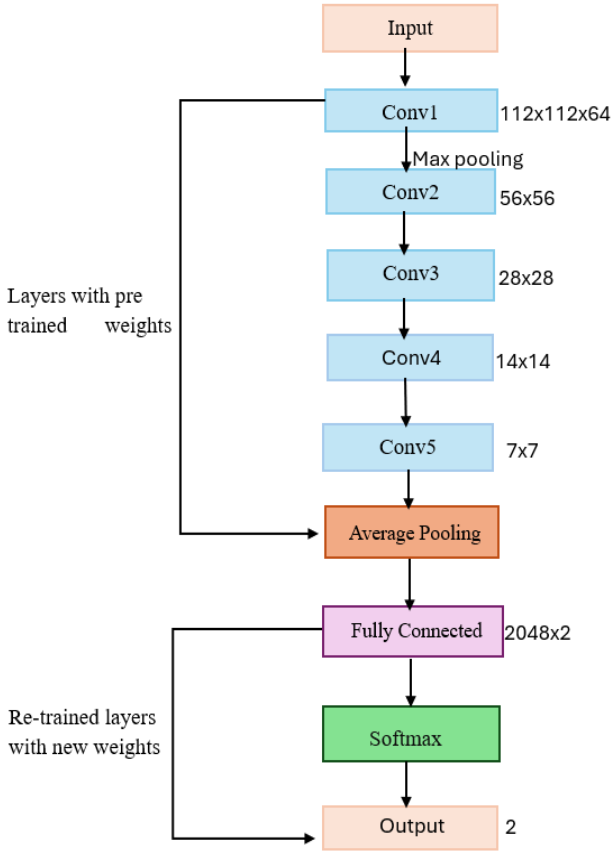


Figure 3: ResNeXt50 Architecture

LSTM

Long short Term Memory networks can be identified as specialized Recurrent Neural Network (RNN) that are designed to manage long term dependencies in sequential data through diminishing vanishing gradient problems using gated memory structures. The structure of LSTM is basically based on the memory blocks with three gates and one cell state. The first gate is called Forget Gate that settles which information should be deleted from the previous cell state through sigmoid activation. Second one is called Input Gate which controls the newly added information to the cell state using a combination of sigmoid and tanh activation. The third gate is Output Gate that based on the cell state and the previous hidden state, adjusts the final hidden state output. The cell state works as persistent memory that makes retrieving information possible after several sequences. In comparison to RNNs, LSTMs can manage memory better since they use cell state and gates for long term retention but RNNs rely on hidden states. Furthermore, LSTMs are better at sequence handling. They can handle dependencies which are composed of more than 1000steps, but RNNs efficiency is for short sequences. LSTM can be applied in speech recognition , time series analysis , Natural Language Processing (NLP), video captioning, healthcare data analysis and efficiency innovations. The down point of LSTM is that they have high computational overhead and their interpretability is low [19-22].

The way LSTM works can be explained as follows: In a sequence of words $(w_1, w_2, w_3, \ldots, w_n)$,the state of LSTM at time step t is declared as (h_t, c_t), is the hidden layer and is cell state. First the input vector and previous state(h_{t-1}, c_{t-1}) is received by LSTM. Then it is decided by the forget gate () that what information should be discarded from cell state. A number between 0 and 1 for each number in in cell state (c_{t-1}) by input vector and previous hidden state. Number 1 here is representing "completely keep it" and number 0 is representing "completely delete it".

$$\text{Forget Gate}(f_t) : \sigma(W_f.[h_{t-1}, x_t] + b_f) \tag{2}$$

Then, in order to decide what information should be stored in the cell state by input gate () which is composed of two parts. The first part is a sigmoid layer named "input gate layer" that decides the values which will be updated and second part is tanh layer which generates the new candidate values (ct~) vector and these values could be added to the state.

$$\text{Input Gate: } i_t = \sigma(W_i.\left[h_{(t-1)}, x_t\right] + b_i) \tag{3}$$

$$\text{CellStateUpdate:} C_t^{\sim} = \tanh(W_c.\left[h_{(t-1)}, x_t\right] + b_c \tag{4}$$

Then the old cell state) will be updated to the new cell state (). The old cell state will be multiplied by and then the new candidate values will be added.

$$\text{FinalCellState: } c_t = f_t \times c_{(t-1)} + i_t \times c_t^{\sim} \tag{5}$$

Then the output will be decided. First a sigmoid layer will run and later the cell state will be put through tanh and will be multiplied by sigmoid gate output.

$$\text{OutputGate:} o_t = \sigma(w_o.\left[h_{(t-1)}, x_t\right] + b_0) \tag{6}$$

$$\text{HiddenState:} h_t = o_t \times \tanh\,(c_t) \tag{7}$$

Databases

Apart from the algorithms used in detecting deepfake, the datasets can also play an important role in the verge of training and evaluating the presented detection methods. Some of the datasets that can be used in this regard are:

•Face Forensics++: A standard dataset with videos which are manipulated and are created using different methods

•DFDC (Deep Fake Detection Challenge): An extensive dataset which is designed to benchmark the state-of-the-art techniques.

•Celeb-DF: Deepfake videos with high quality that test the current detection techniques [8,10].

Table 1 shows the Datasets that are commonly used for Deepfake Detection. In this study, the largest dataset, DFDC, was used for training and test of the models. DFDC full data set consists of 124k videos.

Table 3: Datasets for Deepfake Detection

Dataset	Year	Number of real videos	Number of fake videos
Face Forensics++	2019	1000	4000
DFDC	2020	23,654	104,500
Celeb-DF	2019	590	5,639

Evaluation Metrics

In order to evaluate and compare the performance of LSTM and ResNeXt50 models some experiments were conducted on the results. The evaluations were performed using standard classification metrics, TP (true positives), TN (true negatives), FP (false positives), and FN (false negatives) were defined based on model predictions which were derived from the confusion matrix. These metrics are as follows: Accuracy, Precision, Recall, F_1 and Score. These evaluations provide an insight into the strengths and weaknesses of both the models in predicting accurately and minimizing false positives and negatives. The confusion matrix also allowed detailed analysis of the actual prediction mistakes of the models, further informing their usability in practical deepfake detection tasks. Precision is model's positive classifications which are in fact positive. high precision declares having more relevant results.

Accuracy is a metrics that declares how often a model is predicts the outcomes that are correct.

Accuracy = (8)

Precision is a metric that shows the quality of a positive predictions made by the model.

Precision = (9)

Recall is a metric which can measure how good a model can identify positive instances among all the actual positive instances.

Recall = (10)

Score is the harmonic mean of precision and recall for measuring the accuracy of the model. It is a number between 0 and 1. Best performance for the model is shown by Score of 1 and poor performance is showed by 0.

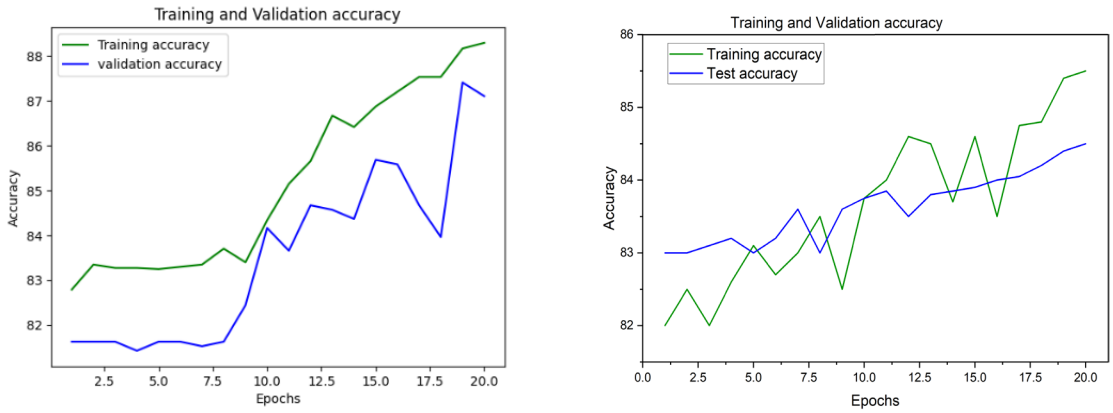
= (11)

Results and findings

While the spatial information is learnt by convolutional neural networks (CNNs), long short-term memory (LSTM) is utilized to evaluate temporal sequences to address deepfake artifacts' multi-dimensional nature. CNN component extracts spatial features from individual video frames, with its main focus on textures, edges, and facial anomalies (e.g., unnatural eye blinking or skin textures). Regarding spatial-temporal feature fusion, CNN extracts frame-level spatial patterns including abnormal facial textures, inconsistency in lit shadows and misshaped effects in facial areas with application of its architecture. Fur-

thermore, CNN detects subtle facial abnormalities by using convolutional filters which are customized to facial landmark areas such as eyes and mouth. In the other hand, LSTM inspects the time coherence for 30-50 frame sequences in order to detect abnormal blinking pattern, abnormal head movements, abnormal changes in the lighting throughout the frames and process features extracted by CNN sequentially in order to model dynamic facial expressions.

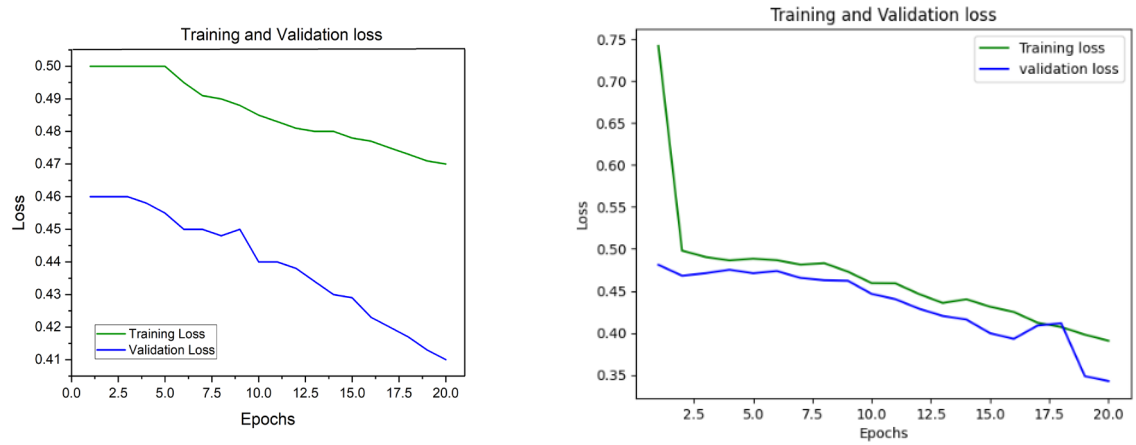
In order to compare the performance of both networks in detecting deepfake videos, evaluation metrics such as Accuracy, Loss and Confusion Matrix have been applied. The results show that the accuracy of the ResNeXt50 model is almost 85% and the accuracy of LSTM model is approximately 87%. Figure 4 shows the training performance based on the accuracy throughout the epochs. The accuracy trends, shown in Figure 4a-b, illustrate the correctness of predictions. The accuracy trends demonstrate the superiority of LSTM over the ResNeXt50.



a b

Figure 4: a) LSTM Accuracy, b) ResNeXt5o Accuracy

Figure 5 shows the loss trend of both LSTM and ResNeXt50 models, declaring that loss of LSTM model has been much less than ResNeXt50 model illustrating that it had a better performance over the course of training.



a b

Figure 5: a) ResNeXt50 Loss, b) LSTM Loss

Confusion matrix can help to understand the overall accuracy and distribution of misclassifications. The confusion matrix's diagonal represents correct classifications. The off-diagonal represents misclassifications of each class. According to the information from confusion matrix, Precision for this LSTM model is 0.96, Recall is 0.3 and F1Score is 0.45. in the other hand these metrics for ResNeXt50 model are as follows: Precision for ResNeXt50 is 0.98, Recall is 0.86 and F1Score for this model is 0.91. In our experiments on the DFDC test set, the LSTM model achieves a very high precision of 96 percent—meaning only 4 false positives for every 100 videos flagged—but its recall is just 30 percent, so it fails to identify 70 percent of actual deepfakes, yielding an F_1score of 0.45. By contrast, ResNeXt50 boasts a precision of 98 percent (only 2 false alarms per 100 flags) and dramatically higher sensitivity, detecting 86 percent of deepfakes; its combined F_1score of 0.91 more than doubles that of the LSTM. In concrete terms, if you run 1000 videos through each model (assuming a balanced mix), the LSTM would correctly catch only 300 deepfakes while missing 700, whereas ResNeXt50 would catch 860 and miss just 140. This stark difference in falsenegative rate and the modest gain in falsepositive rate show that ResNeXt50 not only maintains extremely high confidence in its positive predictions but also offers a far more reliable detection rate, making it statistically and practically superior for realworld deepfake screening.

Table 4: Performance of LSTM and ResNeXt50 models

Precision	Recall	F_1-score	Model
0.96	0.3	0.45	LSTM
0.98	0.86	0.91	ResNeXt50

A better context of the performance of our models can be provided when we place them against state-of-the-art deepfake detection methods. For instance, Al-Dulaimi and Kurnaz [23] presented a hybrid CNN-LSTM model with an F_1-score of 0.87 on video datasets, demonstrating a good balance between temporal and spatial features. Similarly, Tipper et al. [24] utilized a CNN-LSTM model and observed precision and recall values which were more than 90% on controlled datasets. Our ResNeXt50 model achieves an F_1-score of 0.91, outperforming these hybrid approaches, especially in recall (86%), which is crucial for reducing false negatives in real-world applications. On the hand, our baseline LSTM model falls far behind with an F_1-score of only 0.45, confirming findings in studies such as Saikia et al. [28] that vanilla(basic) LSTMs have difficulty extracting spatial features in video data.

Furthermore, hybrid models like the model presented by Suresh et al. [25] reported F_1-scores between 0.85 and 0.89 depending on the dataset, indicating the growing reliability of CNN-based approaches when combined with sequential modeling. Compared to these, ResNeXt50 alone—without temporal layers—shows that deeply optimized spatial networks can achieve better performance comparing complex hybrid models, likely for having better generalization and training stability. Our findings are also consistent with the REFERENCES presented by Pei et al. [6], who emphasizes on the importance of robust feature extraction in achieving high recall and minimizing the deepfake threats.

Overall, our results reinforce the literature's trend toward using advanced CNN architectures (e.g., ResNet, InceptionResNetV2) as a backbone for deepfake detection. How-

ever, our experimental comparison reveals that ResNeXt50, despite not being a hybrid model, delivers state-of-the-art performance, and could serve as a strong candidate for future research focusing on efficiency and deployment in real-time detection systems.

Conclusion

In this study, we evaluated and compared the performance of two deep learning architectures, CNN-based ResNeXt50 and RNN-based LSTM, on the DFDC dataset. Performance metrics including accuracy, loss, and confusion matrix analysis, precision, recall and F1score were utilized for a comprehensive assessment. Our results indicate that the ResNeXt50 model outperformed the LSTM model. Additionally, a comparative analysis of loss trends and confusion matrices further supports the superior performance of the ResNeXt50 model. Therefore, we conclude that the CNN-based ResNeXt50 model is more effective for deepfake detection tasks on the DFDC dataset compared to the RNN-based LSTM.

Declaration of Interest: The authors declare that there are no conflicts of interest regarding the publication of this paper.

Author Contributions: Conceptualization: [Leili Eslam JamalGolzari], [Aydın Çetin] Methodology: [Leili Eslam JamalGolzari], [Aydın Çetin] Software Development and Implementation: [Leili Eslam JamalGolzari] Validation and Experimental Analysis: [Leili Eslam JamalGolzari] Formal Analysis and Visualization: [Leili Eslam JamalGolzari], [Aydın Çetin] Writing—Original Draft Preparation: [Leili Eslam JamalGolzari] Writing—Review and Editing: [Leili Eslam JamalGolzari], [Aydın Çetin] Supervision and Project Administration: [Aydın Çetin].

REFERENCES

[1] Meredith Somers, "Deepfakes, explained." [Online]. Available: https://mitsloan.mit.edu/ideas-made-to-matter/deepfakes-explained

[2]Y. B. Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, "Generative Adversarial Nets," Adv. Neural Inf. Process. Syst. 27 (NIPS 2014), pp. 3063–3071, 2014, doi: 10.1109/ICCVW.2019.00369.

[3]I. Kalpokas and J. Kalpokiene, Deepfakes: A Realistic Assessment of Potentials, Risks, and Policy Regulation. 2022. [Online]. Available: https://doi.org/10.1007/978-3-030-93802-4_7

[4]A. Kishore, "What is a Deepfake and How Are They Made?" [Online]. Available: https://www.online-tech-tips.com/what-is-a-deepfake-and-how-are-they-made/

[5]F. Juefei-Xu, R. Wang, Y. Huang, Q. Guo, L. Ma, and Y. Liu, Countering Malicious DeepFakes: Survey, Battleground, and Horizon, vol. 130, no. 7. Springer US, 2022. doi: 10.1007/s11263-022-01606-8.

[6]G. Pei et al., "Deepfake Generation and Detection: A Benchmark and Survey," 2024, [Online]. Available: http://arxiv.org/abs/2403.17881

[7]L. Stroebel, M. Llewellyn, T. Hartley, T. S. Ip, and M. Ahmed, "A systematic literature review on the effectiveness of deepfake detection techniques," J. Cyber Secur. Technol., vol. 7, no. 2, pp. 83–113, 2023, doi: 10.1080/23742917.2023.2192888.

[8]M. S. Rana, M. N. Nobi, B. Murali, and A. H. Sung, "Deepfake Detection: A Systematic Literature Review," IEEE Access, vol. 10, pp. 25494–25513, 2022, doi: 10.1109/ACCESS.2022.3154404.

[9]G. Gupta, K. Raja, M. Gupta, T. Jan, S. T. Whiteside, and M. Prasad, "A Comprehensive Review of DeepFake Detection Using Advanced Machine Learning and Fusion Methods," Electron., vol. 13, no. 1, pp. 1–27, 2024, doi: 10.3390/electronics13010095.

[10]A. Naitali, M. Ridouani, F. Salahdine, and N. Kaabouch, "Deepfake Attacks: Generation, Detection, Datasets, Challenges, and Research Directions," Computers, vol. 12, no. 10, pp. 1–26, 2023, doi: 10.3390/computers12100216.

[11]A. Raza, K. Munir, and M. Almutairi, "A Novel Deep Learning Approach for Deepfake Image Detection," Appl. Sci., vol. 12, no. 19, 2022, doi: 10.3390/app12199820.

[12]K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit., vol. 2016-Decem, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.

[13]M. S. Ebrahimi and H. K. Abadi, "Study of Residual Networks for Image Recognition," Intell. Comput. - Proc. 2021 Comput. Conf., pp. 754–763, 2021, doi: 10.1007/978-3-030-80126-7_53.

[14]I. Bello et al., "Revisiting ResNets: Improved Training and Scaling Strategies," Adv. Neural Inf. Process. Syst., vol. 27, no. NeurIPS, pp. 22614–22627, 2021.

[15]M. I. Moreno-Lozano, E. J. Ticlavilca-Inche, P. Castañeda, S. Wong-Durand, D. Mauricio, and A. Oñate-Andino, "A Performance Evaluation of Convolutional Neural Network Architectures for Pterygium Detection in Anterior Segment Eye Images," Diagnostics, vol. 14, no. 18, 2024, doi: 10.3390/diagnostics14182026.

[16]E. J. TiclavilcaInche, M. M.-L. Isabel, P. Castañeda, S. Wong-Durand, Alejandra, and Oñate-Andino, "Online and Biomedical Engineering," Int. J. Online Biomed. Eng., vol. 19, no. 9, pp. 122–130, 2023.

[17]M. K. Choi and H. Jung, "Development of Fast Refinement Detectors on AI Edge Platforms," Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics), vol. 12664 LNCS, pp. 592–606, 2021, doi: 10.1007/978-3-030-68799-1_43.

[18]S. Kolekar, S. Gite, B. Pradhan, and A. Alamri, "Predicting Vehicle Pose in Six Degrees of Freedom from Single Image in Real-World Traffic Environments Using Deep Pretrained Convolutional Networks and Modified Centernet," Int. J. Smart Sens. Intell. Syst., vol. 17, no. 1, 2024, doi: 10.2478/ijssis-2024-0025.

[19]A. Ayadi, T. M. D. Truong, C. Frydman, and C. Wemmert, "Explaining LSTM neural networks: A DEVS-based approach," Eur. Model. Simul. Symp. EMSS, vol. 2024-Septe, pp. 1–10, 2024, doi: 10.46354/i3m.2024.emss.035.

[20]D. Lee et al., "Long short-term memory recurrent neural network-based acoustic model using connectionist temporal classification on a large-scale training corpus," China Commun., vol. 14, no. 9, pp. 23–31, 2017, doi: 10.1109/CC.2017.8068761.

[21]S. Hochreiter and J. Schmidhuber, "LSTM paper," Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997.

[22]N. Elsayed, Z. Elsayed, and A. S. Maida, "LiteLSTM Architecture for Deep Recurrent Neural Networks," Proc. - IEEE Int. Symp. Circuits Syst., vol. 2022-May, no. c, pp. 1304–1308, 2022, doi: 10.1109/ISCAS48785.2022.9937585.

[23]O. A. H. H. Al-Dulaimi and S. Kurnaz, "A Hybrid CNN-LSTM Approach for Precision Deepfake Image Detection Based on Transfer Learning," Electron., vol. 13, no. 9, pp. 1–22, 2024, doi: 10.3390/electronics13091662.

[24]S. Tipper, H. F. Atlam, and H. S. Lallie, "An Investigation into the Utilisation of CNN with LSTM for Video Deepfake Detection," Appl. Sci., vol. 14, no. 21, 2024, doi: 10.3390/app14219754.

[25]M. B. Suresh, N. Dinesh, and T. N. Abhishek, "HYBRID RESNET CNN-LSTM FOR DEEPFAKE VIDEO DETECTION," vol. 12, no. 4, pp. 588–596, 2024.

[26]P. Yadav, I. Jaswal, J. Maravi, V. Choudhary, and G. Khanna, "DeepFake Detection using InceptionResNetV2 and LSTM," CEUR Workshop Proc., vol. 3058, pp. 1–6, 2021.

[27]B. As, C. K. Jha, and S. Asopa, "Advanced deepfake detection using hybrid CNN-MLP models with feature extraction from facial landmarks," vol. 5, no. 2, pp. 136–149, 2024.

[28]P. Saikia, D. Dholaria, P. Yadav, V. Patel, and M. Roy, "A Hybrid CNN-LSTM model for Video Deepfake Detection by Leveraging Optical Flow Features," Proc. Int. Jt. Conf. Neural Networks, vol. 2022-July, 2022, doi: 10.1109/IJCNN55064.2022.9892905.

DESIGN OF A PATIENT MONITORING DEVICE FOR EFFECTIVE INHALER MEDICATION USAGE

Erhan Sesli*
Karadeniz Technical University, Of Technology Faculty, Department of Electronics and Telecommunication Engineering, Trabzon, Türkiye

Hamdi Eren Cagilci
Karadeniz Technical University, Of Technology Faculty, Department of Electronics and Telecommunication Engineering, Trabzon, Türkiye

Mustafa Bagatur
Karadeniz Technical University, Of Technology Faculty, Department of Electronics and Telecommunication Engineering, Trabzon, Türkiye
**(Corresponding author) erhansesli@ktu.edu.tr*

ABSTRACT

In this paper, a smart inhaler monitoring device is introduced for designing to improve medication adherence of inhalers among asthmatic patients. Asthma is a chronic and prevalent respiratory disease globally, and inhaled therapy is the main part of the treatment strategy. The correct administration of inhaler medication and adherence to treatment are, however, often poor, which affects asthma control in a negative way. A smart inhaler monitor device consisting of a small electronic module that could be mounted on the inhaler drug was developed in this work and its prototype was fabricated. The device has an accelerometer sensor to detect shaking of the medicine in the correct manner and administration of the same at the correct angle, an ESP8266 Wi-Fi module to provide data communication to patient relatives and healthcare professionals, and an OLED and warning LEDs display to provide user feedback. The performance of the developed device was verified using various test cases under lab conditions. It is proven that the accelerometer sensor can measure correctly the proper shaking time and angle of the inhaler medicine. The device information was successfully transmitted to a cloud platform (ThingSpeak) using the Wi-Fi module and the patient's relative/doctor can access it remotely. It was observed that the OLED screen and warning LEDs display medication instructions and device status to the user in an understandable way. The results of this study show that the smart inhaler tracking device has the potential to enhance inhaler medication adherence in asthma patients and follow-up of patients. The device is thought to be capable of assisting patients to improve medication use skills, increase adherence to treatment, and provide healthcare providers with better patient management. In future works, the device's performance on patients in the clinical setup and patient satisfaction will be evaluated and its commercial potential will be decided.

Keywords: Smart inhaler, asthma monitoring, medication use compliance, internet of things.

Introduction

Asthma is a chronic respiratory disease that affects more than 300 million people worldwide and poses a significant socioeconomic burden [1]. The use of inhaler medication is one of the most com- monly used treatment methods to keep asthma symptoms under control and improve quality of life [2]. The effectiveness of inhaler therapy is greatly increased when patients use correct usage techniques and adhere to treatment regularly. However, studies have shown that a significant number of asthma patients use inhalers incorrectly and do not adhere to treatment [2]. For example, Hesso et al. 2020 examined inhaler use in patients with COPD or asthma in the UK. Using an electronic inhaler assessment device (EMD), they found that the actual adherence rate was low at 42.7% and only 17% of patients used the medication at the right time and with the right technique. These results reveal that technical errors in use are common, highlighting the importance of objective assessment of inhaler use [3]. This makes asthma control more difficult, increases emergency room visits and hospitalization rates, and thus increases healthcare costs [4].

Various approaches have been developed to improve medication use in asthma patients. These in- clude patient education programs, use of visual and audio materials, telehealth services, and mobile- based reminder systems [5]. However, these approaches may be insufficient to ensure long-term patient participation and make it difficult to objectively monitor medication use data. Therefore, technological solutions are needed to ensure patients use inhaled medications correctly and to support patient follow-up with objective data.

In this study, a smart inhaler tracking device was developed to track the use of inhaled medication by asthma patients, to verify that the medication is used with the correct technique, and to inform the relatives of the patient. The developed device consists of a compact electronic module that can be in- tegrated onto the inhaler. This module includes an accelerometer sensor that detects the shaking of the medication and its use at the correct angle, an ESP8266 Wi-Fi module that transmits data wirelessly, and an OLED screen that provides instructions and status information to the user.

The main contribution of this study is that it offers an innovative approach to improve patient follow- up by monitoring inhaler medication use with objective data. In addition, it provides remote follow- up thanks to the feature of transmitting medication use information to patient relatives. In this way, medication use errors of patients will be minimized and treatment processes will be managed more effectively.

In the following sections of the paper, studies in this scope in the literature were examined; the elements to be considered in the hardware and software design of the smart inhaler tracking device were listed in the method section, and prototype development processes were explained. In the implementation studies and system tests section, laboratory tests and patient use simulations were discussed in detail. In the conclusion section, the contributions of the system developed in line with the findings obtained to asthma management and possible future improvements were discussed.

Literature Review

Studies in the literature on asthma management and the effectiveness of inhaler medication use reveal the diversity of technological and clinical approaches in this area. It is a fact that the correct use of inhalers in the treatment process of asthma patients plays a critical role in controlling the disease and improving quality of life. We can summarize the studies conducted in the literature over the last decade as follows.

Chan et al. in their 2015 study examined the use of inhaler compatible electronic monitoring devices (EIMs) in the clinical setting, and suggested that EIMs support the clinical decision-making process by providing accurate and detailed compliance data. The article discusses the characteristics of different EIMs and the useful features of these devices for clinicians and patients. In this sense, the article is more of a guide to help physicians use electronic devices (EIMs) used to measure inhaler use more effectively [6].

In their 2016 study, Sulaiman et al. aimed to develop a new method to measure adherence to inhalers used in asthma treatment. Unlike existing methods, this new method aimed to evaluate adherence to treatment more accurately by taking into account the intervals between doses and errors in inhaler use. While Sulaiman et al. developed a method that analyzes the timing and technique of inhaler use in asthma treatment in more detail, this study offers a more applicable solution that directly shares medication use data to increase patient follow-up and adherence to treatment [7].

In their 2017 study, Grossman et al. designed a mobile application-based personalized intervention to increase adherence to asthma treatment in disadvantaged youth. They aimed to improve adherence to daily medication regimen by combining electronic monitoring devices with informative and persuasive elements. Grossman et al. increased adherence to treatment in youth with asthma by providing person- alized treatment on a mobile platform using a rarely used treatment and monitoring device. However, while Grossman et al.'s study focused on motivation and persuasion, this study emphasizes monitoring medication use and ensuring compliance with instructions [8].

In their 2021 study, Lavanya et al. aimed to develop a low-cost, wearable asthma monitoring system and smart inhaler. The system continuously monitors the patient's heart rate and environmental condi- tions (humidity, smoke, temperature). This data is uploaded to a cloud server that a doctor and caregiver can access, and when the system detects an adverse asthma environment, it sends an alert message to the patient. Both the monitoring device and the smart inhaler focus on improving asthma treatment by improving overall lung function, reducing daily symptoms, and preventing acute asthma attacks. While Lavanya et al. presented a low-cost wearable device and smart inhaler system for monitoring and treating asthma, this study aims to directly increase treatment adherence and facilitate remote patient monitoring by developing a system that monitors the correct use (angle and shake) and timing (record and track) of the medication [9].

In their 2022 study, Takano et al. proposed an inhalation monitoring device using an inertial mea- surement unit (IMU) to increase the correct use of inhalers used in the treatment of asthma and COPD. In the study, IMU is used to measure the patient's inhalation

movement and faulty inhalation is determined by comparing the measurement data with correct use. Considering the results they obtained, Takano et al. stated the usefulness of the proposed device, but emphasized that further improvements and clinical studies are needed. While Takano et al. proposed an IMU-based device to measure inhalation technique, this study offers a more practical system that monitors the correct use of the drug (angle and shaking) and timing (recording and tracking), and also allows information sharing with the patient's relative and the doctor, thus facilitating remote patient monitoring [10].

In their 2023 publication, Eikholt et al. examined how current electronic monitoring devices (EMDs) assess inhalation technique, their ability to provide feedback to patients, and the integration of these de- vices into clinical practice. In their study, it was understood that most EMDs can detect inhalation errors but support patient monitoring to a limited extent and generally use short-range connections such as Bluetooth. This study, unlike existing systems, was developed to both ensure correct use of the drug (with additional sensor support such as correct angle and shaking control) and to facilitate patient mon- itoring (remote data access through WiFi and cloud integration). In this respect, a more comprehensive and patient-centered system is proposed that increases patient compliance, offers instant monitoring, and provides continuous data to healthcare professionals [11].

Methodology

In this section, followed methodological approach to achieve for this study is explained in detail. The methodology covers a systematic process starting from requirements analysis, system design, hardware and software development, prototyping and finally testing stages.

Requirements Analysis

Literature review highlights the critical role of the inhalers and correct usage of inhalers by the patients. However, current studies show that there are common errors in inhaled medication use and poor adher- ence to treatment. This complicates asthma control, increases emergency room visits and hospitalization rates. Therefore, a smart inhaler tracking device must meet the following basic requirements:

1.The device must reliably detect whether the inhaler is shaken correctly and used at the correct angle. This should be achieved by determining the sensitivity of the accelerometer sensor and the correct threshold values.

2.The times of medication use (day, time) should be recorded and this data should be viewable by the patient or the patient's relative. This can help to encourage patients to take their medications regularly.

3.The device should provide the user with instant feedback on medication use technique. This feed- back should be presented visually (OLED screen and LEDs).

4.Medication use data should be shared wirelessly with the patient's relative or healthcare profes- sionals. This facilitates patient follow-up and enables early detection of adherence problems.

5.The device should not make the inhaler difficult to use and should be easily integrated into daily life. Its compact size, light weight and long battery life should meet this requirement.

6.The cost of the device should be kept as low as possible to ensure its accessibility to the masses

The requirements listed are guiding points in the design and development of the smart inhaler track ing device. The device should be both technically reliable and user-friendly, which will contribute to improving the use of medication by asthma patients and increasing their compliance with treatment.

System Design

This section describes the hardware and software design details of the smart inhaler tracking device. The system consists of three main parts: (1) Sensors and Data Collection, (2) Data Processing and Feedback,(3) Wireless Communication and Data Transfer. The working mechanism of the system is shown below.

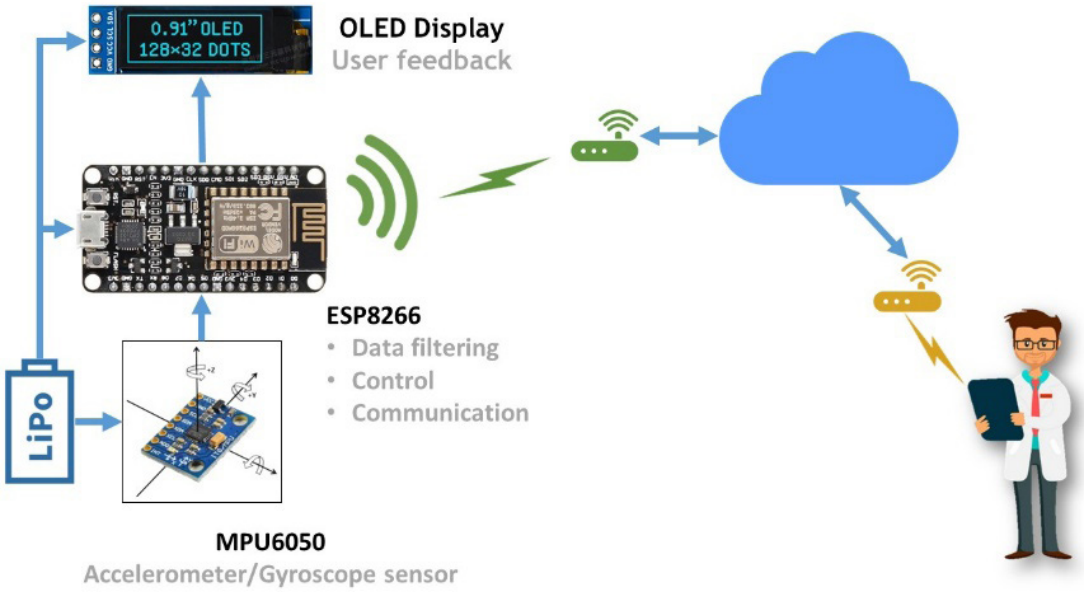


Figure 1: General system architecture

Data Collection

MPU6050 accelerometer sensor has been used to detect shaking and angle, being the basic param- eters of inhaler use. It is capable of detecting acceleration and angular velocity along three axes (x, y, z). For detecting shaking, whether motion of the shaking of the inhaler is long enough duration and intensity is examined by analyzing the accelerometer data. Variations in acceleration beyond a particular threshold value over a particular time period (10s) are referred to as shaking. In angle detection, whether the inhaler should be administered at the correct angle or not is determined based on the inclination angle calculated using the values from the accelerometers. There is a defined tolerance value in the correct use angle.

Data Processing and Feedback

ESP8266 is used as the control and communication unit. ESP8266 is a low-cost, Wi-Fi-capable microcontroller module and is widely used in IoT applications. Because of its built-in TCP/IP stack, it can provide wireless communication with network devices. Because of its small size, low power requirement and compatibility with microcontrollers such as Arduino, it becomes the best choice for reliable data transfer in different smart systems such as sensor networks, home automation systems, and remote monitoring systems. MPU6050 sensor data is obtained by this method via the ESP8266 microcontroller. ESP8266 and MPU6050 between them are communicated using I2C. A moving average filter is used to smooth the sensor data noise in the data filtering stage.

A moving average filter creates an output signal $y(n)$ by averaging the previous values (N) over a given window length for each sample $x(n)$ of a signal. The output signal of a moving average filter is defined as in Equation 1.

$$y(n) = \frac{1}{N} \sum_{k=0}^{N-1} x(n-k) \quad (1)$$

Here, $x(n)$ is the input signal, $y(n)$ is the filtered output signal, (N) is the window length (number of samples), and k is the index of past samples. When performing shaking and holding angle analysis with

MPU6050, determining the sampling period is a critical issue in terms of both accuracy and efficiency. The moving average filter suppresses high-frequency noise and smoothes the sensor data, producing a more stable output. However, it can also filter out sudden changes, so choosing a very large N may not be correct. On the other hand, choosing a small N value reduces the noise to a small extent. The sampling rate of the MPU6050 accelerometer is 1kHz [12]. Considering the above-mentioned issues, the value of N is selected as 10. The algorithm embedded in ESP8266 evaluates whether the drug use is correct or not based on the shaking and angle data. The algorithm result is presented to the users as visual feedback via the I2C communication established between ESP8266 and the OLED display. In addition, the status of the device (connection status, battery level, etc.) is also shown on the screen.

Wireless Communication and Data Transfer

Wireless communication and data transfer are performed using the ESP8266's WiFi module. Peri- odically transferred inhaler usage data is transferred to a cloud platform and its logs are kept. Thus, the patient's relative or doctor's appropriate use of the device or the necessary dose data can be observed.

Hardware Development

At this stage, firstly the circuit connections of the system components are created. The circuit schematic of the created system is shown in the figure below.

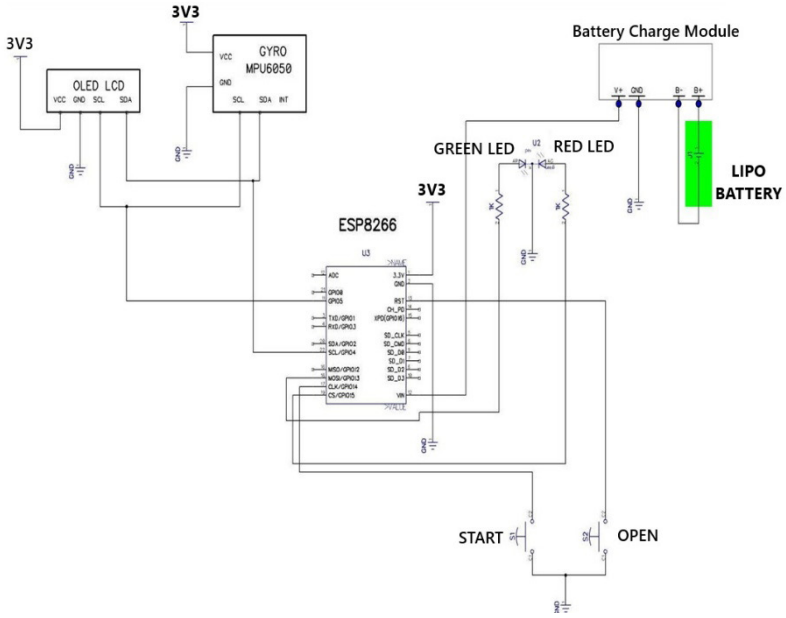


Figure 2: Circuit schematic of the system

A box has been designed to fit the medication inhaler and house all the electronic components. The box design has been optimized to ensure patient comfort and not hinder the use of the inhaler. The design has been printed using a suitable material with a 3D printer.

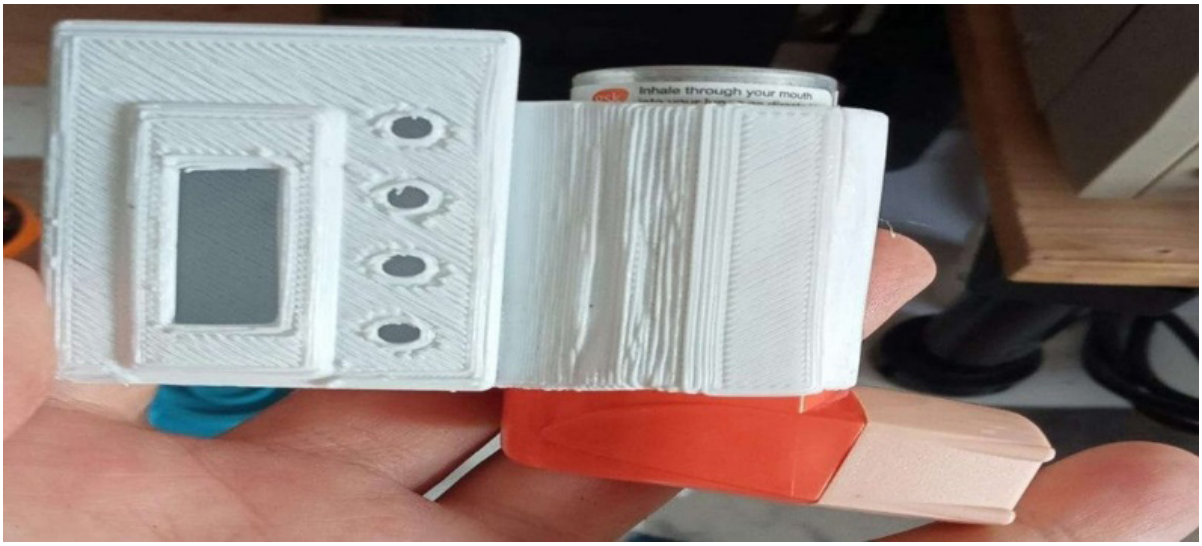


Figure 3: 3D printed inhaler box

Software Development

This system connects the ESP8266 module to a predefined wireless network, starting with the start button activation; the connection status is periodically monitored. After the connection is successfully established, users are asked to shake the device, during which axial acceleration data is obtained using the MPU6050 inertial measurement unit, and acceleration changes exceeding a predetermined threshold value are evaluated as shaking. Then, the device is asked to be held horizontally, and again, after the MPU6050 data is

purified from noise by means of a moving average filter, the flat position is confirmed by comparing it with the gravity vector within the specified tolerance range. Finally, the sensor data is transmitted to a remote database (ThingSpeak) via WiFi with a timestamp, and a retry mechanism is activated in case of data transmission failure to prevent data loss. The main reasons for utilizing ThingSpeak are that the platform is low-cost, open-source, and uniquely designed for Internet of Things (IoT) purposes. ThingSpeak offers an ideal solution for monitoring and analyzing inhaler usage data of asthma patients in real time. Having a simpler interface compared to other cloud platforms has accelerated our prototype development process and minimized costs. The entire algorithm of the system mechanism is shown in Figure 4.

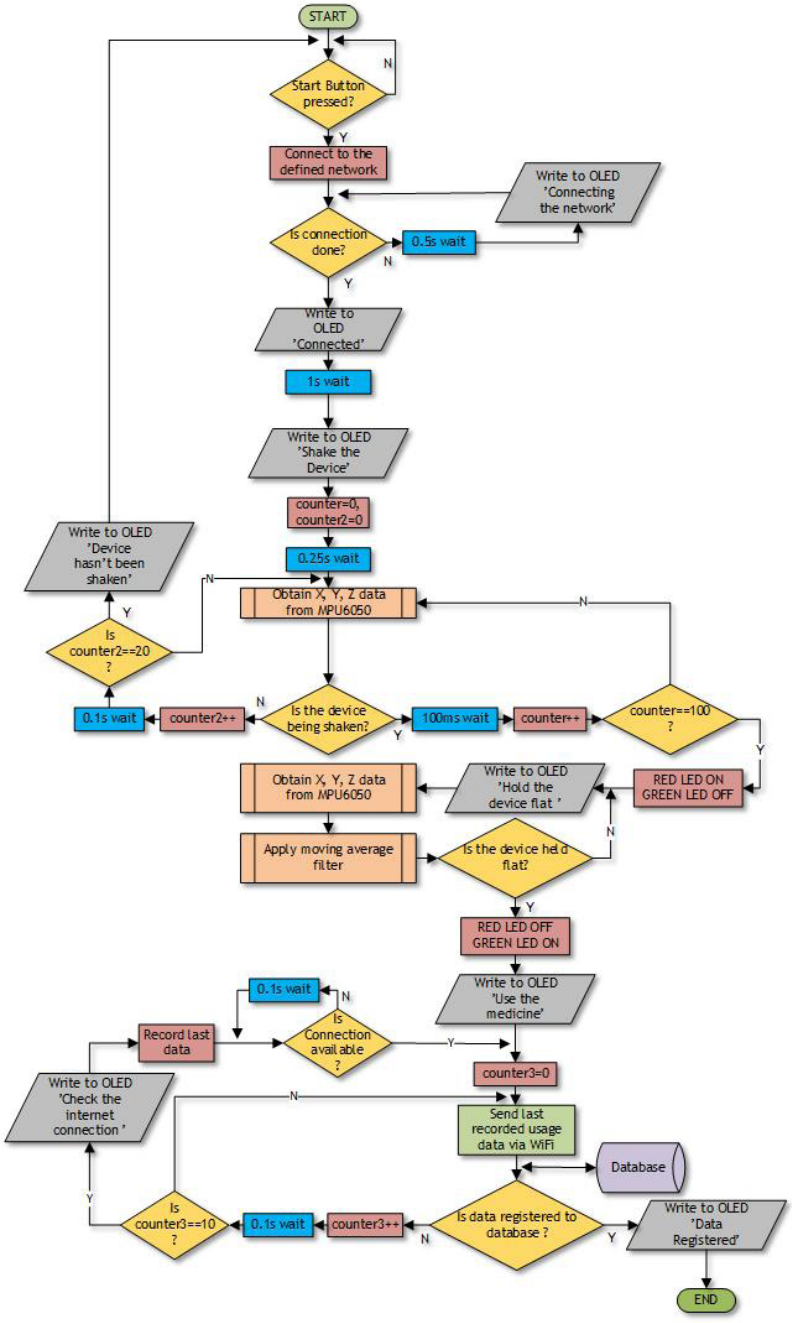


Figure 4: Entire system working algorithm

Implementation Studies and System Tests

In the application phase, firstly, the necessary circuit connections were made and the circuits were placed in the device case taken from the 3D printer. This phase is shown in Figure 5.

Testing the Device's Network Connection

Images showing the device's network connection tests are shown in Figure 6. The WiFi library required for the use of the ESP8266 WiFi module has been added to the code and the operation of the code that will establish the WiFi connection has been tested.

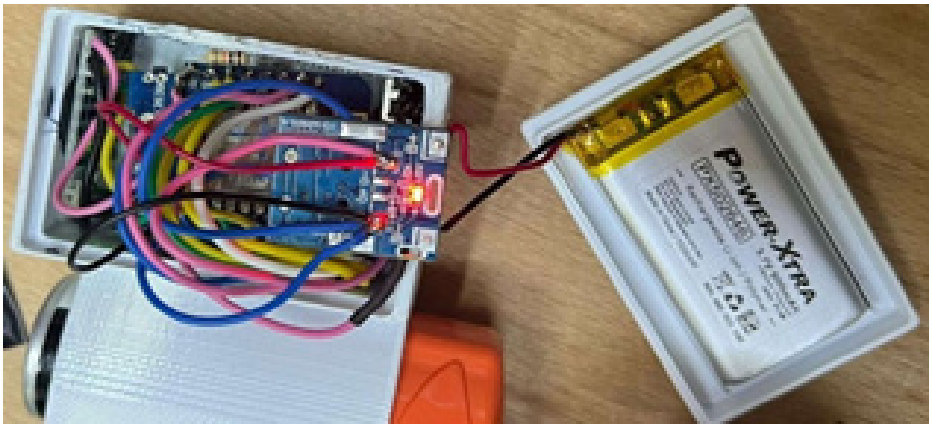


Figure 5: Integration of circuits into the case

Figure 6: WiFi connection tests

Inhaler Shake Process Test

A visual showing the process of properly shaking the inhaler is shown in Figure 7. During this process, when the device is shaken continuously for 10 s, it transitions to the next process of keeping the device in the correct position.

Figure 7: Device shaking control

Device correct grip control test

At this stage, proper and necessary shaking is now performed and the device is expected to be held at the correct angle. In the left image in Figure 8., while the user is given the command to hold the device in a straight position, the red warning LED on the right side is on, indicating that it is not in a straight position. In the right image, when the device is held correctly, the green LED is on and the information screen shows that the drug can be used.

Figure 8: Device correct grip test

Device Recording Tests

In the last stage, after the usage is completed, the data is recorded and the data log is observed on the doctor or the patient caregiver screen. These test stages are shown in Figure 9. In (A), the process of transferring the status information to the other party is seen, while in (B), the feedback sent to the user after being recorded is seen. In (C) and (D), the usage details observed by the doctor or the patient's relative can be observed with date tags.

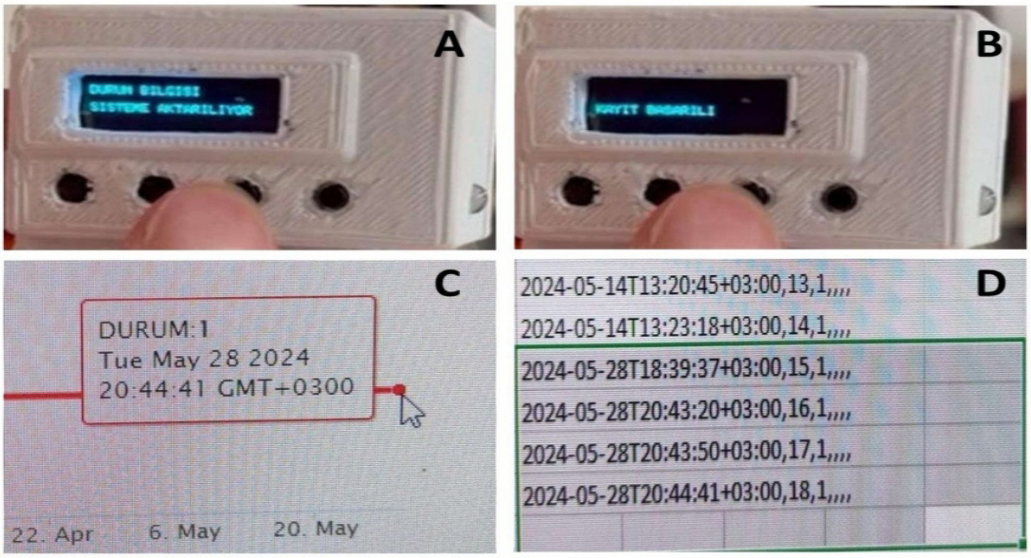


Figure 9: Device record/control test

Conclusion

This study presents the design and prototype of a smart inhaler tracking device developed to improve inhaler medication use by asthma patients. The basic requirements determined at the beginning of the study, such as the ability to use the medication at the correct angle and by shaking for a sufficient period of time, recording the times of use, providing instant feedback to the user, sharing data with the patient's relative/doctor, and the device being portable and low-cost, were successfully met in the developed de- vice. Tests conducted in a laboratory environment showed that the device can reliably detect the correct shaking and angle of the medication thanks to its accelerometer sensor, transfer data wirelessly thanks to the Wi-Fi module, and provide understandable feedback to the user thanks to the OLED screen and warning LEDs.

These results support the potential of the smart inhaler tracking device to increase medication com- pliance in asthma patients and improve patient follow-up. Thanks to its compact design and low cost, the device can easily integrate into the daily lives of patients and reach a wide mass of the people. In the following stages, the device's effectiveness on patients in the clinical environment and patient satisfac- tion will be evaluated and its commercial potential will be investigated. Additionally, the compatibility of the device with different inhaler types and the evaluation of its performance in long-term use will be an important focus for future studies.

Declaration of Interest: The authors declare that there is no conflict of interest.

Author Contributions: Hamdi Eren Cagilci and Mustafa Bagatur were responsible for data curation, software development, and hardware implementation of the system. Erhan Sesli led the conceptual development of the study, designed the methodology, conducted the formal analysis, and played a key role in writing the manuscript and supervising the research project and article.

REFERENCES

[1] GBD 2015 Chronic Respiratory Disease Collaborators, "Global, regional, and national deaths, preva- lence, disability-adjusted life years, and years lived with disability for chronic obstructive pulmonary disease and asthma, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015," The Lancet. Respiratory Medicine, 5(9), 691, 2017.

[2] Dudvarski Ilic, A., Zugic, V., Zvezdin, B., Kopitovic, I., Cekerevac, I., Cupurdija, V., and Barac, A., "Influence of inhaler technique on asthma and COPD control: a multicenter experience," Interna- tional Journal of Chronic Obstructive Pulmonary Disease, 11, 2509-2517, 2016.

[3] Hesso, I., Nabhani Gebara, S., Greene, G., Costello, R. W., and Kayyali, R., "A quantitative evalua- tion of adherence and inhalation technique among respiratory patients: an observational study using an electronic inhaler assessment device," International Journal of Clinical Practice, 74(2), e13437, 2020.

[4] Usmani, O. S., Lavorini, F., Marshall, J., Dunlop, W. C. N., Heron, L., Farrington, E., and Dekhuijzen, R., "Critical inhaler errors in asthma and COPD: a systematic review of impact on health outcomes," Respiratory Research, 19, 1-20, 2018.

[5] Van der Kamp, M. R., Hengeveld, V. S., Brusse-Keizer, M. G., Thio, B. J., and Tabak, M., "eHealth technologies for monitoring pediatric asthma at home: scoping review," Journal of Medical Internet Research, 25, e45896, 2023.

[6] Chan, A. H. Y., Harrison, J., Black, P. N., Mitchell, E. A., and Foster, J. M., "Using electronic monitoring devices to measure inhaler adherence: a practical guide for clinicians," The Journal of Allergy and Clinical Immunology: In Practice, 3(3), 335-349, 2015.

[7] Sulaiman, I., Seheult, J., MacHale, E., Boland, F., O'Dwyer, S. M., Rapcan, V., and Costello, R. W., "A method to calculate adherence to inhaled therapy that reflects the changes in clinical features of asthma," Annals of the American Thoracic Society, 13(11), 1894-1903, 2016.

[8] Grossman, B., Conner, S., Mosnaim, G., Albers, J., Leigh, J., Jones, S., and Kenyon, R., "Application of human augmentics: a persuasive asthma inhaler," Journal of Biomedical Informatics, 67, 51-58, 2017.

[9] K. Lavanya, K. Nikitha, S. R. Deepthi, "Low-cost wearable asthma monitoring system with a smart inhaler," 2021 Asian Conference on Innovation in Technology (ASIANCON), IEEE, August 27-29, pp. 1-7, PUNE, India.

[10] Takano, S., Shimono, T., Masaki, K., Fukunaga, K., Kabata, H., Nishie, M., and Hasegawa, A., "An inhalation device with inertial measurement unit for monitoring inhaler technique," IEEE/ASME Transactions on Mechatronics, 27(4), 2204-2211, 2022.

[11] Eikholt, A. A., Wiertz, M. B., Hew, M., Chan, A. H., and van Boven, J. F., "Electronic monitor- ing devices to support inhalation technique in patients with asthma: a narrative review," Current Treatment Options in Allergy, 10(1), 28-52, 2023.

[12] InvenSense. (2013). MPU-6000 and MPU-6050 Register Map and Descriptions, Revision 4.0. Retrieved from https://cdn.sparkfun.com/datasheets/Sensors/Accelerometers/RM-MPU-6000A.pdf

ENHANCING ANCILLARY PRODUCT CLASSIFICATIONS IN THE AVIATION INDUSTRY: A COMPARATIVE ANALYSIS OF MACHINE LEARNING ALGORITHMS

Yavuz Selim Emir
Bahçeşehir University, Graduate School, İstanbul, Turkey

Tamer Uçar*
Bahçeşehir University, Department of Software Engineering, İstanbul, Turkey
**(Corresponding author) tamer.ucar@bau.edu.tr*

ABSTRACT

The escalating rivalry in the aviation sector compels companies to explore innovative strategies to enhance their profitability. One such strategy is the provision of ancillary products in addition to their core services. However, the challenge lies in identifying the right product for the customer who truly requires it. The recent advancements in recommender systems, particularly in the e-commerce sector, prompt an inquiry into their effectiveness in suggesting ancillary products in the aviation industry. This research aims to develop a classification module to recommend suitable ancillary to passengers, utilizing online customer reviews. These reviews, sourced from various websites, were categorized based on their topics using the BERTOPIC topic modeling algorithm. Reviews labeled with ancillary product recommendations by experts were then put into the algorithms to construct classification module. The objective of this research is to compare classification models that are built using different machine learning algorithms. The findings of the study indicate that the model employing Neural networks yielded the highest accuracy results, with a score of 0.85.

Keywords: Recommender Systems, Machine Learning, Deep Learning, SVM, Classification.

Introduction

The aviation sector is fiercely competitive, forcing companies to continually upgrade their service quality to attract and retain customers. With numerous options available, airlines must provide personalized and convenient recommendations. Recommender systems, which have recently become prominent for their ability to analyze customer data and deliver customized suggestions, represent a viable strategy.

Recommender systems are software tools widely utilized across various industries, such as e-commerce, music, and movie streaming services. The increasing number of available products and services can complicate the decision-making processes of customers. In these sectors, personalized recommender systems, built using diverse machine learning

techniques, analyze vast data volumes, streamlining decision-making when customers are uncertain about their needs or pREFERENCES.

In aviation, recommender systems can significantly enhance customer satisfaction and loyalty. By examining data such as previous purchases, flight destinations, and feedback, these systems can provide recommendations tailored to individual customer pREFERENCES and needs. Analyzing customer behavior offers valuable insights for service improvement. For instance, an online review expressing dissatisfaction with a specific service can highlight areas for enhancement.

The extensive use of social media and online review platforms has generated a wealth of customer feedback and ratings regarding services in aviation. Analyzing this feedback yields crucial insights into customer behavior, expectations, experiences, and attitudes toward the airline company. This data can improve the precision and effectiveness of recommender systems. However, the vast volume and complexity of this data pose significant challenges for system design and implementation, requiring substantial processing power. Thus, developing efficient models that require less processing power is advantageous.

Numerous studies have shown the potential of recommender systems that incorporate online customer reviews in aviation. For instance, Aciar et al. proposed a recommendation model based on customer reviews that enhanced recommendation accuracy significantly [1]. Zhang et al. developed a personalized recommendation system that considers customer pREFERENCES and feedback [2]. Heidari and Rafatirad utilized a hybrid approach, combining customer reviews and flight itinerary data to deliver personalized recommendations [3].

This study aims to investigate the implementation of a classification model to find ways to recommend ancillary services in aviation. By analyzing online customer reviews which are written in English language, we seek to enhance the recommendation of ancillary services. Our study includes a thorough review of relevant literature, examining both the challenges and opportunities in this field. We propose a methodology that employs topic modeling and sentiment analysis, to analyze customer reviews and improve service recommendations for airline companies. Additionally, we integrate various machine learning algorithms, such as artificial neural networks (ANNs) and collaborative filtering methods, into our approach to comparing their performance in recommending ancillary services. Our analysis focuses on the following research questions:

- How can different classification algorithms improve recommendation accuracy using customer reviews?
- How effective are various classification algorithms when applied to recommender systems?
- What are the potential benefits and limitations of using classification models with online customer reviews in the airline industry?

This study offers valuable insights into implementing classification models through the analysis of online customer reviews in aviation. It also contributes to developing actionable

recommendations for airlines, demonstrating the potential use of recommender systems to increase ancillary product sales.

Literature Review

The literature review was conducted to align with the research questions outlined in the preceding section. The initial focus was on the general concept of recommender systems and their applications, before narrowing down to their specific use within the airline industry. Studies that analyzed online customer reviews to understand customer needs and opinions were particularly sought. Recent research post-2015 was given significant attention, although foundational definitions were considered from as early as 1988. The primary keywords used included recommender systems, recommender systems in the airline industry, ancillary products, online customer review analysis for the airline industry, and recommendations for ancillary products in the airline industry.

Recommender systems are algorithms that use data on past behavior, pREFERENCES, and feedback of customers to generate personalized suggestions. These systems have become increasingly popular as they help customers navigate the extensive and complex array of products and services available online. The complexity of choosing a product from numerous options can be simplified with recommender systems, which have been shown to enhance customer loyalty and increase sales [4]. The rise of online e-commerce and social platforms has resulted in an increase in user-generated review data, which can be used to build recommender systems that improve user experiences [5].

Various platforms, including e-commerce, social media, streaming services, and tourism, use recommender systems to enhance user experiences. The goal of these systems is to help customers find relevant items that match their needs or pREFERENCES, thus reducing the burden of searching through vast numbers of products [6].

Recommender systems typically follow three sequential steps: information gathering, learning, and recommendation. In the information gathering stage, data is collected to extract business insights. The learning stage involves identifying patterns and relationships within the data to make accurate recommendations. These relationships can be between users or products, recognizing similar users by examining their past purchases, ratings, or reviews. In the recommendation stage, the system presents a list of products or services based on these patterns. [7].

There are three primary types of recommendation methods: content-based, collaborative-based, and hybrid systems. Content-based recommender systems analyze past behavior and interests of a user to suggest similar items [8]. Collaborative filtering-based systems make recommendations based on similarities between users with similar pREFERENCES or interests [1]. Hybrid systems combine multiple recommendation approaches or algorithms to leverage the strengths of each method [9].

Most recommender systems using user reviews employ Natural Language Processing (NLP) techniques like sentiment analysis and opinion mining to extract insights [10]. Some systems focus on the positive or negative polarity of reviews, while others use multi-label classification for more complex models, resulting in better recommendations [11].

To provide high-quality service, companies must understand the expectations of their customers, which significantly influences service quality and satisfaction evaluations. Companies should listen to customers across various channels and integrate their feedback into their design process. In many industries, techniques such as design experimentation, quality function deployment, and value engineering are used. After providing a service, companies assess whether they have met customer expectations. Service quality is measured as the gap between expectations and perceptions, with smaller gaps indicating higher quality. The SERVQUAL method, developed by Parasuraman et al., addresses this gap in the airline industry by measuring ten dimensions of service quality [12]. Pakdil and Aydin extended this method with a 35-question survey [13].

Baker examined data from the top 14 U.S. airlines between 2007 and 2011 to compare service quality and customer satisfaction between low-cost and legacy carriers [14]. Tusar and Islam analyzed Twitter data from customers of six major U.S. airlines, using NLP techniques and various machine learning algorithms for sentiment analysis [15].

Wang et al. analyzed online airline passenger reviews to gain insight about recommendation tendency between passengers [16]. The study indicated that joy, trust, anger and disgust emotions affected the intention of the passenger to recommend. Logistic regression was used to find the most effective emotions.

Recommender systems in the airline industry assist customers in finding suitable flights based on pREFERENCES like destination, travel dates, and budget. Their performance can be improved by incorporating online customer reviews. Dadoun et al. discussed integrating recommender systems into airline business processes [17]. Khaturia et al. compared Naive Bayes and sentiment analysis approaches for recommender systems using customer tweets [18]. Lu et al. proposed a method using Gradient Boosting Decision Tree (GBDT) for suggesting ancillary services [19].

Mottini et al. noted that recommender systems are underutilized in the airline industry due to the perceived complexity of customer decision processes [20]. They compared three choice modeling techniques for recommender systems. Jain et al. used K Nearest Neighbors (KNN), SVM, and decision trees to predict customer feedback from online reviews [21]. Srinivas and Ramachandiran conducted unsupervised text analytics on passenger reviews to extract business intelligence and summarize reviews [22].

Heidari and Rafatirad proposed a method using Convolutional Neural Networks (CNN) with online reviews and flight data to recommend tickets [3]. Dundar et al. developed a system analyzing airport evaluations with linguistic summarization [23]. Zhang et al. proposed a contextualized recommender system to address the cold start problem, using factors like time, location, and weather for ancillary service recommendations [2].

Germán Pardo González et al. implemented a linear programming model to optimize revenue from seat arrangement which is one of ancillary products which we used in our study [24]. The main strategy is to arrange seating in a way that passengers will be inclined to pay for changing seats. This study deals with ancillary products; however, it does not take customer reviews into consideration.

Cai et al. used deep learning techniques such CNN and attention mechanisms to classify user complaints of airline companies. They combined semantic features of texts with label information and then classified them. Their results showed that the proposed model achieved better results than the models they compared in the study. Their study, however, does not take ancillary services into account [25].

Despite existing research on user review analysis and ancillary service recommender systems, a significant gap remains in integrating these areas. Studies either analyze ancillary services or passenger reviews, however they do not connect the two areas together. This study aims to bridge this gap by using online user reviews for sentiment analysis and topic extraction, feeding these insights into a machine-learning model to classify personalized recommendations for ancillary services in the airline industry.

Material and Methods

In this section, a comprehensive overview of the dataset is provided, detailing the columns and relevant statistical information. The effectiveness of various machine learning models is assessed by comparing Support Vector Machines (SVM), Extreme Gradient Boosting (XGBoost), and Artificial Neural Networks (ANN). The study employs Bidirectional Encoder Representations from Transformers (BERT) for sentiment analysis, BERTopic for topic extraction, and utilizes TensorFlow and Keras for implementing the artificial neural network frameworks. The design of the proposed model is illustrated, and the methods employed are thoroughly explained.

Data

Publicly available user reviews were gathered using web scraping methods from travel websites where customers share their experiences on airline review pages. These reviews include feedback on both positive and negative experiences with airline companies and service quality ratings. Approximately 300,000 unique reviews were collected, encompassing five columns: rating, review title, review text, route, and flight class.

- Rating: Users scored their overall flight experience from 1 (worst) to 5 (best).
- Review Title: A brief title assigned by users to their review.
- Review Text: Detailed descriptions of the experiences with the airline.
- Route: Combines the departure and destination cities in one column.
- Class: Indicates the fare class purchased by the customer, such as economy, business, etc.

The average rating across all airlines was 3.5 in the dataset. Table 1 presents statistical information on airline ratings, revealing that 81% of the reviews belong to economy class fares. To enhance the dataset, the route column was split into two new columns: “from city” and “to city.” Additionally, a new column was created to indicate the distance between the two cities, calculated using the “geopy” library, which uses city names to return the distance based on latitude and longitude values.

Table 1: Statistical Information of Ratings

Statistic	Rating
count	308418
mean	3.362
std	1.516
min	1.000
25%	2.000
50%	4.000
75%	5.000
max	5.000

The training data for the recommender system is a subset of review data with an extra label of an ancillary product column. Training data consists of 600 rows. A featured product column was introduced to pair reviews and products. Reviews were labeled by an expert in the airline industry via reading the review text and choosing the suitable ancillary product. Four of the available ancillary products used in this study are "Exit Seat", "Special Meal", "Priority Boarding", and "Extra Baggage". The data labeling process was done by choosing appropriate keywords for each ancillary product. For example, "legroom", "cramped", and "narrow" keywords were chosen for "Exit Seat" products to find related reviews. After finding reviews that contained keywords, they were read manually to check it was the correct label.

Methods

This section provides a comprehensive explanation of the methods, libraries, and metrics utilized in our study.

Support Vector Machines

Introduced by Cortes and Vapnik in 1995, Support Vector Machines (SVM) are a category of supervised machine learning algorithms, widely used for solving both regression and classification problems [26]. The core principle of SVM is to identify a potential boundary, or hyperplane, that maximizes the distance between different classes in the dataset. The goal of SVM is to determine not just any decision boundary, but the optimal one. The margin is defined as the distance between the hyperplane and the closest data points from each class, known as support vectors.

Extreme Gradient Boosting

XGBoost, introduced by Chen and Guestrin in 2016, is a supervised machine learning algorithm that employs multiple decision trees to make predictions [27]. It was developed based on the fundamental principles of gradient boosting. As an ensemble technique, its goal is to build a more robust learning model by combining several smaller models.

A decision tree is a tree-structured supervised machine learning algorithm consisting of a root node, branches for decision rules, internal nodes for features, and leaf nodes for outcomes.

Artificial Neural Networks

Artificial Neural Networks (ANNs), inspired by the learning capabilities of the human brain, consist of interconnected nodes or neurons organized into layers. The architecture of ANNs, as illustrated in Figure 3, includes an input layer, one or more hidden layers, and an output layer.

Bidirectional Encoder Representations from Transformers (BERT)

BERT, an NLP technique introduced by Google, is a deep learning algorithm designed for various tasks such as question-answering, sentiment analysis, and language translation. BERT's unique feature is its ability to pre-train a bidirectional transformer encoder, which can then be fine-tuned for specific NLP tasks.

BERTopic

Inspired by BERT, BERTopic is a topic modeling technique that leverages transformers and c-TF-IDF to generate clusters and produce easily interpretable topics while retaining key terms from the topic descriptions. It utilizes BERT to identify topics within text documents. BERTopic operates on the premise that topics can be represented as clusters of similar documents in high-dimensional space.

Keras

Keras is an open-source high-level deep learning API designed to operate seamlessly with the TensorFlow library. It was developed to prioritize user-friendliness and modularity, making it accessible for both beginners and experienced practitioners in the field of deep learning.

Evaluation Metrics

True positive results are when the model classifies a service as to recommend, and it is indeed a product which a customer would need. The true negatives are when the model classifies a service as not recommended and it is a product which a customer would not need. Whereas false positives and false negatives are when the model classifies a product as to recommend, however it is a product a customer would not need and vice versa. Furthermore, the following evaluation metrics are used to measure model performances:

- $Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)}$
- $Recall = \frac{TP}{TP + FN}$
- $Precision = \frac{TP}{TP + FP}$
- $F - 1\,Score = 2 \cdot \frac{(Recall \cdot \mathrm{Pr}\,ecision)}{(Recall + \mathrm{Pr}\,ecision)}$

Proposed Method

Following the problem identification, a structured approach was developed to utilize online reviews to classify ancillary products either as to recommend or not to recommend

to customers. The proposed method, outlined in Figure 1, involves reviewing processing and matching products for recommendation purposes. A pre-trained BERT model was employed and fine-tuned with the review data to improve sentiment analysis on the review texts. This model features 12 transformer blocks with 768 hidden units. Additionally, a model was constructed to extract topics from review texts using the BERTopic method. Subsequently, reviews, the topic extraction model, and product-review matchings were put into the classification module to classify ancillary products as recommendable or not.

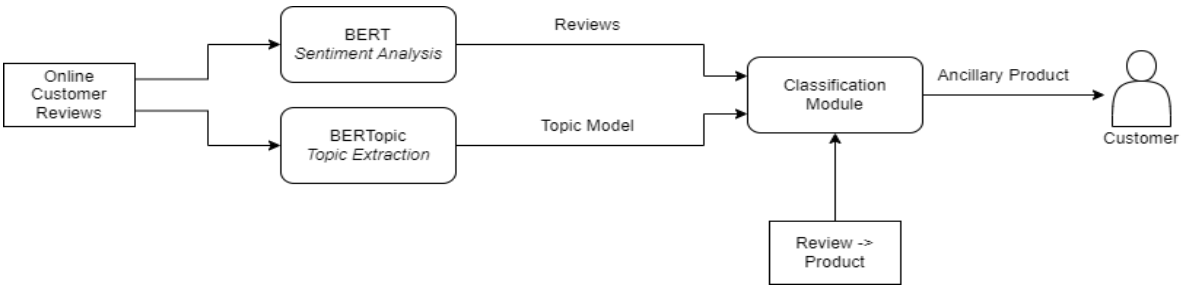


Figure 1: Proposed Method

The review preprocessing involved several steps. Initially, the review text was lowercased, and then all punctuation, links, and HTML code were removed using regular expressions. Stop words from the NLTK library for the English language were eliminated, and additional stop words were identified from the review text, such as city names and dates, which had no value for building the topic extraction model. Next, the reviews were lemmatized using the WordNetLemmatizer from the NLTK library. Lemmatization groups inflected forms of words, enabling analysis as a singular entity based on their common root or lemma. For example, both "happiness" and "happy" become "happi" after lemmatization. The lemmatized text was then tokenized using the same library, splitting sentences into smaller units to facilitate easier interpretation of their meaning.

The BERTopic method was utilized to extract topics from the given reviews. The text column from the entire review dataset was input into the algorithm to construct the topic extraction model. The UMAP algorithm was employed to reduce the dimensionality for improved clustering performance. The "n_neighbors" parameter, crucial for UMAP, determines the size of the local neighborhood, impacting the algorithm's focus on local versus global structures. Different values for "n_neighbors" were experimented with optimizing cluster size. Additionally, the "min_dist" parameter controlled the minimum distance between points allowed to pack together in lower dimensions, set to zero in this case for building clustering models. Figure 2 displays the top five words' scores for each topic extracted. Based on this initial iteration, additional words, such as 'one', 'mr', 'bbc', 'image', 'getty', 'de', 'en', 'caption', and 'also', were incorporated into the stop words dictionary as they were irrelevant in the context of ancillary services.

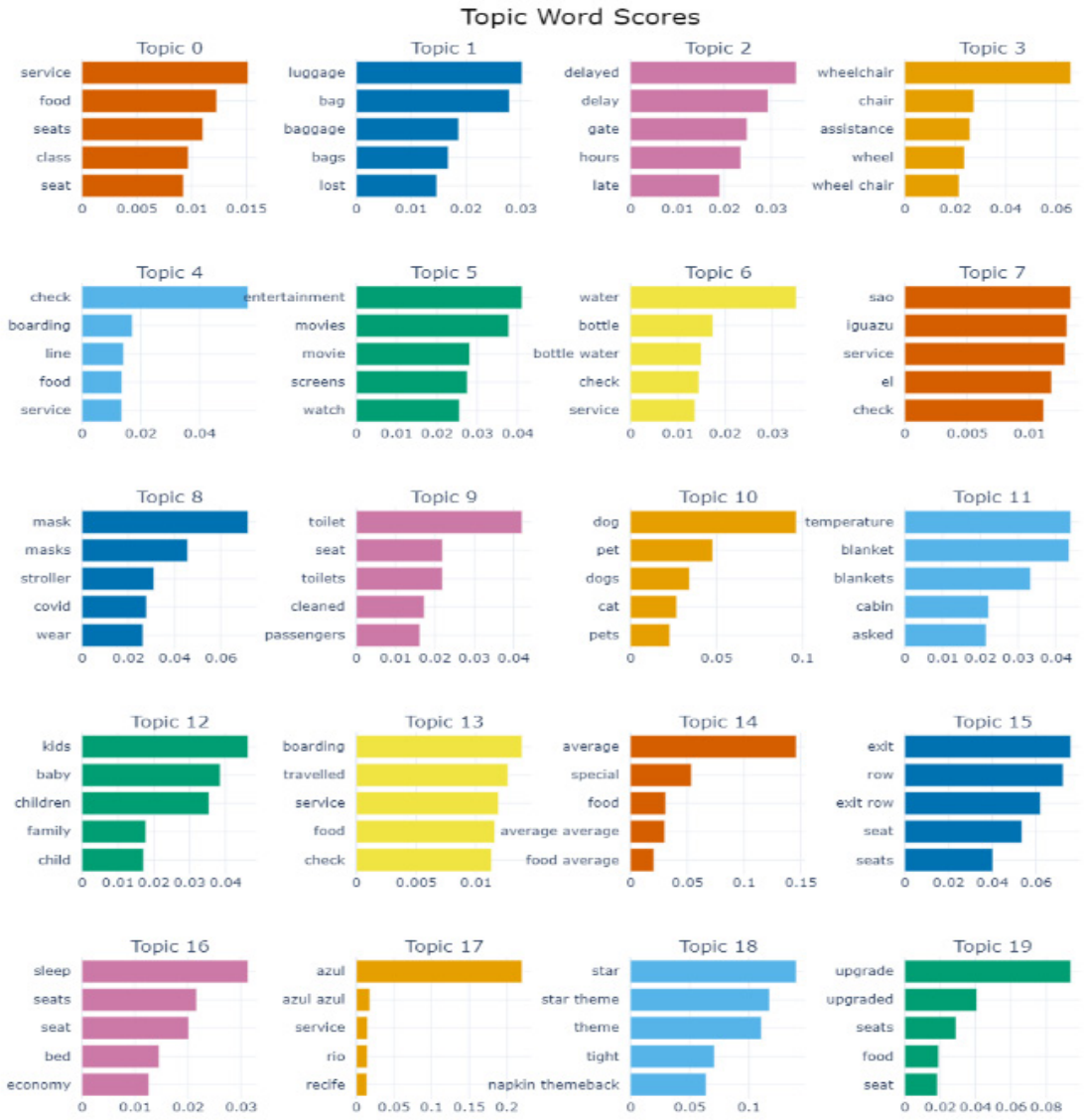


Figure 2: Topics

Once the sentiment analysis and topic extraction models were prepared, the dataset containing review and product matchings was passed into the topic extraction model. Subsequently, a vector was generated for each review, containing the assigned topics and distance information.

To improve results, the data vector was scaled using a standard scaler. The scaled data was then input into the support vector classifier of the sklearn library, employing different kernel methods such as linear, polynomial, and radial basis functions. Hyperparameter tuning played a crucial role in enhancing the overall performance of the machine learning algorithms. Hyperparameters, set before the learning process begins, were optimized using the GridSearchCV method. This method systematically tests different parameter values to find the optimal combination for model training.

The XGBoost algorithm was employed to evaluate the performance of tree-based methods. Initially, the algorithm was run with default parameter values, followed by hyperparameter tuning to achieve better prediction results. Parameters such as "max_depth", "learning_rate", "n_estimators", "colsample_bytree", and "gamma" were adjusted and compared based on changes in accuracy and RMSE scores.

As the third method artificial neural network was utilized to classify ancillary products. The TensorFlow library facilitated the construction of a network with hidden layers. Various hyperparameter configurations were tested to determine the optimal structure. Since the data vector comprised 20 members, the input layer consisted of 20 nodes. Different combinations were explored for the second and third layers. The "argmax" and "softmax" activation functions were tested for the output layer, with "softmax" yielding better results. As there were 4 classes in the target variable, the output layer consisted of 4 nodes. The hidden layer contained 128 nodes, with the output layer utilizing the "softmax" activation function. The sparse categorical cross-entropy function was selected as the loss function due to the target value being integers (0, 1, 2, 3). Additionally, the Adam optimizer function was chosen.

Findings

In this section, the results obtained from SVM, XGBoost, and ANN methods are stated through evaluation metrics. Table 2 presents the classification report of models used, depicting precision, recall, and F-1 score for each class. The accuracy rate achieved by the SVM method was 0.79 when utilizing the linear kernel and a C value of 0.5. The second method employed for recommending ancillary products was the XGBoost algorithm. The accuracy rate achieved using XGBoost was also 0.79. Finally, employing the ANN methods in our model yielded the highest accuracy rate of 0.85. Figure 3 illustrates the training and validation accuracies over time, while Figure 4 depicts the loss rate over time.

Table 2: Comparison of Model Performances

Model	Accuracy	RMSE	Recall	F-1 Score
SVM	0.79	0.23	0.79	0.79
XGBOOST	0.79	0.32	0.80	0.80
ANN	0.85	1.70	0.76	0.42

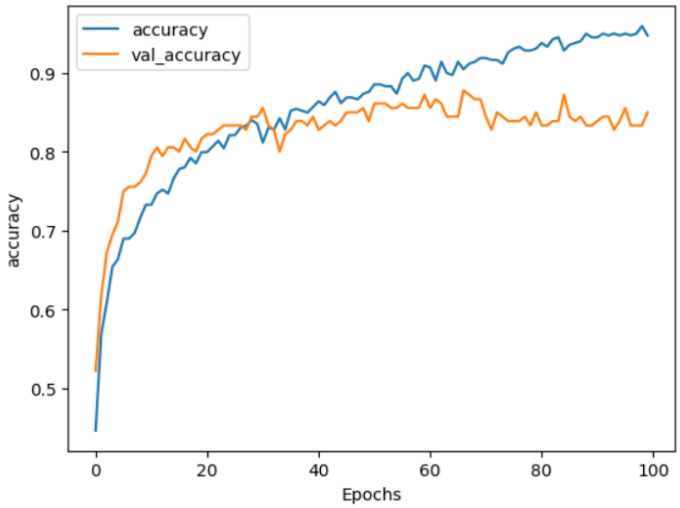


Figure 3: Accuracy Graph of NN

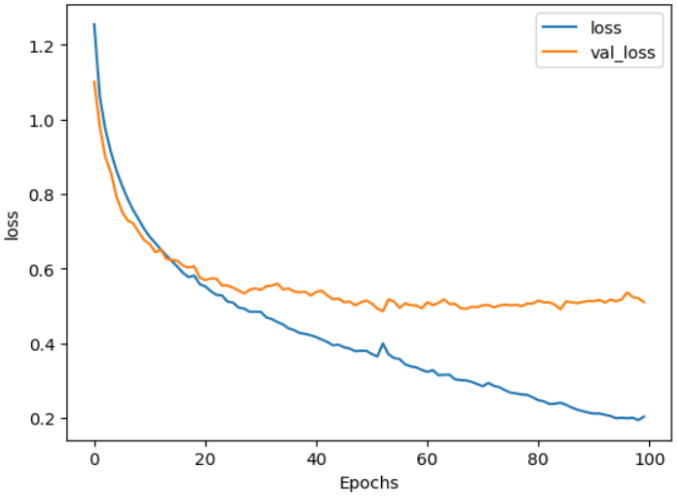


Figure 4: Loss Value Over Time

Discussion and Conclusion

The use of recommender systems based on online reviews in the airline industry holds great potential for enhancing customer satisfaction and driving business growth. By leveraging customer feedback and pREFERENCES, these systems can provide personalized recommendations for flights, destinations, and ancillary services, thereby improving the overall travel experience.

However, implementing such systems comes with its own set of challenges. One major obstacle is accurately analyzing and interpreting the vast amount of customer feedback available online. While natural language processing (NLP) techniques can assist in this process, their effectiveness may be limited, and ensuring accurate results can be challenging.

Another challenge lies in managing the large volumes of data generated by online reviews. This requires significant resources and expertise, as well as careful consideration of privacy and ethical concerns regarding the use of customer data.

Despite these challenges, many airlines are already successfully using recommender systems based on online reviews to improve customer satisfaction and drive business growth. Our study aimed to address a gap in existing literature by investigating the use of such systems to classify which ancillary products to recommend to the customers in the airline industry.

Table 2 summarizes the performance of three different models – SVM, XGBoost, and ANN – in our recommender system. While all three models demonstrated promising results, the ANN model outperformed the others with an accuracy rate of 0.85.

This highlights the potential effectiveness of deep learning algorithms in this context.

However, one of the main challenges we encountered during our study was the manual labeling of data by reading review content. Since there was no existing dataset containing customer choices for ancillary services, we had to rely on domain knowledge to label the review set. This manual process limited the amount of available data and may have affected the performance of deep learning algorithms.

Despite these challenges, our study contributes to the growing body of research on recommender systems in the airline industry and underscores the importance of leveraging online reviews to enhance customer satisfaction and drive business growth. Future research could focus on automating the data labeling process and further optimizing deep learning algorithms for this task.

Declaration of Interest: The authors declare that there is no conflict of interest.

Acknowledgements: This paper presents research conducted for the MSc thesis of Yavuz Selim Emir under the supervision of Tamer Uçar at Bahçeşehir University.

Author Contributions: Yavuz Selim Emir: Conceptualization, Methodology, Software, Investigation, Writing – Original Draft. Tamer Uçar: Conceptualization, Methodology, Supervision, Writing – Review & Editing

REFERENCES

[1] S. Aciar, D. Zhang, S. Simoff, and J. Debenham, "Informed recommender: Basing recommendations on consumer product reviews," IEEE Intell. Syst., 22(1), 39-47, 2007.

[2] Y. Zhang, W. Luo, M. Li, and T. Chen, "Contextualized recommendation of aviation ancillary services based on passenger portraits," IEEE Trans. Aerosp. Electron. Syst., 58(6), 5078-5088, 2022.

[3] M. Heidari and S. Rafatirad, "Using transfer learning approach to implement convolutional neural network model to recommend airline tickets by using online reviews," in Proc. 15th Int. Workshop Semantic and Social Media Adaptation and Personalization (SMA), 1-6, Zakynthos, Greece, 2020.

[4] T.L. Keiningham, B. Cooil, L. Aksoy, T.W. Andreassen, and J. Weiner, "The value of different customer satisfaction and loyalty metrics in predicting customer retention, recommendation, and share-of-wallet," Managing Service Quality: An International Journal, 17(4), 361-384, 2007.

[5] Y. Lu, R. Dong, and B. Smyth, "Coevolutionary recommendation model: Mutual learning between ratings and reviews," In Proceedings of the 2018 World Wide Web Conference (WWW), 773-782, 2018.

[6] G. Adomavicius and A. Tuzhilin, " Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," IEEE Transactions on Knowledge and Data Engineering, 17(6), 734-749, 2005.

[7] F. Isinkaye, Y. Folajimi, and B. Ojokoh, "Recommendation systems: Principles, methods, and evaluation," Egyptian Informatics Journal, 16(1), 233-250, 2015.

[8] M.J. Pazzani and D. Billsus, "Content-based recommendation systems," in P. Brusilovsky, A. Kobsa, and W. Nejdl, Eds., The Adaptive Web, 325-341, Springer, 2007.

[9] V. Danilova and A. Ponomarev, "Hybrid recommender systems: The review of state-of-the-art research and applications," In Proceedings of the 20th Conference of FRUCT Association, 2016.

[10] M. Chelliah and S. Sarkar, "Product recommendations enhanced with reviews," Product recommendations enhanced with reviews. In Proceedings of RecSys 2017 2017, 398-399, ACM, 2017.

[11] Q. Diao, M. Qiu, C. Wu, A.J. Smola, J. Jiang, and C.J. Wang, "Jointly modeling aspects, ratings and sentiments for movie recommendation (JMARS)," In Proceedings of SIGKDD 2014, 193-202, ACM, 2014.

[12] A. Parasuraman, V.A. Zeithaml, and L.L. Berry, "SERVQUAL: A multiple item scale for measuring consumer perceptions of service quality," Journal of Retailing, 64(1), 12-40, 1988.

[13] F. Pakdil and O. Aydin, "Expectations and perceptions in airline services: An analysis using weighted SERVQUAL scores," Journal of Air Transport Management, 13(4), 229-237, 2007.

[14] D.M.A. Baker, "Service quality and customer satisfaction in the airline industry: A comparison between legacy airlines and low-cost airlines," American Journal of Tourism Research, 2(1), 1-11, 2013.

[15] T.H.K. Tusar and T. Islam, "A comparative study of sentiment analysis using NLP, and different machine learning techniques on US airline Twitter data," arXiv preprint arXiv:2110.00859, 2021.

[16] X. Wang, J. Zheng, L. Tang, and Y. Luo, "Recommend or not? The influence of emotions on passengers' intention of airline recommendation during COVID-19," Tourism Management, 95, 104675, 2023.

[17] A. Dadoun, M. Defoin-Platel, T. Fiig, et al., "How recommender systems can transform airline offer construction and retailing," Journal of Revenue and Pricing Management, 20(4), 301-315, 2021.

[18] D. Khaturia, A. Saxena, M. Basha, N.C. Iyenger, and R. Caytiles, "A comparative study on airline recommendation system using sentimental analysis on customer tweets," International Journal of Advanced Science and Technology, 111, 107-114, 2018.

[19] H. Lu, J. Cao, Y. Tan, Q. Xiao, et al., "Auxiliary service recommendation for online flight booking," in Web Information Systems Engineering 40-49, Springer, 2017.

[20] A. Mottini, A. Lhéritier, R. Acuna-Agost, and M.A. Zuluaga, "Understanding customer choices to improve recommendations in the air travel industry," in RecTour 2018, 2018.

[21] P.K. Jain, R. Pamula, S. Ansari, D. Sharma, and L. Maddala, "Airline recommendation prediction using customer generated feedback data," in Proceedings of ISCON, 376-379, Mathura, India, 2019.

[22] S. Srinivas and S. Ramachandiran, "Discovering airline-specific business intelligence from online passenger reviews: An unsupervised text analytics approach," arXiv preprint arXiv:2012.08000, 2020.

[23] B. Dundar, D. Akay, and S. Ozdemir, "Linguistic descriptions of airport evaluation data for recommendation strategies," in Proceedings of ISNCC 2021, 1-8, 2021.

[24] G. Pardo González, A. Tabares Pozos, C. Quiroga, and D. Álvarez-Martínez, "Seat assignment recommendation in airlines purchase flow to increase ancillary revenue considering weight and balance constraints," Journal of Air Transport Management, 117, 102582, 2024.

[25] H. Cai, X. Shao, P. Zhou, and H. Li, "Multi-label classification of complaint texts: Civil aviation service quality case study," Electronics, vol. 14, no. 3, p. 434, 2025.

[26] C. Cortes and V.N. Vapnik, "Support-vector networks," Machine Learn., 20(3), 273-297, 1995.

[27] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in Proc. 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16), 785-794, ACM, 2016.

DİNAMİK SINIFLANDIRICI SEÇİM YÖNTEMLERİ İLE METİN TABANLI DUYGU SINIFLANDIRMASI: MAĞAZA YORUMLARI ÜZERİNE BİR ÇALIŞMA

İsmail Terzi*

Zonguldak Bülent Ecevit Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Zonguldak, Türkiye

**(İletişimden sorumlu yazar) ismailterzi@beun.edu.tr*

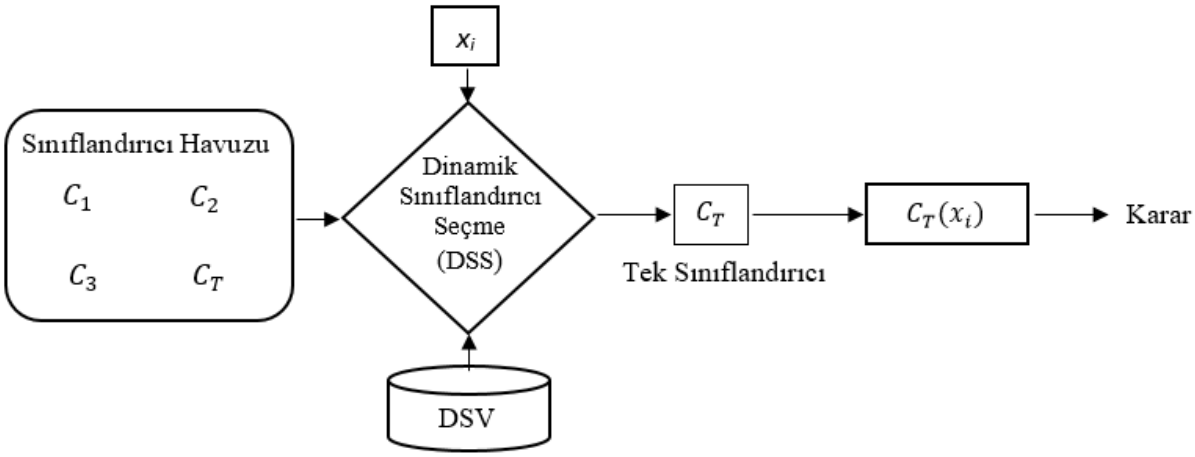
ÖZET

Bu çalışmada, e-ticaret ve mağaza değerlendirme platformlarında kullanıcılar tarafından paylaşılan yorumlar üzerinden duygu analizi gerçekleştirilmiştir. Yorumların içerdiği olumlu ya da olumsuz duygu durumlarının doğru bir şekilde sınıflandırılması, hem müşteri memnuniyetinin analiz edilmesi hem de işletmelerin hizmet kalitesini iyileştirmeye yönelik stratejiler geliştirebilmesi açısından büyük önem taşımaktadır. Bu kapsamda, çalışmada makine öğrenmesi algoritmaları kullanılarak, kullanıcı yorumları üzerinde metin sınıflandırma işlemi uygulanmıştır. Kullanılan veri seti, toplam 8485 kullanıcı yorumundan oluşmakta olup, veriler dengeli bir şekilde dağıtılmıştır; yorumların %50'si olumlu, %50'si ise olumsuz duygu içermektedir. Sınıflandırma sürecinde, yorumların içerdiği duygu durumunu en yüksek doğrulukla tespit edebilmek amacıyla, farklı sınıflandırıcı algoritmaların bir arada değerlendirildiği Dinamik Sınıflandırıcı Seçme (DSS) yöntemleri kullanılmıştır. Bu yöntemler sayesinde, her bir örnek için en uygun sınıflandırıcının otomatik olarak belirlenmesi sağlanmıştır. Çalışmada kullanılan dinamik sınıflandırıcı seçme yöntemleri arasında A posteriori, A priori, Local Class Accuracy (LCA), Multiple Classifier Behaviour (MCB), Modified Local Accuracy (MLA), Overall Local Accuracy (OLA) ve Modified Rank stratejileri yer almaktadır. Bu yöntemler, sınıflandırıcıların farklı bağlamlardaki performanslarını dikkate alarak, sınıflama sürecini örnek düzeyinde optimize etmeyi amaçlamaktadır. Yöntemlerin performansı, doğruluk, hassasiyet, özgüllük ve F1 skoru gibi metrikler üzerinden karşılaştırmalı olarak değerlendirilmiştir. Elde edilen sonuçlar, dinamik sınıflandırıcı seçme yöntemlerinin, özellikle metin temelli duygu analizinde tekil sınıflandırıcılara göre daha başarılı sonuçlar verdiğini ortaya koymuştur.

Anahtar Kelimeler: Duygu analizi, metin sınıflandırma, dinamik sınıflandırıcı seçme, yorum makine öğrenmesi

Giriş

Son yıllarda dijitalleşmenin etkisiyle kullanıcılar, ürün ve hizmetlere ilişkin görüşlerini sıklıkla çevrimiçi platformlarda paylaşmaktadır. Özellikle e-ticaret siteleri ve mağaza değerlendirme platformları, kullanıcı yorumları açısından zengin ve anlamlı veri kaynakları haline gelmiştir. Bu yorumlar, işletmelerin müşteri memnuniyetini analiz etmesi, hizmet kalitesini değerlendirmesi ve stratejik kararlar alması açısından büyük önem taşımaktadır [1,2].



Şekil 1: Dinamik Sınıflandırıcı Seçme(DSS)

Şekil 1, dinamik sınıflandırıcı seçme sürecinin temel işleyişini göstermektedir. Sınıflandırıcı havuzunda yer alan farklı modeller arasından, her bir test örneği için en uygun sınıflandırıcının seçilmesi hedeflenmektedir. Sisteme giren x_i örneği, dinamik seçim mekanizmasına yönlendirilir ve burada, Dinamik Seçim Verisi (DSV) kullanılarak sınıflandırıcıların lokal ve genel performanslarına dayalı bir analiz gerçekleştirilir. Seçilen en yetkin sınıflandırıcı , ilgili örnek üzerinde sınıflandırma işlemini gerçekleştirir ve nihai karar üretilir.

Doğal dil işleme (NLP) ve makine öğrenmesi (ML) teknikleri, metinlerden anlam çıkararak sınıflandırma yapılmasını mümkün kılmakta ve bu sayede kullanıcı yorumları üzerinden duygu analizi gerçekleştirilmesine olanak tanımaktadır [3]. Metin sınıflandırma, geleneksel olarak tekil sınıflandırıcılar (örneğin Naive Bayes, SVM, Decision Trees vb.) ile gerçekleştirilse de, bu algoritmaların farklı veri örnekleri üzerinde değişken performans sergilediği bilinmektedir [4,5]. Bu nedenle, son dönemde birden çok sınıflandırıcının birlikte kullanıldığı topluluk (ensemble) yöntemleri ve özellikle DSS yaklaşımları ön plana çıkmaktadır. DSS, her bir test örneği için en uygun sınıflandırıcıyı seçmeyi amaçlamakta ve bu sayede sınıflandırma başarımını artırmaktadır [6]. Bu yaklaşım, özellikle heterojen sınıflandırıcı havuzlarının bulunduğu durumlarda daha etkili sonuçlar vermektedir [7]. DSS yöntemleri sınıflandırılacak örnek için sınıflandırıcı havuzundan tek bir sınıflandırıcı seçerek sınıflandırma yapmaktadır.

Bu çalışmada, e-ticaret platformlarında yer alan kullanıcı yorumları üzerinden duygu analizi gerçekleştirilmiş ve farklı DSS yöntemlerinin etkinliği karşılaştırılmıştır. Kullanılan yöntemler arasında A posteriori, A priori, Local Class Accuracy (LCA), Multiple Classifier Behaviour (MCB), Modified Local Accuracy (MLA), Overall Local Accuracy (OLA) ve Modified Rank stratejileri yer almaktadır. Bu yöntemler, sınıflandırıcıların yerel performanslarına göre örnek bazlı seçim yaparak, genel doğruluğu artırmayı hedeflemektedir [8].

Çalışmanın temel amacı, dengeli bir veri kümesi üzerinde çeşitli DSS yöntemlerinin performansını doğruluk, hassasiyet, özgüllük ve F1 skoru gibi metrikler aracılığıyla analiz etmek ve tekil sınıflandırıcılarla elde edilen sonuçlarla karşılaştırarak bu yöntemlerin duygu analizi görevlerindeki etkinliğini ortaya koymaktır.

İlgili Çalışmalar

Duygu sınıflandırma, metin madenciliği ve doğal dil işleme (NLP) alanlarında öne çıkan bir araştırma konusudur. Kullanıcıların metinler aracılığıyla ifade ettikleri duyguların otomatik olarak sınıflandırılması, özellikle e-ticaret, sosyal medya, kamuoyu araştırmaları ve müşteri hizmetleri gibi alanlarda yaygın biçimde uygulanmaktadır. Bu alanda gerçekleştirilen çalışmalar, hem geleneksel makine öğrenmesi algoritmalarını hem de derin öğrenme temelli yaklaşımları kapsamaktadır. Duygu analizi üzerine yapılan ilk sistematik çalışmalardan biri Pang, Lee ve Vaithyanathan tarafından gerçekleştirilmiş ve film yorumları üzerinde Naive Bayes, Maximum Entropy ve Support Vector Machines (SVM) gibi yöntemler kullanılarak duygu sınıflandırması yapılmıştır [9]. Bu çalışma, sonraki birçok çalışmaya temel oluşturmuştur. Liu ise kapsamlı bir literatür taramasıyla hem sözlük tabanlı hem de makine öğrenmesi temelli yöntemleri karşılaştırmalı biçimde incelemiş ve alanın gelişim seyrine ışık tutmuştur [2]. Zamanla derin öğrenme temelli yöntemler de duygu sınıflandırma problemlerinde etkili biçimde kullanılmaya başlanmıştır. Kim, tek bir konvolüsyonel sinir ağı (CNN) modeliyle cümle düzeyinde duygu sınıflandırması gerçekleştirmiştir [10]. Tang ve arkadaşları ise, LSTM tabanlı modeller ile kullanıcı ve ürün bilgilerini entegre ederek duygu analizi doğruluğunu artırmayı başarmıştır [11]. Duygu analizi alanında bir diğer önemli yönelim ise çoklu sınıflandırıcı sistemlerin kullanılmasıdır. Go, Bhayani ve Huang, Twitter verisi üzerinde farklı sınıflandırıcıları değerlendirerek basit özellik mühendisliği ile etkili sonuçlar elde edilebileceğini göstermiştir [12]. Alyari ve arkadaşları ise film yorumları üzerinde gerçekleştirdikleri çalışmada, ensemble (topluluk) yöntemlerinin duygu analizi performansını artırabildiğini göstermiştir [13]. Ancak geleneksel tekil sınıflandırıcılar veya sabit ensemble yapıları, her örneğe aynı sınıflama yaklaşımını uyguladıkları için bağlama duyarsız kalabilmektedir. Bu soruna çözüm olarak, dinamik sınıflandırıcı seçme (Dynamic Classifier Selection - DSS) yöntemlerinin kullanımı önerilmiştir. Cruz ve arkadaşları, DSS yöntemlerinin sınıflandırma doğruluğunu örnek bazında artırabileceğini ve özellikle metin sınıflandırma gibi karmaşık problemlerde daha esnek yapılar sunduğunu ortaya koymuştur [6]. Ancak duygu sınıflandırma bağlamında DSS yöntemlerinin kullanımı hâlen sınırlı sayıda çalışmada ele alınmıştır. Bu nedenle, mevcut çalışmanın temel katkılarından biri, farklı DSS stratejilerinin duygu analizi görevlerinde karşılaştırmalı olarak değerlendirilmesidir.

Yöntem

Bu bölümde, çalışmada kullanılan veri seti, önişleme adımları, sınıflandırıcı havuzu ve uygulanan DSS yöntemleri detaylandırılmaktadır.

Veri Seti

Çalışmada, 8485 kullanıcı yorumu içeren mağaza yorumları veri seti kullanılmıştır[14]. Kullanılan mağaza yorumları veriseti iki sınıflı dengeli bir verisetidir. Yorumların %50'si olumlu, %50'si ise olumsuz duygu içermektedir. Dengeli dağılım, sınıflandırma algoritmalarının performansının adil şekilde değerlendirilebilmesine olanak sağlamıştır.

Veri Önişleme

Ham kullanıcı yorumları üzerinde aşağıdaki önişleme adımları uygulanmıştır:

- Tüm metinler küçük harfe dönüştürülmüştür.
- Noktalama işaretleri, özel karakterler ve sayılar kaldırılmıştır.
- Türkçe durak (stop) kelimeleri metinden çıkarılmıştır.
- Kelimeler köklerine indirgenmiş (lemmatizasyon/stemming).
- TF-IDF (Term Frequency-Inverse Document Frequency) yöntemi kullanılarak sayısal vektör temsilleri oluşturulmuştur.

Sınıflandırıcılar metinler üzerinde işlem yapabilmeleri için metinler önce kelimelere, kelimeler de sayısal değerlere dönüştürülmelidir. Bu alanda kullanılan farklı yöntemler vardır, bu çalışmada Sözcük Çantası (The Bag of Words)[21] modelini kullanılmıştır. Bu modelde metin kümesi matrise dönüştürülür. Bu matrisin satırları metinleri, sütunları ise kelimelerin oluşturduğu öznitelikleri temsil eder. Matristeki hücrelerdeki sayısal değerler Terim Frekansı Ters Doküman Frekansı(TFTDF) (Term Frequency Inverse Document Frequency) (TFIDF) değerleridir.

Metin sınıflandırma problemlerinde öznitelik sayıları oldukça yüksektir. Çok fazla öznitelik kullanarak sınıflandırma yapmak sınıflandırma doğruluğunu düşürebilir, bu sebeple sınıflandırmaya katkısı en yüksek olan öznitelikler seçilmelidir. Bu nedenle daha doğru bir sınıflandırma yapabilmek için öznitelik seçim yöntemlerinden, Chi-square kullanılmıştır [21]. Öznitelik seçim yöntemi uygulandıktan sonra elde edilen matris, rastgele olacak şekilde %50 oranında eğitim, %25 oranında test ve %25 oranında dinamik sınıflandırma verisine ayrıldı. Sınıflandırıcı havuzundaki tüm sınıflandırıcılar aynı eğitim verisi kullanılarak eğitildi. Deneyler 50, 100,300,500,1000 öznitelik sayıları kullanılarak yapıldı. Sınıflandırıcı havuzundaki sınıflandırıcıların ortalama doğruluğu hesaplandı.

Sınıflandırıcı havuzu

Dinamik seçim yöntemlerinde kullanılmak üzere bir sınıflandırıcı havuzu oluşturulmuştur. Sınıflandırıcı havuzu aşağıdaki beş temel makine öğrenmesi algoritmasından oluşmaktadır, bu algoritmalar literatürde yaygın olarak kullanılan yüksek doğrulukta sınıflandırma yapan algoritmalardır.

- **K-En Yakın Komşu (K-NN)**
- **Destek Vektör Makineleri (SVM)**
- **Karar Ağaçları (Decision Tree - DT)**
- **Naive Bayes (NB)**
- **Lojistik Regresyon (LR)**

Bu sınıflandırıcılar, Python programlama diliyle birlikte **scikit-learn** kütüphanesi kullanılarak eğitilmiş ve test edilmiştir.

Dinamik sınıflandırıcı seçme yöntemleri

Her bir test örneği için en uygun sınıflandırıcının seçilebilmesi amacıyla **DESLib** (Dynamic Ensemble Selection Library) [15] kütüphanesi aracılığıyla çeşitli dinamik seçim stratejileri uygulanmıştır. Kullanılan DSS yöntemleri aşağıda listelenmiştir:

- **A Posteriori [16]:** Aday sınıflandırıcıların doğruluk tahminlerini kullanarak seçim yapar.
- **A Priori [16]:** Her sınıflandırıcının genel doğruluk düzeyine göre seçim yapar.
- **Local Class Accuracy (LCA) [17]**: Belirli bir sınıfa özgü doğruluk değerlerine dayanır.
- **Multiple Classifier Behaviour (MCB) [18]**: Sınıflandırıcıların benzer davranış örüntülerine göre seçim yapar.
- **Modified Local Accuracy (MLA) [19]**: LCA yönteminin iyileştirilmiş bir versiyonudur.
- **Overall Local Accuracy (OLA) [17]**: Sınıflandırıcıların lokal doğruluk ortalamasını temel alır.
- **Classifier Rank (DCS-Rank) [20]** : Sınıflandırıcıların sıralamasını lokal doğruluklara göre günceller.

Her bir yöntem, sınıflandırıcı havuzu içerisinden test örneğine en uygun olanı seçerek sınıflama kararını bu seçilen sınıflandırıcıya dayandırır. Bu sayede, sınıflandırma süreci örnek düzeyinde özelleştirilmiş olur.

Deney Sonuçları

Bu bölümde, deney sonuçları sunulmaktadır. Amaç, farklı DSS stratejilerinin duygu sınıflandırma görevlerindeki başarımını karşılaştırmalı olarak analiz etmektir. Deneyler, 8485 kullanıcı yorumundan oluşan dengeli bir veri seti üzerinde gerçekleştirilmiş olup, yorumların %50'si olumlu, %50'si ise olumsuz sınıf etiketine sahiptir. Sınıflandırma sürecinde oluşturulan sınıflandırıcı havuzu, k-NN, SVM, Karar Ağaçları, Naive Bayes ve Lojistik Regresyon modellerinden oluşmaktadır.

Model performanslarını değerlendirmek amacıyla doğruluk (%), hassasiyet (precision), özgüllük (specificity) ve F1 skoru gibi yaygın sınıflandırma metrikleri kullanılmıştır. Sonuçlar, 10 katlı çapraz doğrulama yöntemiyle elde edilerek ortalama değerler halinde raporlanmıştır. Tablo 1'de her bir DSS yönteminin ve sınıflandırıcı havuzundaki tekil sınıflandırıcıların ortalama doğrulukları gösterilmiştir.

Deneyler 100,300,500,1000,3000,5000 öznitelik sayıları kullanılarak yapılmıştır. Deney sonuçları sınıflandırma doğruluğu Tablo 1'de, hassasiyet değerleri Tablo 2'de, özgüllük değerleri Tablo 3'te ve F1 değerleri Tablo 4'te verilmiştir. He öznitelik sayısında en yüksek skoru elde eden yöntem koyu ve altı çizili olarak gösterilmiştir.

Tablo 1: Farklı öznitelik sayılarında sınıflandırma doğruluğu(%)

	50	**100**	**200**	**300**	**500**	**1000**
Havuz Ort.	79,23	81,77	84,33	85,37	85,67	84,40
Rank	79,80	81,63	84,10	85,83	85,93	81,30
OLA	77,97	82,20	84,70	85,93	85,90	83,03
LCA	77,83	82,20	84,57	85,30	83,73	75,77
Apriori	79,23	80,90	82,93	84,23	84,77	83,00
APosteriori	**81,60**	**83,33**	84,33	84,73	84,97	82,57
MCB	77,57	82,23	**85,00**	**86,40**	**86,43**	**85,53**
MLA	77,83	82,20	84,57	85,30	83,73	75,77

Tablodaki sonuçlar, farklı öznitelik sayılarının (50'den 1000'e kadar) sınıflandırma doğruluğu üzerindeki etkisini ve çeşitli DSS yöntemlerinin başarımını göstermektedir. İlk satırda yer alan "Havuz Ort." değeri, sınıflandırıcı havuzundaki tüm modellerin ortalama doğruluğunu temsil ederken, diğer satırlar ilgili DSS yöntemlerinin doğruluk oranlarını yansıtmaktadır. Öznitelik sayısı 50'den 500'e doğru arttıkça genel olarak tüm yöntemlerde belirgin bir doğruluk artışı gözlenmektedir. Örneğin, havuz ortalaması %79.23'den %85.67'ye yükselmektedir. Ancak 1000 öznitelik seviyesine gelindiğinde bazı yöntemlerde düşüşler görülmekte, bu da aşırı öznitelik artışının performansı olumsuz etkileyebileceğini göstermektedir (örneğin Rank yöntemi %85.93'ten %81.30'a gerilemektedir).

Genel olarak, dinamik sınıflandırıcı seçme yöntemlerinin büyük çoğunluğu, sınıflandırıcı havuzunun ortalama doğruluğuna kıyasla belirli öznitelik seviyelerinde daha yüksek başarı sağlamaktadır. Özellikle 300-500 öznitelik aralığı, hemen tüm yöntemlerde doğruluk açısından en iyi sonuçların elde edildiği aralık olarak öne çıkmaktadır. Sonuç olarak Tablo 1, dinamik sınıflandırıcı seçme yöntemlerinin duygu sınıflandırma görevlerinde tekil sınıflandırıcılara göre belirgin bir performans avantajı sunduğunu ve özellikle veri setinin öznitelik boyutuna göre farklı stratejilerin öne çıktığını göstermektedir.

Tablo 2: Farklı öznitelik sayılarında hassasiyet(%)

	50	**100**	**200**	**300**	**500**	**1000**
Havuz Ort.	80,90	82,80	85,00	85,80	85,50	85,57
Rank	82,57	83,77	85,34	86,07	85,00	84,06
OLA	**83,20**	83,94	85,61	85,94	84,02	85,08
LCA	81,24	**84,64**	84,78	85,77	84,26	82,15
Apriori	81,88	82,01	83,42	84,27	84,67	84,87
APosteriori	83,07	84,38	85,42	85,13	**85,58**	85,65
MCB	83,08	83,67	**85,82**	**86,72**	85,58	**86,33**
MLA	81,23	**84,64**	84,76	85,77	84,26	82,15

Tablo 2'de farklı öznitelik sayılarının (50'den 1000'e kadar) sınıflandırma hassasiyet değerleri üzerindeki etkisini ve çeşitli DSS yöntemlerinin performansını göstermektedir. İlk satırdaki "Havuz Ort.", sınıflandırıcı havuzundaki tüm modellerin ortalama hassasiyet değerlerini, diğer satırlar ise ilgili DSS yöntemlerinin hassasiyet performanslarını ortaya koymaktadır. Hassasiyet değerleri genel olarak 100 öznitelikten sonra yükseldiği ve 300-500 öznitelik aralığında zirve yaptığı görülmektedir. En yüksek hassasiyet 300 öznitelikte MCB yöntemiyle (%86.72) elde edilmiş. 1000 öznitelikte, genel eğilim istikrarını korumaktadır; bu durum aşırı

öznitelik artışının hassasiyet üzerinde dramatik olumsuz etkiler yaratmadığını göstermektedir. DSS yöntemlerinin büyük kısmı, havuz ortalamasına kıyasla hassasiyet değerlerinde kayda değer bir iyileşme sağlamış. Hassasiyet sonuçları, dinamik sınıflandırıcı seçme yöntemlerinin duygu sınıflandırmada yalnızca doğruluk değil, aynı zamanda kesinlik açısından da önemli avantajlar sunduğunu ortaya koymaktadır. MCB ve APosteriori yöntemleri, yüksek öznitelik sayılarında istikrarlı biçimde öne çıkarken, OLA yöntemi de orta öznitelik seviyelerinde güçlü bir alternatif olarak dikkat çekmektedir.

Tablo 3: Farklı öznitelik sayılarında *özgüllük* (%)

	50	**100**	**200**	**300**	**500**	**1000**
Havuz Ort.	79,33	81,93	84,50	85,40	85,23	85,20
Rank	82,09	83,50	85,14	85,91	84,91	83,89
OLA	**82,67**	83,55	85,34	85,70	83,78	85,01
LCA	80,36	83,70	84,18	84,64	83,55	80,61
Apriori	81,40	81,82	83,33	84,19	84,60	84,82
APosteriori	82,66	**83,88**	84,73	84,39	85,22	85,22
MCB	82,66	83,41	**85,68**	**86,63**	**85,40**	**86,26**
MLA	80,34	83,70	84,17	84,64	83,55	80,61

Tablo 3'te, farklı öznitelik sayılarının (50–1000 arası) sınıflandırma **özgüllük** değerleri üzerindeki etkisini ve çeşitli DSS yöntemlerinin performansını göstermektedir. İlk satırdaki "Havuz Ort.", sınıflandırıcı havuzundaki tüm modellerin ortalama özgüllük değerlerini sunarken, diğer satırlar DSS yöntemlerinin performanslarını göstermektedir. Genel olarak, özgüllük değerlerinin 50 öznitelikten 300 özniteliğe kadar belirgin bir artış gösterdiği görülüyor. Örneğin Havuz Ort. %79.33'ten %85.40'a yükselmiş. 500 ve 1000 öznitelikte sonuçlar genel olarak istikrarlı kalmış veya hafif dalgalanmalar göstermiş. En yüksek özgüllük MCB yöntemiyle 300 öznitelikte (%86.63) elde edilmiş. 1000 öznitelikte genel olarak hafif düşüşler göze çarpsa da, MCB ve APosteriori yöntemleri yüksek seviyelerini koruyabilmiştir. DSS yöntemleri genel olarak Havuz Ortalaması'na göre daha yüksek özgüllük sağlamış; bu da dinamik seçim stratejilerinin duyarlılığı artırmada etkili olduğunu göstermektedir. Özgüllük sonuçları, dinamik sınıflandırıcı seçme yöntemlerinin, özellikle veri örneklerinin doğru şekilde pozitif sınıfa atanmasında önemli başarı sağladığını ortaya koymaktadır. MCB yöntemi yine en güçlü aday olarak öne çıkarken, APosteriori ve OLA yöntemleri de rekabetçi ve istikrarlı sonuçlar vermektedir.

Tablo 4: Farklı öznitelik sayılarında F1 (%)

	50	**100**	**200**	**300**	**500**	**1000**
Havuz Ort.	78,77	81,80	84,70	85,53	85,80	85,77
Rank	79,20	82,59	84,82	86,39	84,68	84,67
OLA	79,70	81,60	85,92	86,14	84,28	86,04
LCA	77,67	83,09	84,70	85,17	83,84	83,20
Apriori	80,58	80,89	83,69	84,38	84,66	84,64
APosteriori	82,76	82,70	84,57	84,97	84,17	84,73
MCB	79,23	81,70	86,29	86,64	85,91	87,16
MLA	77,67	83,06	84,70	85,17	83,84	83,20

Tablo 4'te, (50–1000 arası) farklı öznitelik sayılarında sınıflandırma F1 skoru üzerindeki etkisini ve DSS yöntemlerinin başarımını göstermektedir. "Havuz Ort.", sınıflandırıcı havuzu-

nun ortalama F1 değerlerini sunarken diğer satırlar, ilgili DSS yöntemlerinin performansını yansıtmaktadır. 50'den 300 özniteliğe kadar genel olarak tüm yöntemlerde F1 skorlarında belirgin bir artış gözlenmektedir. Örneğin, Havuz Ort. %78.77'den %85.53'e çıkmış. 500 ve 1000 öznitelikte ise genel eğilim istikrar veya hafif dalgalanma şeklinde olmuştur. En yüksek F1 skoru MCB yöntemiyle 1000 öznitelikte (%87.16) elde edilmiş. Genel olarak DSS yöntemleri, Havuz Ortalaması'na kıyasla hemen her öznitelik sayısında F1 skoru açısından belirgin avantaj sağlamıştır. 300 öznitelik civarında hemen tüm yöntemlerin güçlü sonuçlar verdiği, bu aralığın optimum öznitelik boyutu olabileceğini düşündürmektedir. F1 skoru analizine göre dinamik sınıflandırıcı seçme yöntemleri, yalnızca doğruluk ve özgüllük değil aynı zamanda F1 skoru açısından da belirgin avantajlar sağlamaktadır. MCB yöntemi öne çıkan en güçlü aday olurken, OLA ve Rank yöntemleri de dengeli ve rekabetçi performanslarıyla dikkat çekmektedir.

Yapılan deneyler sonucunda farklı DSS yöntemlerinin performansı doğruluk, hassasiyet, özgüllük ve F1 skoru metrikleri açısından kapsamlı biçimde değerlendirilmiştir. Genel olarak, tüm metriklerde 50'den 300 özniteliğe kadar belirgin bir iyileşme gözlemlenmiş; 300-500 öznitelik aralığı, yöntemlerin en yüksek başarıyı gösterdiği kritik bölge olarak öne çıkmıştır. MCB yöntemi, özellikle yüksek öznitelik seviyelerinde tüm metriklerde en yüksek performansı sergileyerek en güçlü aday olmuştur. APosteriori ve OLA yöntemleri de genel olarak istikrarlı ve güçlü sonuçlar sunmuş, hassasiyet ve özgüllük açısından dengeli performans göstermiştir. LCA ve MLA yöntemleri orta düzey öznitelik sayılarında iyi sonuçlar üretirken, yüksek boyutlarda nispeten düşüş yaşamıştır. Havuz ortalamasına kıyasla tüm DSS yöntemleri, örnek bazlı karar mekanizması sayesinde metin sınıflandırma görevlerinde anlamlı iyileşmeler sağlamıştır. Elde edilen bulgular, dinamik sınıflandırıcı seçme stratejilerinin duygu analizinde hem doğruluk hem de denge açısından tekil sınıflandırıcılara kıyasla önemli avantajlar sunduğunu ortaya koymaktadır.

Sonuçlar ve Gelecekteki Çalışmalar

Bu çalışma kapsamında, e-ticaret ve mağaza değerlendirme platformlarında yer alan kullanıcı yorumları üzerinden duygu analizi gerçekleştirilmiş ve DSS yöntemlerinin sınıflandırma başarısına etkisi detaylı olarak incelenmiştir. Özellikle, sınıflandırıcı havuzunda yer alan temel makine öğrenmesi algoritmaları (k-NN, SVM, Karar Ağaçları, Naive Bayes ve Lojistik Regresyon) ile çeşitli DSS stratejileri (MCB, APosteriori, OLA, Rank, LCA, Apriori ve MLA) karşılaştırılmıştır. Deneyler 50'den 1000'e kadar değişen öznitelik sayıları için gerçekleştirilmiş ve doğruluk, hassasiyet, özgüllük ile F1 skoru gibi performans metrikleri kullanılarak analiz yapılmıştır.

Elde edilen sonuçlar, dinamik sınıflandırıcı seçme yöntemlerinin tüm metriklerde sınıflandırıcı havuzunun ortalama performansına kıyasla önemli iyileşmeler sağladığını ortaya koymuştur. Özellikle MCB yöntemi, yüksek öznitelik seviyelerinde tüm ölçütlerde en iyi sonuçları elde ederek öne çıkmıştır. APosteriori ve OLA yöntemleri ise tutarlı ve dengeli performanslarıyla dikkat çekmiştir. LCA ve MLA yöntemleri orta seviyede öznitelik sayısında güçlü sonuçlar sunarken, çok yüksek boyutlarda belirgin düşüşler yaşamıştır. Rank yöntemi ise farklı öznitelik seviyelerinde genel olarak rekabetçi bir performans sergilemiştir.

Öznitelik sayısına bağlı olarak tüm yöntemlerde genel bir performans artışı gözlemlenmiş, özellikle 300-500 öznitelik aralığında yöntemlerin en yüksek başarıyı sağladığı belirlenmiştir. Ancak 1000 öznitelik seviyesinde bazı yöntemlerde performansın stabilize olduğu veya hafif düşüşler yaşandığı görülmüştür. Bu durum, aşırı yüksek boyutların zaman zaman sınıflandırıcıların etkinliğini olumsuz etkileyebileceğini göstermektedir.

Sonuç olarak, dinamik sınıflandırıcı seçme yöntemleri, duygu analizi gibi örnek bazlı karar gerektiren metin sınıflandırma görevlerinde güçlü bir alternatif sunmakta ve sınıflandırma doğruluğunu anlamlı biçimde artırmaktadır. Çalışma, DSS yöntemlerinin yalnızca doğruluk değil, aynı zamanda hassasiyet, özgüllük ve F1 skoru gibi temel ölçütlerde de kayda değer gelişmeler sağladığını göstermektedir. Gelecekte, bu yaklaşımların daha büyük, çeşitli ve çok dilli veri setlerinde test edilmesi, ayrıca derin öğrenme tabanlı sınıflandırıcıların DSS çerçevesine entegre edilerek daha gelişmiş hibrit sistemlerin oluşturulması önerilmektedir.

Çıkar Beyanı**:** Yazar çıkar çatışması olmadığını beyan etmektedir.

REFERANSLAR

[1] Pang, B., & Lee, L. (2008). *Opinion mining and sentiment analysis*. Foundations and Trends® in Information Retrieval, 2(1–2), 1–135.
[2] Liu, B. (2012). *Sentiment analysis and opinion mining*. Synthesis Lectures on Human Language Technologies, 5(1), 1–167.
[3] Cambria, E., Schuller, B., Xia, Y., & Havasi, C. (2013). *New avenues in opinion mining and sentiment analysis*. IEEE Intelligent Systems, 28(2), 15–21.
[4] Sebastiani, F. (2002). *Machine learning in automated text categorization*. ACM Computing Surveys (CSUR), 34(1), 1–47.
[5] Kowsari, K., Meimandi, K. J., Heidarysafa, M., Mendu, S., Barnes, L. E., & Brown, D. (2019). *Text classification algorithms: A survey*. Information, 10(4), 150.
[6] Cruz, R. M. O., Sabourin, R., & Cavalcanti, G. D. C. (2018). *Dynamic classifier selection: Recent advances and perspectives*. Information Fusion, 41, 195–216.
[7] Ko, A. H. R., Sabourin, R., & Britto, A. S. (2008). *From dynamic classifier selection to dynamic ensemble selection*. Pattern Recognition, 41(5), 1718–1731.
[8] Cruz, R. M. O., Cavalcanti, G. D. C., & Sabourin, R. (2018). *Dynamic classifier selection: Recent advances and perspectives*. Pattern Recognition, 41(5), 195–216.
[9] Pang, B., Lee, L., & Vaithyanathan, S. (2002). Thumbs up? Sentiment classification using machine learning techniques. *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing*, 79–86.
[10] Kim, Y. (2014). Convolutional neural networks for sentence classification. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*,
[11] Tang, D., Qin, B., & Liu, T. (2015). Document modeling with gated recurrent neural network for sentiment classification. *Proceedings of EMNLP*, 1422–1432.
[12] Go, A., Bhayani, R., & Huang, L. (2009). Twitter sentiment classification using distant supervision. *CS224N Project Report*, Stanford University.
[13] Dashtipour, K., Ieracitano, C., Carlo Morabito, F., Raza, A., & Hussain, A. (2021). An ensemble based classification approach for persian sentiment analysis. *Progresses in Artificial Intelligence and Neural Systems*, 207-215.
[14] Bilen, B., & Horasan, F. (2022). LSTM Network based Sentiment Analysis for Customer Reviews. Politeknik Dergisi, 25(3), 959-966.
[15] Cruz, R. M., Hafemann, L. G., Sabourin, R., & Cavalcanti, G. D. (2018). DESlib: A dynamic ensemble selection library in python. *arXiv preprint arXiv:1802.04967*.
[16] Giacinto, G., & Roli, F. (1999). Methods for dynamic classifier selection. In *Proceedings 10th International Conference on Image Analysis and Processing* (pp. 659-664): IEEE.
[17] Woods, K., Kegelmeyer, W. P., & Bowyer, K. (1997). Combination of multiple classifiers using local accuracy estimates. *IEEE transactions on pattern analysis and machine intelligence, 19*, 405-410.

[18] Giacinto, G., & Roli, F. (2001). Dynamic classifier selection based on multiple classifier behaviour. *Pattern Recognition, 34*, 1879-1881.
[19] Smits, P. C. (2002). Multiple classifier systems for supervised remote sensing image classification based on dynamic classifier selection. *IEEE Transactions on Geoscience and Remote Sensing, 40*, 801-813.
[20] Sabourin, M., Mitiche, A., Thomas, D., & Nagy, G. (1993). Classifier combination for hand-printed digit recognition. In *Proceedings of 2nd International Conference on Document Analysis and Recognition (ICDAR'93)* (pp. 163-166): IEEE.
[21] Han, J.,2005. DataMining:Concepts and Techniques. *Morgan Kaufmann Publishers Inc., SanFrancisco, CA, USA.*

ENKAZ ARAMA ALANINDA KARINCA SİSTEMİ KULLANILARAK GELİŞTİRİLEN BİR ALGORİTMA İLE YAPILAN İNCELEME

Zeynep Karabulut*
Milli Savunma Üniversitesi, Temel Bilimler Bölümü, İstanbul, Türkiye

Mehmet Dinçer Erbaş
Bolu Abant İzzet Baysal Üniversitesi, Bilgisayar Mühendisliği Bölümü, Bolu, Türkiye
**(İletişimden sorumlu yazar) zeynep.karabulut@msu.edu.tr*

ÖZET

TDK' ya göre felaket, "Büyük zarar, üzüntü ve sıkıntılara yol açan olay veya durum; yıkım, katastrof" olarak tanımlanır. Felaketler, doğal ve insan kaynaklı olmak üzere iki ana gruba ayrılabilir. Doğal veya insan hatası sonucu oluşan afetler sonrasında büyük tahribatlar oluşabilmektedir. Oluşan bu afetlerden sonra can kaybını azaltmak için arama/ kurtarma çalışmaları gerçekleştirilir. Bu çalışmanın ilk bölümünde felaketler ile ilgili genel bir bakışa, ikinci bölümünde teleoperasyon mimarisine ve arama/ kurtarma operasyonlarında kullanımlarını içeren bazı çalışmalara, üçüncü bölümünde bazı sürü zekâsı algoritmalarına (parçacık sürü optimizasyonu ve karınca kolonisi optimizasyonu), stigmerji kavramına ve feromon kullanımını içeren bazı çalışmalara yer verilmiştir. Son bölümde ise enkaz altı sürü robotiği için önerilen yeni algoritma tanıtılmış ve yirmi çalıştırma sonrası elde edilen feromonlu ve feromonsuz hareketin oluşturduğu ortalama hedefe ulaşma sayılarını içeren bir grafiğe yer verilmiştir.

Anahtar Kelimeler: Sürü robotiği, stigmerji, karınca sistemi.

Giriş

Felaket, "Büyük zarar, üzüntü ve sıkıntılara yol açan olay veya durum; yıkım, katastrof" olarak tanımlanır [1]. Felaketler, doğal ve insan kaynaklı olmak üzere iki ana gruba ayrılır. Doğal felaketler, depremler, volkanlar, kasırgalar, seller ve yangınlar, insan kaynaklı olanlar, savaş, kirlilik, nükleer patlamalar, yangınlar, tehlikeli maddelere maruz kalma, patlamalar ve ulaşım kazalarıdır. Doğal veya insan hatası sonucu oluşan bu tür afetler büyük tahribatlara neden olabilmektedir [2]. Oluşan bu afetlerden sonra can kaybını azaltmak için arama/kurtarma çalışmaları gerçekleştirilir. Bu çalışmaların gerçekleştiği önemli operasyonel ortamlardan biri, deprem ve sel gibi doğal afetler ve patlamalar gibi insan hatası sonucu oluşan binaların yapısal olarak çökmesidir. Bu alanlarda canlı kalan afetzedelere ulaşım için ilk 24 saat çok önemlidir [3]. Afet alanlarında, asbest, toz, zehirli gazlar, tehlikeli maddeler, radyasyon ve aşırı sıcaklıkların varlığı kurtarma görevlilerinin bu alanlara erişimini güçleştirir [4]. Bu tür sebeplerden dolayı arama/kurtarma çalışmaları için robotlara gereksinim artar. Afetlerden etkilenen insanlara yardım etmek için insan-robot arama/kurtarma çalışmalarını

birlikte yürütür. Bu nedenle, insan hayatını kurtarmak için, çağdaş teknolojiyi (robotlar ve bilişsel sistemleri) arama/kurtarma çalışmalarına entegre etmek büyük önem taşır [5]. Teknik olarak arama/kurtarma çalışmalarında ilk robot kullanımı, 11 Eylül 2001'de Dünya Ticaret Merkezi (WTC) kulelerine yapılan saldırılar sonrasında olmuştur [6]. Yaşanan bu olaydan günümüze kadar, farklı boyut, fonksiyon ve iletişim çeşitlerine sahip robot türleri dünya çapında pek çok afette ve arama/kurtarma çalışmalarında kullanılmak üzere tasarlanmış/geliştirilmiştir [7,8].

Teleoperasyon Mimarisi

Yapısal çökmeler sonucu oluşan enkaz alanlarında arama/kurtarma için kullanılan yöntemlerden biri teleoperasyondur. Teleoperasyon yönteminde, enkaz altının araştırılması için kullanılan robot, uzaktan bir operatör tarafından kontrol edilir. Bu amaçla, robotun üzerinde yer alan kamera ve diğer sensörlerden gelen veriler, operatörler tarafından incelenir ve gerekli komutlar, kablosuz bağlantı yolu ile robota iletilir. Teleoperasyon mimarisi, sistemin etkili bir şekilde çalışmasını sağlayan bileşenler ve kontrol algoritmalarından oluşur. Bu mimaride, işletim altyapısı, iletişim altyapısı, operatör bileşenleri, hız, görüş, konum gibi çeşitli sensörler, farklı algoritma türleri ve aktüatörler vardır [9]. Yapısal çökme sonrası enkaz alanlarında arama aktivitesini yöneten operatörler üç boyutlu ortamda görsel geri bildirim yolu ile çalıştıklarında önemli algısal zorluklar yaşarlar. Bu durumdan dolayı operatörler, toz ve gaz yoğunluğunun yüksek olduğu enkaz alanlarında yönlerini ve ortam farkındalıklarını kaybeder [10]. Dolayısıyla, gerekli komutları alamayan robotlar ise enkaz alanında fiziksel olarak sıkışırlar. Bu durumu çözmek için teleoperasyon ile üretilen özgün yöntemlerde kontrol döngüsünde operatör tutulur ve bunun yanı sıra robotların otonom özellikleri arttırılır [11]. Afetzedenin varlığının tespiti, harita oluşturmak, görsel gözlem ve görüş sürecini yürütmek, keşif gibi görevleri içeren yarı otonom-ayarlanabilir otonomi kontrol mimarisinin kullanımı üzerine bazı çalışmalar mevcuttur [10,11].

Bazı Sürü Algoritmaları ve Stigmerji

Enkaz altı gibi arama ortamları için geliştirilen çoklu robot sistemleri sürü zekâsı algoritmalarından yararlanır. Sürü zekâsı algoritmaları kullanımının bazı güçlü yönleri mevcuttur. Bunlardan ilki, daha yüksek ölçeklenebilirliğe olanak tanıyan esnek birey sayısına sahip olmaları, ikincisi, nispeten büyük ölçekli bir arama gerçekleştirme ve çözümleri tek bir arama sürecinde iyileştirmeye olanak tanımaları, üçüncüsü, kendi aralarında yerel olarak etkileşime giren birkaç bireyin, sonuç olarak karmaşık bir global davranışa yol açmaları ve sonuncusu ise, sürünün esas olarak, merkezi kontrol olmadan çalışması ve işleyişini sürdürebilmesi için hiçbir belirli bireyin vazgeçilmez olmamasıdır [12]. Buna karşın, sürü zekâsı algoritmalarının sınırlılıkları da vardır. Bunlardan birincisi, popülasyon büyüklüğü, yineleme sıklığı ve şekli gibi faktörlerden etkilenen zaman alıcı süreçler olmaları, ikincisi, merkezi koordinasyon eksikliğinden ötürü durgunluk ve yerel optimuma düşmenin zamanından önce gerçekleşme sorunu ve sonuncusu ise, koordinasyon mekanizmalarının bloke edilmesinin hesaplama sürelerini arttırmasıdır [12]. Son yıllarda yapılan araştırmalarda, sürü zekâsı algoritmaları içerisinde birinci sırada Parçacık Sürü Optimizasyonu (PSO), ikinci sırada ise Karınca Kolonisi Optimizasyonu (ACO) kullanımı gelmektedir [12].

Parçacık Sürü Optimizasyon algoritması

Parçacık sürü optimizasyonunun kökenleri, kuş sürüsü, balık sürüsü ve sürü teorisi ile bağlarından dolayı yapay yaşam ve evrimsel hesaplamaya dayanır [13]. Bu algoritmaya göre her robot bir parçacığı gösterir. Robotlar, uygunluk değerlerini (İng: fitness function), robot topluluğunun uygunluk değeri ile kıyaslar ve kendi yer değiştirmelerini, grubun en yüksek uygunluk değerine sahip robota yaklaşacak şekilde günceller.

Tablo 1: Parçacık Sürü Optimizasyon tabanlı algoritmaları içeren bazı çalışmalar

Referans	Algoritma	Görev	Test Tipi
[14]	Robotik Darwinci Parçacık Sürü Optimizasyonu (RDPSO), Genişletilmiş Parçacık Sürü Optimizasyonu (EPSO), Fiziksel olarak gömülü Parçacık Sürü Optimizasyonu (PPSO)	Arama	MRSim (Multi Robot Simulator) simülatörü kullanılmıştır. 14 e-puck robot ile deney gerçekleştirilmiştir.
[15]	Robotik Parçacık Sürü Optimizasyonu (RPSO), Robotik Darwinci Parçacık Sürü Optimizasyonu (RDPSO)	Arama ve Kurtarma	Simülasyon için Gazebo kullanılmıştır. Görselleştirme için Rviz kullanılmıştır.
[16]	Dağıtılmış Parçacık Sürü Optimizasyonu (DPSO)	Arama kurtarma ortamında afetzede tespiti ve çok robotlu bir sistemin navigasyonu	Simülasyonlar için, Python ile bağlantılı VRep kullanılmıştır.

Tablo 1'de, Parçacık Sürü Optimizasyonu ve ona bağlı geliştirilen bazı algoritmalar yer alır. Parçacık sürü optimizasyon algoritması ethernet, radyo ve kablosuz gibi bazı ortamları gerektiren direkt iletişim kanallarına ihtiyaç duyar [17]. Yöntemin direkt iletişime sahip olması, dezavantajlarından biridir. Bir diğeri ise, global topluluk içerisinde global bilgi alışverişini kullanmasıdır. Bu yöntem ve türevleri, enkaz altı araştırmaları için ele alındığında, ortamda yüksek oranda metal ve benzeri alaşımlar ile beton kütleleri yer alabileceği için kablosuz iletişime dayalı direkt mesajlaşma yöntemleri yeterli dayanıklılığa sahip olmaz.

Stigmerji

Robotlar arasında bilgi alışverişinin sınıflandırılması için, doğrudan ve dolaylı iletişim ayrımı kullanımlardan biridir. Doğrudan iletişim, bazı yerleşik donanım aygıtlarını kullanırken, dolaylı iletişim stigmerjiyi kullanır [18]. Dolaylı iletişim mekanizması olan stigmerji (İng: Stigmergy), bazı bireylerin gerçekleştirdiği eylemlerin sonucunda çevrenin aracılık ettiği ve bir başkasının davranışını dolaylı olarak etkileyen bir iletişim şeklidir [19]. Yaban arıları, termitler ve karıncalar gibi bazı sosyal canlılarda görülür. Karıncalar ve termitlerin çevreye bıraktıkları mesajlar, bir kimyasal madde olan feromon izleri biçimindedir [20]. Stigmerji tabanlı iletişim sistemi, sürü zekâsı alanında geliştirilen algoritmalardan biri olan ve birçok farklı problemde verimli sonuçlar üreten Karınca Kolonisi Optimizasyonu algoritmasının temelini oluşturur [21].

Karınca Kolonisi Optimizasyon algoritması

Karınca kolonilerinde stigmerji örnekleri gözlemlenir. Birçok karınca türü, bir besin kaynağına gidip gelme sırasında yere feromon adı verilen bir madde bırakır. Diğer karıncalar yere bırakılan bu feromonu algılar ve konsantrasyonun yüksek olduğu yolu tercih eder.

Bu yöntem, karıncaların, yiyecekleri yuvalarına etkin bir şekilde taşımasını sağlar. Karıncaların yiyecek arama davranışını açıklamak için önerilen modelden ilham alarak geliştirilen Karınca kolonisi optimizasyonu, bir dizi yapay karınca kullanarak, ele alınan optimizasyon problemine çözümler üretir. Bu metasezgisel algoritma üç aşamalı olarak tekrarlanır. Her yinelemede, karıncalar tarafından bir dizi çözüm oluşturulur. Daha sonra bu çözümler yerel arama ile iyileştirilir (bu adım isteğe bağlı gerçekleştirilir) ve son olarak feromon güncellenir. Feromon konsantrasyonu, algoritmanın stokastik olarak yürütülmesine olanak tanır. Çözüm bileşenleri için stokastik seçim kuralı farklı karınca kolonisi algoritmaları için değişiklik gösterse de hepsi, gerçek karıncaların davranış modelini temel alan bir bağıntıdan (Bkz: [21]; Syf:29; Denklem (1)) faydalanır [21].

Tablo 2: Karınca Kolonisi Optimizasyonu tabanlı algoritmaları içeren bazı çalışmalar

Referans	Algoritma	Görev/Amaç	Test Tipi
[22]	Karınca kolonisi optimizasyonu algoritması (ACO)	Kimyasal gaz konsantrasyonu haritasını oluşturmak ve yüksek gaz konsantrasyonuna sahip alanları belirlemek.	Deneyler laboratuvar ortamında 3 Khepera III ve bağımsız olarak 5 MiniQ mobil robotları kullanılarak gerçekleştirilmiştir.
[23]	Değiştirilmiş Karınca kolonisi algoritması	Küçük otonom robotlardan oluşan bir robot sürüsü kullanarak ACO'nun küçük ölçekli bir uygulamasına yer vermektir	Deneyler, SimpleIDE yazılımı ile C dilinde programlanabilen 3 tane Propeller Activity Bot robotları kullanılarak gerçekleştirilmiştir.
[24]	Robot hareketi ve yönelimi için: bir bulanık mantık denetleyici -FLC- (İng: Fuzzy logic controller), bir dinamik bulanık bilişsel harita -DFCM- (İng: Dynamic fuzzy cognitive map), karınca koloni optimizasyonu metasezgise-linden esinlenilen bir dinamik bulanık bilişsel harita -DFCM-ACO (İng: Dynamic fuzzy cognitive map- Ant colony optimization) kullanılmıştır.	Simüle edilmiş ortamlarda arama görevlerini yerine getirebilen otonom bir çoklu robot sistemi geliştirmektir. Çoklu robot sisteminin üstlendiği üç görev şu şekildedir: statik ve dinamik (diğer robotlar) engelleri tespit etmek ve bunlardan kaçınmak, tüm afetzedeleri bulmak ve kurtarmak, ortamları haritalamaktır.	Simülasyon için Matlab kullanılmıştır. Üç simülasyon ortamı kullanılmıştır. Simülasyon ortamı 8 robot ve 6 afetzedeyi içerir.
[25]	Karınca algoritması ve karınca algoritmasının üç farklı biçimi kullanılmıştır. Bunlar, ters karınca algoritması -iAA (İng: inverted Ant Algorithm), işaret başlatmalı ters karınca algoritması -iAA-B (İng: inverted-AA with beacon initialization), algılama aralığı artırılmış ters karınca algoritması -iAA-R (İng: inverted-AA with an increased sensing range) dir.	Karınca algoritmalarının -AA (İng: Ant algorithms), sıkışmış bir afetzedeye giden bir yol için arama süresini optimize etmede uygulanabilirliğini ve kurtarma aşamasında Dijkstra algoritmasının uygulamasını araştırmaktır.	Simülasyon için Matlab kullanılmıştır. Simülasyonda beş farklı labirent benzeri ortam ve sayıları 100,200,300,400,500,600 ile değişen grup büyüklükleri yer alır.

Tablo 2'de Karınca kolonisi optimizasyonu ve ona bağlı geliştirilen algoritmaları içeren bazı çalışmalar yer almaktadır.

Feromon

Stigmerji mekanizması, robotlar arasında bilgi aktarımını sağlamak için ortama bırakılan feromonu kullanır. Feromon kullanımı, robotların iletişim yeteneği gereksinimini azaltarak sürüyü daha ölçeklenebilir hale getirir [17]. Bu mekanizma sayesinde sürü robotları hedefi daha kısa sürede bulur ve farklı sayıda robot ile mekanizmanın uygulanabilirliği devam eder [17]. Sistemin uygulanabilirliğinin devam etmesi ölçeklenebilirliğinin bir sonucudur. Oluşan bu durum, stratejiyi farklı karmaşık senaryolarda da kullanılabilir hale getirir [17]. Karmaşık senaryonun oluştuğu durumlardan biri de enkaz altı arama/kurtarma çalışmalarıdır. Stigmerji stratejisini kullanan robot toplulukları, direkt mesajlaşma protokollerinin kesintiye uğrayarak kullanım dışı kaldığı, enkaz altı gibi ortamlarda avantaj sağlar. Böylece, direkt bir kablosuz iletişime ihtiyaç duymayan robot topluluklarının boyut ve enerji tüketimi de azaltılabilmektedir. Buradan sağlanan katkı, daha fazla sayıda, küçük boyutlu robot toplulukları kullanımı ile, enkaz altı arama/kurtarma operasyonlarında yüksek öneme sahip olan robot topluluklarının bireysel hata ve bozulmalarına karşı dayanıklılıklarının arttırılması olur. Bu önemli teorik ve pratik avantajların etkisiyle, son dönemde arama/kurtarma amaçlı oluşturulan robot topluluklarında stigmerji kullanıma yönelik araştırmalar yapılmıştır. Yapılan son araştırmalar [26], feromon izlerinin ziyaret edilen alanlara ilişkin bir tür sürü zekâsı sağladığı yönündedir.

Tablo 3: Feromon kullanımını içeren bazı çalışmalar

Referans	Feromon Türü (varsa modeli)	Görev/Amaç	Test Tipi
[27]	Sanal feromon	GÖZETleme, keşif, tehlike tespiti ve yol bulma gibi konularda büyük ölçekli sonuçlar elde etmek için çok sayıda küçük ölçekli robotun eylemlerini koordine etmeye yönelik teknikleri araştırmak.	20 ferobot sürü robotu kullanılmıştır. Kontrol sistemi Java'da ve algoritmaların simülasyonu Teambots ortamında yazılmıştır.
[28]	Feromon görevi yapan bir kimyasal olarak etanol kullanılmıştır.	Robot sürülerinde grup arama davranışını gerçekleştirmek. Feromon görevi yapan bir kimyasal tarafından koordine edilen bir robot sürüsünün işbirlikçi davranışını hem simülasyonlarda hem de gerçek robot deneylerinde test etmek.	Feromon iletişiminin varlığının etkisi için sayıları 1-40 arasında değişen robot sürüsü, feromonlu ve feromonsuz iletişim için sayıları 10,40, 90,160 ve 250 arasında değişen robotlar kullanılmıştır. Simülasyon sonuçlarını doğrulamak için ARGOS-01 robotu geliştirilmiştir. Bu aşamada kullanılan robot sürüsü büyüklüğü 1,2,3,4,7 ve 10 olarak değişir.
[29]	İtici feromon	Stigmerjik etkileşim kuralının işleyişi ve stigmerjik alan kapsama algoritmasının kollektif performansını araştırmaktır.	Görüntü işleme için Matlab, simülator olarak Kilobox, deneysel uygulama için Kilobot robotları kullanılmıştır. Simülasyon için 400'e kadar, deneysel olarak 324'e kadar sayıda robot kullanılmıştır.

Tablo 3' de feromon kullanımını içeren bazı çalışmalar yer almaktadır. Yukarıda belirtilen çalışmalarda arama/kurtarma operasyonları için oluşturulmuş çoklu robot sistemlerinde stigmerji kullanımının mümkün olduğu gösterilmiş olmakla birlikte, fiziksel feromon salınımı ve enkaz altında çalışan bir robot sistemi oluşturulması gözlenmemiştir. Belirtilen çalışmalarda

çoğunlukla simülasyon üzerinde ve 2 boyutlu ortamlarda robot sistemleri test edilmiştir. Fiziksel feromon benzeri salınımların ve algılamanın mevcut olmaması, bu çalışmalarda geçen yöntemlerin gerçek dünya senaryolarına uygulamalarını güçleştirir. Fiziksel feromon ve benzeri salınımları içeren stigmerji tabanlı yaklaşımların kullanımını içeren az sayıda çalışma mevcuttur. Bu çalışmalardan biri olan [30], karıncaların feromon izi takip yeteneklerinin robotik sistemlere aktarılmasını içeren bir proje araştırmasıdır. Bu çalışmada, bir çift koku algılayıcı anten ile donatılan robotik bir karınca inşa edilmiş ve *Lasius fuliginosus* karıncasında gözlenen iz takip algoritması ile sistem programlanarak deneysel sonuçlara yer verilmiştir [30]. Tablo 3' de yer alan ve bu alanda yapılan ender araştırmalardan biri olan [28] de, ARGOS-01 robotlarından oluşan bir robot topluluğunun ethanol kimyasal maddesi salınımı sayesinde stigmerji yoluyla haberleşebildikleri ve 2 boyutlu bir robot arenasında basit bir hedef nokta bulma problemini kümülatif olarak çözebildikleri gösterilmiştir [28]. Öte yandan ilgili robot topluluğu ve çalışmada geliştirilen stigmerji tabanlı kontrol algoritması, enkaz benzeri herhangi bir 3 boyutlu alanda arama/kurtarma amaçlı bir görev çerçevesinde test edilmemiştir.

Önerilen Çalışma ve Bazı Bulgular

Yukarıda incelenen literatür göz önüne alındığında enkaz altı gibi ortamların oluşması durumunda teleoperasyon yöntemi ve doğrudan iletişim kullanan sürü optimizasyon tabanlı algoritmaların kullanımı işlevsel olmayacaktır. Görüntü işlemenin ve robotlar arası veri transferinin zor olduğu durumlarda kullanılmak üzere tam otonom özgün bir kontrol algoritması bu çalışmada önerilmektedir. Geliştirilen yeni algoritma, klasik karınca sistemi algoritmasına yeni iki terim eklenerek (olasılıksal seçim yöntemine göre ses yüksekliği ve lokal etkileşimler sonucu elde edilen sayım bilgisi) oluşturulmuştur. Bu algoritmaya göre numaralı robotun arama hücresinden arama hücresine geçme ihtimali

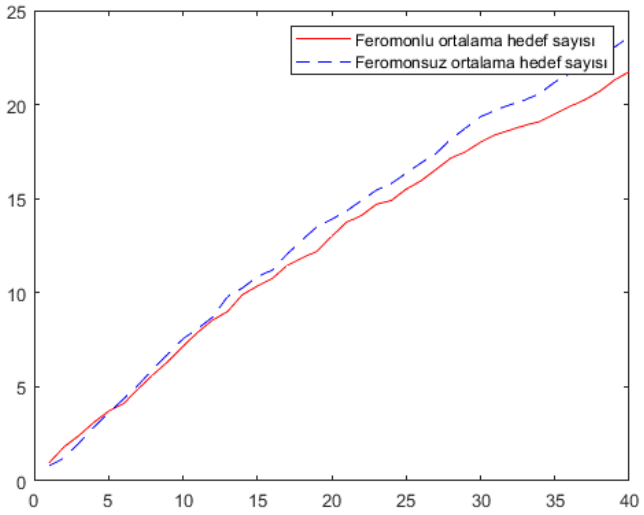
$$P_{i,j}^{k} = \begin{cases} \dfrac{\tau_{i,j}^{\sigma}.\pi_{i,j}^{\delta}.\eta_{i,j}^{\beta}}{\sum_{l \in S_k^i} \tau_{i,l}^{\sigma}.\pi_{i,l}^{\delta}.\eta_{i,l}^{\beta}} & j \in S_k^i \text{ ise} \\ 0 & \text{aksi halde} \end{cases} \quad (1)$$

ile gösterilir. Denklem (1) de, S_k^i, i arama hücresinde geçiş yapılabilecek, henüz robotu tarafından ziyaret edilmemiş komşu arama hücrelerinin kümesini, $\tau_{i,j}^{\sigma}$, algılanan feromon miktarını, ise algılanan ses yüksekliğini gösterir. σ, δ ve β, sırasıyla, feromon miktarı, ses yüksekliği ve sayım bilgisinin görece ağırlığını kontrol ederler. Denklem (1) de yer alan $\eta_{i,j}$ sayım bilgisi,

$$\eta_{i,j} = \frac{1}{1+r_{i,j}} \quad (2)$$

ile hesaplanır. Denklem (2) de yer alan $r_{i,j}$, i arama hücresinden j arama hücresine geçiş yapacağını belirten robot sayısını gösterir.

Şekil 1 de, yirmi çalıştırma sonucu t=4000 için her t=100 zaman diliminde bir elde edilen feromonlu ve feromonsuz olarak ortalama hedefe ulaşma sayıları için oluşturulan grafik yer almaktadır.



Şekil 1: Feromonlu ve feromonsuz olarak 20 çalıştırma sonrası elde edilen ortalama hedef bulma verileri ile oluşturulan grafik

Grafik üzerinden yirmi çalıştırma sonucu elde edilmiş olan ortalama hedef bulma sayıları göz önüne alındığında, feromonlu ve feromonsuz şekilde hareket için daha farklı ve fazla sayıda test işleminin gerçekleşmesi sonucunda, yöntemlerin avantaj ve dezavantajlarının daha iyi değerlendirileceği noktasında daha fazla katkı sağlanacaktır. Bu doğrultuda çalışmalarımız devam etmektedir.

Çıkar Beyanı**:** Yazarlar arasında çıkar çatışması yoktur.

Yazar Katkıları: Bildirinin hazırlanmasında yazarlar eşit derecede katkı vermiş olup, yazım aşaması Zeynep Karabulut, kontrol aşaması ise Mehmet Dinçer Erbaş'a aittir.

REFERANSLAR

[1] Türk Dil Kurumu (TDK), https://sozluk.gov.tr/, (09.04.2025 tarihinde erişildi).

[2] J. Zibulewsky, "Defining disaster: The emergency department perspective", In Baylor University Medical Center Proceedings, 14, 2, 144-149, 2001.

[3] D.P. Stormont, "Autonomous rescue robot swarms for first responders", In CIHSPS 2005. Proceedings of the 2005 IEEE International Conference on Computational Intelligence for Homeland Security and Personal Safety, March 31- April 1, 2005, Orlando.

[4] Y. Liu, and G. Nejat, "Robotic urban search and rescue: A survey from the control perspective", Journal of Intelligent & Robotic Systems, 72, 147-165, 2013.

[5] H. Chitikena, F. Sanfilippo, and S. Ma, "Robotics in search and rescue (SAR) operations: An ethical and Design Perspective Framework for Response Phase", Applied Sciences, 13, 3, 1800, 2023.

[6] J. Casper, and R. R. Murphy, "Human-robot interactions during the robot-assisted urban search and rescue response at the World Trade Center", IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 33, 3, 367-385, 2003.

[7] A.W. Ko, and H.Y. Lau, "Intelligent robot-assisted humanitarian search and rescue system", International Journal of Advanced Robotic Systems, 6, 2, 121-128, 2009.

[8] C. Cruz Ulloa, G. Prieto Sánchez, A. Barrientos, and J. Del Cerro, "Autonomous thermal vision robotic system for victims recognition in search and rescue missions", Sensors, 21, 7346, 2021.

[9] S. Opiyo, J. Zhou, E. Mwangi, W. Kai, and I. Sunusi, "A review on teleoperation of mobile ground robots: Architecture and situation awareness", International Journal of Control, Automation and Systems, 19, 1384-1407, 2021.

[10] B. Doroodgar, Y. Liu, and G. Nejat, "A learning-based semi-autonomous controller for robotic exploration of unknown disaster scenes while searching for victims", IEEE Transactions on Cybernetics, 44, 12, 2719-2732, 2014.

[11] A. Finzi, and A. Orlandini, "Human-robot interaction through mixed-initiative planning for rescue and search rovers", 9th Congress of the Italian Association for Artificial Intelligence, September 21-23, 2005, Milan.

[12] J. Tang, G. Liu, and Q. Pan, "A review on representative swarm intelligence algorithms for solving optimization problems: Applications and trends", IEEE/CAA Journal of Automatica Sinica, 8, 10, 1627-1643, 2021.

[13] J. Kennedy, and R. Eberhart, "Particle swarm optimization", In Proceedings of ICNN'95- International conference on neural networks, November 27- December 1, 1995, Perth.

[14] M.S. Couceiro, P.A. Vargas, R.P. Rocha, and N.M. Ferreira, "Benchmark of swarm robotics distributed techniques in a search task", Robotics and Autonomous Systems, 62, 2, 200-213, 2014.

[15] A.S. Kumar, G. Manikutty, R.R. Bhavani, and M.S. Couceiro, "Search and rescue operations using robotic darwinian particle swarm optimization", In 2017 International Conference on Advances in Computing Communications and Informatics (ICACCI), September 13-16, 2017, Udupi.

[16] D. Paez, J.P. Romero, B. Noriega, G.A. Cardona, and J.M. Calderon, "Distributed particle swarm optimization for multi-robot system in search and rescue operations", IFAC-PapersOnLine, 54, 4, 1-6, 2021.

[17] Q. Tang, Z. Xu, F. Yu, Z. Zhang, and J. Zhang, "Dynamic target searching and tracking with swarm robots based on stigmergy mechanism", Robotics and Autonomous Systems, 120, 103251, 2019.

[18] A. Farinelli, L. Iocchi, and D. Nardi, "Multirobot systems: a classification focused on coordination", IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), 34, 5, 2015-2028, 2004.

[19] V. Trianni, T.H. Labella, and M. Dorigo, "Evolution of direct communication for a swarm-bot performing hole avoidance", In Ant Colony Optimization and Swarm Intelligence: 4th International Workshop, ANTS 2004, September 5-8, 2004, Brussels.

[20] M. Salman, D. Garzón Ramos, K. Hasselmann, and M. Birattari, "Phormica: Photochromic pheromone release and detection system for stigmergic coordination in robot swarms", Frontiers in Robotics and AI, 7, 591402, 2020.

[21] M. Dorigo, M. Birattari, and T. Stutzle, "Ant colony optimization", In IEEE Computational Intelligence Magazine, 1, 4, 28-39, 2006.

[22] M. Turduev, G. Cabrita, M. Kırtay, V. Gazi, and L. Marques, "Experimental studies on chemical concentration map building by a multi-robot system using bio-inspired algorithms", Autonomous agents and multi-agent systems, 28, 1, 72-100, 2014.

[23] T. Krentz, C. Greenhagen, A. Roggow, D. Desmond, and S. Khorbotly, "A modified ant colony optimization algorithm for implementation on multi-core robots", In 2015 Swarm/Human Blended Intelligence Workshop (SHBI), September 28-29, 2015, Cleveland.

[24] M. Mendonça, R.H. Palácios, E.I. Papageorgiou, and L.B. de Souza, "Multi-robot exploration using dynamic fuzzy cognitive maps and ant colony optimization", In 2020 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE), July 19-24, 2020, Glasgow.

[25] Z. Husain, A. Al Zaabi, H. Hildmann, F. Saffre, D. Ruta, and A.F. Isakovic, "Search and rescue in a maze-like environment with ant and dijkstra algorithms", Drones, 6, 10, 273, 2022.

[26] F. Fossum, J.M. Montanier, and P.C. Haddow, "Repellent pheromones for effective swarm robot search in unknown environments", In 2014 IEEE Symposium on Swarm Intelligence, December 9-12, 2014, Orlando.

[27] D. Payton, M. Daily, R. Estowski, M. Howard, and C. Lee, "Pheromone robotics", Autonomous Robots, 11, 3, 319-324, 2001.

[28] R. Fujisawa, S. Dobata, K. Sugawara, and F. Matsuno, "Designing pheromone communication in swarm robotics: Group foraging behavior mediated by chemical substance", Swarm Intelligence, 8, 3, 227-246, 2014.

[29] E.R. Hunt, S. Jones, and S. Hauert, "Testing the limits of pheromone stigmergy in high-density robot swarms", Royal Society open science, 6, 11, 190225, 2019.

[30] R.A. Russell, "Ant trails-an example for robots to follow?", In Proceedings 1999 IEEE International Conference on Robotics and Automation, May 10-15, 1999, Detroit.

ETKİN KAYNAK KULLANIMI AMACI İLE GÜVENLİ BİR API YÖNETİM MİMARİSİ TASARIMI

Muhammet Emin
Medyasoft, Yazılım, İstanbul, Türkiye

Gökçen Gökyer Nalcı
Medyasoft, Yazılım, Ankara, Türkiye

Özlem Emin
Medyasoft, Yazılım, İstanbul, Türkiye

Furkan Baran
Medyasoft, Yazılım, İstanbul, Türkiye

Ali Doğan
Medyasoft, Yazılım, İstanbul, Türkiye

Mustafa Kozan
Medyasoft, Yazılım, İstanbul, Türkiye

Önder Demir*
Marmara Üniversitesi, Teknoloji Fakültesi, İstanbul, Türkiye
**(İletişimden sorumlu yazar) odemir@marmara.edu.tr*

ÖZET

Dijital dönüşüm süreçleri birçok yazılımın birbirine entegre olmasını, veri paylaşımı ve aktarımını gerekli kılmıştır. Tüm yazılım türlerini ve uygulamalarını kapsayacak bir standardizasyon uygulanamayacağından yazılım entegrasyonları için arayüz yazılımları geliştirilmiştir. Bir Uygulama Programlama Arayüzü (Application Programming Interface – API) farklı yazılım uygulamalarının birbirleriyle iletişim kurmasını sağlayan bir dizi kural, protokol ve araçtan oluşan yazılımlardır. Web hizmetlerine, mobil uygulamalara, bulut bilişime hatta IoT cihazlarına kadar çeşitli yazılım uygulamalarında kullanılan API'lar farklı yazılım sistemlerinin dahili uygulama şekillerini bilmeden etkileşime girmelerini sağlar. İletişimi standartlaştırmak amacıyla genellikle JSON veya XML formatlarında veri aktarımı ve HTTP, REST, SOAP ile istek gönderimi sağlanır. Bu çalışmada API yönetimini merkezi bir noktada sağlayarak güvenlik, erişim kontrolü, yönlendirme ve yönetim, protokol dönüşümü, izleme ve analiz gibi önemli işlevleri yerine getirecek bir API Gateway uygulaması geliştirilmesi amaçlanmaktadır. Ayrıca, uygulamanın belirli bir kullanıcı trafiği için ek sunucu gerektirmeksizin çalışabilecek şekilde tasarlanması, yeşil mutabakat ilkeleri çerçevesinde enerji tasarrufu sağlamayı ve buna bağlı olarak entegre edilen sistemin karbon ayak izini azaltarak verimliliği artırmayı hedeflemektedir.

Anahtar Kelimeler: Uygulama programlama arayüzü, API, yük dengeleme.

Giriş

Uygulama Programlama Arayüzü (API), yazılım sistemleri için veri alışverişi hizmetleri arasındaki etkileşime izin veren bir dizi rutin, prosedür, kaynak ve protokolden oluşan yazılımsal yapılarıdır [1]. API'ler Web uygulamaları, İşletim Sistemleri (OS), Veri Tabanları ve Donanım dahil olmak üzere geniş bir hizmet yelpazesinde kullanım alanı bulmaktadır. API'ler özellikle web teknolojisini kullanan uygulamaların arasındaki iletişimi düzenlemek için geliştirilen yazılım teknolojileridir [1-3]. Son zamanlarda, kullanıcılar, uygulamalar ve sistemler arasında artan birbirine bağlı olma durumu ve gereksinimi hepsi geliştirilen API'lerin sağladığı arayüzler desteğiyle çözülmektedir. Bu arayüzler artık dijital ekosistemin temel çerçevesi olarak hizmet vererek, değer üreten yenilikçi uygulamaların geliştirilmesini sağlayarak endüstriler ve ekonomiler arasında bağlantı kurmaktadır [4].

API kullanımının önemini artıran bir diğer önemli etken yazılım mimarilerinde hizmet odaklı mimarinin (Service Oriented Architecture – SOA) monolitik yapısının mimari Birçok modern yazılım sistemi, sürdürülebilirliği, ölçeklenebilirliği ve hata toleransını iyileştirmek için bir şekilde bağlanmış hizmetlere bölünmüştür [5-7]. Hizmet Odaklı Mimari, büyük bir monolitik yazılım sistemini birden fazla küçük hizmete dağıtmak için kullanılan, merkezi iletişime dayalı bir yöntemdir. Birbirine bağlı servislerde değişiklik yapmak sistem hizmetleri genelinde entegrasyon ve dağıtım koordinasyonu gerektiriyordu. Bu durum şirketlerin SOA'dan Mikro Hizmet Mimarisine (MSA) geçmesine yol açtı [8, 9]. MSA, paylaşılan hizmet modellerini, her bir mikro hizmet için yalnızca API'ler aracılığıyla açığa çıkan bağımsız etki alanı modelleriyle ve mesajları ileten mesaj aracılarıyla değiştirdi. İşlevselliklerini API'ler aracılığıyla ortaya koyan mikro hizmetlere sağlayıcı, bu API'leri çağıran ve bunlarla etkileşime giren mikro hizmetlere ise tüketici adı verilir. MSA'daki iletişim yaklaşımı mikro hizmetleri bu tür iyi tanımlanmış API'ler aracılığıyla birbirine bağlayarak bir sistem içinde bağımsız olarak geliştirilebilmelerine olanak tanır. Bununla birlikte, her bir mikro hizmetin veri yapıları ve iş mantığı, davranışları bağımsız olarak geliştiğinden, genel sistem işlevselliğini korumak, geliştirme ekipleri arasında daha fazla senkronizasyon çabası gerektirir. Sağlayıcı ekipleri tüketici ekiplerini değişiklikler hakkında önceden bilgilendirmezse, sorun yalnızca çalışma zamanındaki ilk gerçek API çağrısı sırasında ortaya çıkar. Sağlayıcının API'sindeki değişiklikler beklenmedik davranışlara neden olabilir ve bu API ile etkileşime giren bağımlı tüketicilerin yürütülmesini potansiyel olarak bozabilir [9-11].

API yönetimi genellikle manuel olarak gerçekleştirilmekte veya dağınık bir şekilde farklı hizmetler tarafından sağlanan küçük çözümler kullanılmaktadır. API'lerin güvenliği, erişim kontrolü, yönlendirme ve izleme gibi önemli işlevler genellikle her bir hizmetin kendine özgü bir şekilde ele alınmakta, tümleşik bir çözüm sunulamamaktadır. Bu durum, güvenlik, erişim kontrolü karmaşıklığı, trafiği yönlendirme sorunu, protokol uyumsuzluğu ve performans izleme konusunda sorunlara yol açabilmektedir [11-13].

Bu çalışmada API yönetimini merkezi bir API Gateway yazılımı üzerinden sağlayarak; güvenlik, erişim kontrolü, trafik yönlendirme, protokol dönüşümü ve performans izleme gibi tüm önemli işlevlerin bir araya getirildiği bütünleşik bir sistemin geliştirilmesi hedeflenmektedir.

Sistem Mimarisi

Çalışma kapsamında, sisteme veya hizmetine gelen giriş isteklerinin yönetilmesi, kontrol edilmesi ve yönlendirilmesini sağlayan, aynı zamanda, kapsamı fazlara ayıracak şekilde, Güvenlik, Erişim Kontrolü, Protokol Dönüşümü ile İzleme ve Analiz özelliklerini dahil edecek kapsamlı bir API Gateway mimarisi tasarlanmıştır. Bu tasarımın gerekçeleri; devam etmekte olan mevcut operasyonel süreçleri daha verimli, dinamik ve güvenli şekilde yönetilmesi, mevcut kapasitenin yükseltilmesi, sunuculara gelen önceliği düşük isteklerin fazla kaynak ve enerji kullanımına sebep olmasının önlenmesidir.

Mimarı Tasarımın Amacı

Geliştirilmek sistemin amacı, API yönetimi için bir merkezi nokta sağlamaktır. API Gateway olarak adlandırılan bu sistem, gelen istekleri doğrulamak, yetkilendirme yapmak, trafik yönlendirmesi yapmak, protokol dönüşümü gerçekleştirmek ve API kullanımını izlemek gibi bir dizi önemli görevi üstlenmektedir.

1.Güvenlik: API anahtarları, JWT (JSON Web Token) doğrulaması veya OAuth gibi mekanizmalarla gelen istekleri doğrulamak ve yetki kontrolü yapılacaktır. Böylece yetkisiz erişimi önleme ve güvenlik tehditlerine karşı koruma sağlanır.

2.Erişim Kontrolü: Belirli kullanıcı gruplarına veya istemcilere belirli API'lere erişim ve kontrol izni tahsis edilebilecektir. Böylece hassas verilere sadece yetkilendirilmiş kullanıcıların erişmesini sağlanır.

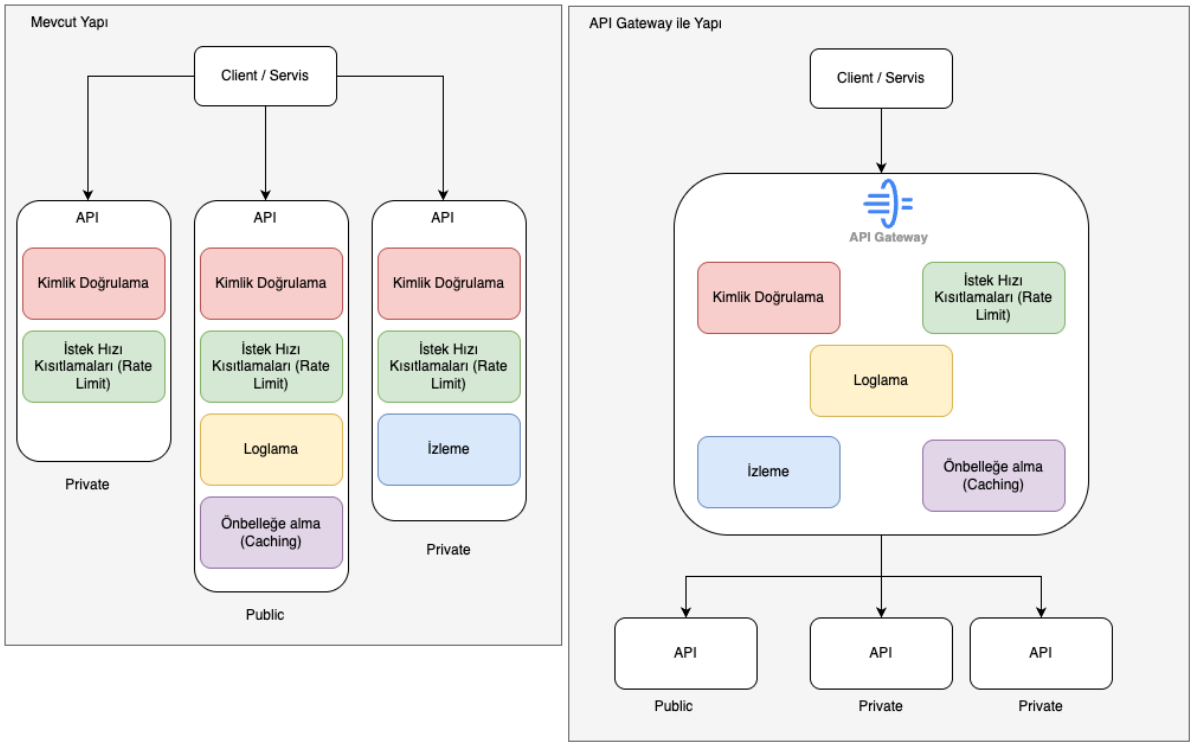
3.önlendirme ve Yönetim: Gelen istekleri birden çok hedef arasında en uygun olana yönlendirilmesidir. Bu, yük dengelemesi, trafik yönlendirmesi ve hata durumlarında geriye düşme gibi senaryoları içerir. Trafiğin dengeli bir şekilde yönlendirilmesini sağlayacak bir yapı kurulması hizmet reddi gibi saldırılara karşı bir önlem olacaktır.

4.Protokol Dönüşümü: Farklı protokoller arasında dönüşüm yapılması sağlanacaktır. Örneğin, HTTP isteklerini HTTPS'e dönüştürmek veya SOAP isteklerini RESTful API'lere dönüştürülmesinin sağlanacaktır.

5.İzleme ve Analiz: Gelen ve giden trafiği izlemek, API kullanımını analiz etmek ve performans ölçümü işlemlerinin gerçekleştirilmesidir. Bu, hataları tespit etmek, kullanım trendlerini anlamak ve hizmet kalitesini iyileştirmek için önemlidir.

Mimari Tasarım

Tasarlanan mimari kapsamında mikro hizmetler, güvenlik katmanları ve yönlendirme stratejileri belirlenmiştir. Mevcut mimari ve tasarımı yapılan API gateway mimarisinin blok diyagramı şekil 1'de görülmektedir. APİ Gateway sistemi çeşitli sistemlere entegre edilecek olup bu sistemlerden farklı türde birçok verinin akışı ve ilgili servislerin veri analizi ve aksiyon süreçlerinin tetiklenmesi sağlanacaktır. Bu amaçla çeşitli kaynaklardan toplanan verilerin veri kaybı olmadan işlenmesi ve ilgili modüllerin tetiklenmesi amacıyla isteklerin yönlendirilebilmesi için veriler bir veri tabanında depolanarak ilgili sisteme aktarılması sağlanacaktır. Veri kaybının önlenmesi amacıyla toplanan veri önce bir kuyruklama yapısıyla sıralanarak veritabanına aktarılacak ve sonra APİ Gateway üzerinden sisteme iletilecektir.



Şekil 1: Mevcut mimari ve API Gateway Mimarisi

Yetkilendirme ve Kimlik Doğrulama katmanı, gelen istekleri doğrulamak ve yetkilendirmek için kullanılacaktır. Bu katmanın geliştirilmesi için JWT (JSON Web Token) veya OAuth gibi kimlik doğrulama mekanizmaları kullanılacaktır. Taşınabilir bir kimlik doğrulama yöntemi olan JWT, güvenliği sağlamak için dijital imzalar kullanmaktadır. OAuth ise güvenli ve standart bir yetkilendirme protokolüdür ve API'lere erişimi kontrol etmek için kullanılması planlanmaktadır.

Erişim Kontrolü katmanı, belirli kullanıcı gruplarına veya istemcilere belirli API'lere erişim kontrolü sağlayacaktır. Bu katman, API gateway üzerinde yetkilendirilmiş kullanıcıların veya uygulamaların API'lere erişimini yönetecektir. Bu amaçla yüksek performanslı bir HTTP sunucusu ve ters proxy hizmeti trafik yönlendirme ve güvenlik özellikleri için tercih edilmiştir.

Trafik Yönetimi katmanı, gelen API isteklerini belirli hedeflere yönlendirerek ve trafik yönetimi sağlanacaktır. Bu katman genellikle yüksek eşzamanlı hizmet verme yeteneği, düşük bellek kullanımı ile yüksek performans sağlayan bir Web sunucusu ve yüksek erişilebilirliğe sahip yük dengeleyici, TCP ve HTTP tabanlı uygulamalar için proxy sunucusu hizmeti veren açık kaynak kodlu bir yazılım ile geliştirilecektir. Böylece yük dengeleme, hata durumlarında geriye düşme ve trafik yönlendirme gibi yeteneklerin geliştirilmesi sağlanacaktır.

Protokol Dönüşümü katmanı, farklı protokoller arasında dönüşümleri gerçekleştirecektir. Örneğin, HTTP isteklerini HTTPS'e dönüştürme veya SOAP isteklerini RESTful API'lere dönüştürme gibi işlevleri sağlayacaktır.

İzleme ve Analiz katmanı, gelen ve giden trafiği izleyecek olup, API kullanımını analiz edecek ve sistem performansını ölçecektir. Bu katman, API trafiğini izleyecek, analiz edecek,

hataları tespit edecek, kullanım trendlerini anlamlandırarak performansı analizi gerçekleştirecektir. Jest, Mocha gibi test çerçeveleri ve Jenkins, Travis CI gibi sürekli entegrasyon ve dağıtım araçları ile desteklenecektir.

Ek olarak denetimsiz makine öğrenimi algoritmalarıyla, API trafiğini izleyerek anormal davranışlar tanımlanarak potansiyel saldırıların belirlenebileceği değerlendirilmektedir. Ayrıca, API trafiği üzerindeki trendleri analiz ederek performans iyileştirmeleri veya kapasite planlaması için önerilerde bulunulabileceği daha sonrada sisteme eklenebilecek ek yetenekler olarak değerlendirilmektedir.

API gateway yönetim arayüzünün kullanıcı dostu ve basitleştirilmiş kullanımının sağlanması için kullanıcı arayüzü tasarımları gerçekleştirilecektir. Geliştirilen kullanıcı yetkilendirme ve kimlik doğrulama işlemlerinin önyüz tarafında entegrasyon çalışmaları gerçekleştirilecektir. Kullanıcı arayüzünün güvenliği ve performansı için önlemlerin alınmasına yönelik sistem testleri gerçekleştirilecektir. Arayüz geliştirmelerinde; HTML, CSS, JavaScript kullanılması planlanmaktadır.

Uygulama ve Bulgular

API'lerin etkin yönetimi amacıyla gerçekleştirilen mimari tasarım, sürdürülebilir yazılım çözümlerinin ön planda olduğu yenilikçi bir yaklaşımla ele alınmaktadır. Tasarımın temel amaçlarında biri, enerji verimliliği ve kaynak kullanımı açısından optimizasyon sağlayarak çevreye duyarlı bir yazılım altyapısı geliştirmektir.

Mimari, yazılım sektöründe enerji tüketimini azaltmak amacıyla geliştirilen çözümlerle uyumlu bir şekilde tasarlanmıştır. Mevcut sistemlerin yüksek işlemci ve bellek kullanımı, enerji tüketimini artırarak çevresel etkileri olumsuz yönde etkilemektedir. Tasarlanan mimarinin kullanılacağı sistem halihazırda yaklaşık 10.000 müşterinin kullandığı ve günlük ortalama 37 milyon isteği işlemekte ve ciddi bir işlem *gücü* ve enerji tüketimine yol açmaktadır. Bu isteklerin %40'ının yüksek *öncelikli* olmayan istekler olduğu tespit edilmiştir. Geliştirilecek API gateway, bu önceliksiz istekleri filtreleyerek sistem kaynaklarının verimli kullanılmasını sağlayacaktır.

Çalışma, donanım ve elektrik kaynak tüketimini minimize etmek için yenilikçi teknolojiler kullanılacaktır. Bu sayede, mevcut donanım altyapısı ile müşteri sayısı artırıldığında dahi ek sunucu ihtiyacı doğmadan yüksek performanslı bir sistem sağlanacaktır. Örneğin, mevcut kullanımda 10.000 müşteriyi destekleyen sistem, proje sonrası 4000 yeni müşteri eklendiğinde bile ek sunucu gerektirmeyecektir. Bu durum, enerji tüketiminin azaltılması ve kaynakların verimli kullanılması açısından önemli bir avantaj sunmaktadır. Bu durum, tüm elektrik tüketim sarfiyatına yönelik karbon ayak izinin azaltılması kapsamında önemli bir katkı sunacaktır.

Geliştirilen sistem için, işlem *gücü kullanımını minimize eden* programlama dilleri ve teknolojiler kullanılarak enerji tüketimi minimize edilecektir. Bu sayede, çevresel farkındalık yüksek, yenilikçi ve sürdürülebilir bir teknolojik çözüm geliştirilecektir. Sistemin kullanıma geçmesiyle hedeflenen iyileşmeler tablo 1'de sunulmuştur.

Tablo 1: İyileştirme Hedefleri

Ölçüt	Hedeflenen Değer
Mevcut sistem üzerindeki performans iyileştirmesi	%40 Verimlilik Artışı
Maksimum İstek Sayısında Sistem Cevap Süresi	400 ms (Mevcut 800ms)
4000 kullanıcıya kadar ek sunucu ihtiyacı	0
Enerji tüketiminde düşüş	%10

Deneysel Çalışma Sonuçları

Bu çalışmada, mikroservis mimarisine sahip bulut tabanlı uygulamada API Gateway entegrasyonunun kaynak kullanımı üzerindeki etkisi incelenmiştir. Uygulama önce doğrudan kullanıcı isteklerini mikroservislere yönlendirirken, sonrasında API Gateway kullanılarak merkezi bir kontrol noktası oluşturulmuştur. API Gateway kullanılmadan sonuçlar Tablo 2'de verilmiştir. M*üşteri sayısındaki artışla birlikte paralel istek sayısının artması nedeniyle CPU kullanımında %75'leri aşmış, bu da ek sunucular eklenmesine neden olmuştur.*

Tablo 2: API Gateway Kullanılmadan Kaynak Gereksinimi

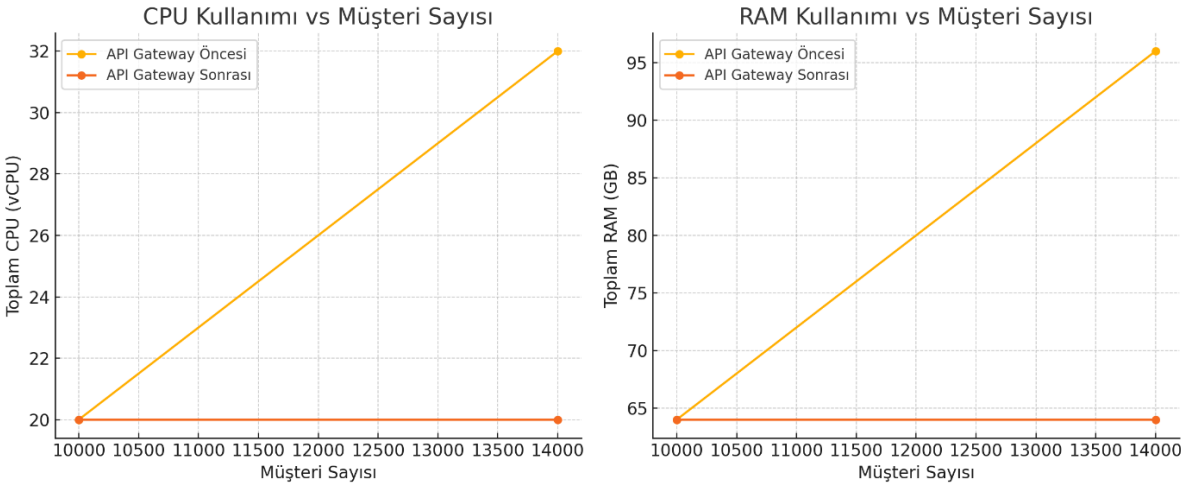
Müşteri Sayısı	Sunucu Sayısı	CPU (Toplam)	RAM (Toplam
10.000	5 Adet VM	20 vCPU	64 GB
14.000	8 Adet VM	32 vCPU	96 GB

API Gateway devreye alındığında ise Rate limiting gibi özelliklerle yük dengelemesi yapılmış ve anlık trafik dağılımı dengelenmiştir. Kaynak artışı gerekmeden 4000 müşteri daha sisteme entegre edilmiştir. API Gateway devreye alındığında elde edilen sonuçlar Tablo 2'de verilmiştir.

Tablo 3: API Gateway Kullanılmadan Kaynak Gereksinimi

Müşteri Sayısı	Sunucu Sayısı	CPU (Toplam)	RAM (Toplam
10.000	5 Adet VM	20 vCPU	64 GB
14.000	5 Adet VM	20 vCPU	64 GB

Şekil 2'de API Gateway kullanılarak ve kullanılmadığı durumlarda ortaya çıkan işlem gücü ihtiyacı ve bellek kullanımı grafikleri yer almaktadır. API Gateway kullanıldığında anlık müşteri sayısı artışında CPU kullanımı ve bellek ihtiyacının artmadığı görülmektedir.



Şekil 2: İşlem Gücü ve Bellek Kullanımı Grafikleri

Sonuç ve Gelecek Çalışmalar

Bu çalışma kapsamında API'lerin etkin yönetimi amacıyla bir mimari tasarım gerçekleştirilmiştir. Geliştirilen tasarım ile yapılan ön çalışma ve deneylerde umut vadeden sonuçlar elde edilmiştir. Sürdürülebilir yazılım çözümlerinin geliştirilmesi için yenilikçi bir yaklaşımla ele alınan tasarımın enerji verimliliği ve kaynak kullanımı açısından optimizasyon sağladığı görülmüştür. Yapılan deneylerde elde edilen veriler makine öğrenmesi algoritmaları ile değerlendirilerek kaynak planlamasının daha optimize gerçekleştirilmesi çalışmaları başlamıştır.

Gelecek dönem çalışmalarında daha fazla sayıda kullanıcı talebinin, farklı API'lar ile eş zamanlı olarak geldiği durumlardaki kaynak yönetimi çalışmaları devam edecektir. Optimizasyon çalışmaları için ve makine öğrenmesi algoritmalarının kullanılabilmesi için veri toplama ve eğitim süreçleri gerçekleştirilecektir.

Çıkar Beyanı**:** Yazarlar çıkar çatışması olmadığını beyan eder.

Teşekkürler: Yazarlar, Medyasoft Bilişim Grubu'na çalışmanın gerçekleştirilmesi için sağladığı imkanlardan dolayı teşekkür eder.

Yazar Katkıları: Muhammet Emin genel sektör ihtiyacına istinaden oluşan ARGE fikrinin Ürün Sahipliğini yapmaktadır. Furkan Baran ve Muhammet Emin mimari tasarım ve önyüz tasarımını, Özlem Emin, Ali Doğan, Mustafa Kozan arka yüz geliştirimlerini, Gökçen Gökyer Nalçı yazılım proje yönetimini, Önder Demir ArGe ve akademik danışmanlığı yapmıştır. Çalışmanın ArGe niteliğine ve test süreçlerine tüm ekip katkı sağlamıştır.

REFERANSLAR

[1] A. Lercher, L. Ferreira, C. Sulzmann, and U. Zdun, "Microservice API evolution in practice: A study on strategies and challenges," J. Syst. Softw., vol. 215, p. 112110, 2024.

[2] L. Ochoa, J. Garcia, G. Robles, and J. M. Gonzalez-Barahona, "Characterising harmful API uses and repair techniques: Insights from a systematic review," Comput. Sci. Rev., vol. 57, p. 100732, 2025.

[3] I. Rauf, E. Troubitsyna, and I. Porres, "A systematic mapping study of API usability evaluation methods," Comput. Sci. Rev., vol. 33, pp. 49–68, 2019.

[4] U. Aharon, Y. Kim, J. Keshet, Y. Yeshurun, and M. Lindenbaum, "A classification-by-retrieval framework for few-shot anomaly detection to detect API injection," Comput. Secur., vol. 150, p. 104249, 2025.

[5] D. Arcolini, Full Lifecycle API Management: Microgateway Infrastructural Pattern Adopting Kong Gateway, M.S. thesis, Politecnico di Torino, 2023.

[6] N. Podgorny, A. Dvoryankin, A. Konev, and A. Pogrebnoy, "End user authorization solutions in modern microservice applications," in Proc. AIP Conf., vol. 2969, no. 1, 2024.

[7] M. Tomić, G. Popović, D. Milinković, and B. Crnobrnja, "Towards applying API gateway to support microservice architectures for embedded systems," in Proc. Int. Conf. Embedded Syst. Appl., 2022, pp. 86–91.

[8] M. Söylemez, B. Tekinerdogan, and A. K. Tarhan, "Challenges and solution directions of microservice architectures: A systematic literature review," Appl. Sci., vol. 12, no. 11, p. 5507, May 2022. [Online]. Available: https://doi.org/10.3390/app12115507

[9] S. Hassan, R. Bahsoon, and R. Kazman, "Microservice transition and its granularity problem: A systematic mapping study," presented at the IEEE Int. Conf. Softw. Archit. (ICSA), Hamburg, Germany, Mar. 2019. [Online]. Available: https://arxiv.org/abs/1903.11665

[10] Z. Mousavi, C. Islam, M. A. Babar, A. Abuadbba, and K. Moore, "Detecting misuse of security APIs: A systematic review," arXiv preprint, Jun. 2023. [Online]. Available: https://arxiv.org/abs/2306.08869

[11] S. Emanet, G. K. Baydogmus, and O. Demir, "An ensemble learning based IDS using Voting rule: VEL-IDS," PeerJ Comput. Sci., vol. 9, p. e1553, 2023.

[12] M. Mathijssen, M. Overeem, and S. Jansen, "Identification of practices and capabilities in API management: A systematic literature review," arXiv preprint, Jun. 2020. [Online]. Available: https://arxiv.org/abs/2006.10481

[13] D. E. Kurniawan and A. Dzikri, "Development of a prototype for microservices architecture and API gateway integration in an eLearning platform," presented at the 6th Int. Conf. Appl. Eng. (ICAE), Batam, Indonesia, Nov. 2023. [Online]. Available: https://eudl.eu/doi/10.4108/eai.7-11-2023.2342930S.J. Randles, "Formulation of Environmentally Acceptable Lubricants", 49th STLE Annual Meeting, May 1-5, 1994, Pittsburgh.

[14] J.W. Strunk, and E.B. White, "The Elements of Style", 4th ed. New York: Longman, 2020.

FORWARDER TRUCK SYSTEM

Eyüp Kerem Malkoç*
Maltepe Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Yazılım Mühendisliği, İstanbul, Türkiye

Emre Atlıer Olca
Maltepe Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Yazılım Mühendisliği, İstanbul, Türkiye
(Corresponding author) malkockerem1312@gmail.com

ABSTRACT

The Forwarder project is a software application developed for freight companies providing transportation services in international logistics. This project aims to simplify the process of finding cargo for trucks. Truck owners will manually enter route, stop, and available space information into the system but will not interact with the system directly. The project manages transportation operations through a forwarder (freight transport intermediary) to ensure more efficient and faster execution of transport processes. Communication will be conducted via email.

Developed with a monolithic architecture, the project is being implemented using Spring Boot and React technologies. The database containing cargo and trucks will be managed through a user-friendly interface. The project aims to offer a significant solution for optimizing transportation processes in the logistics sector.

Ultimately, all necessary infrastructure will be provided to ensure the smooth operation of the system and the delivery of fast, reliable service to users. This software will contribute to increasing operational efficiency in the logistics sector and accelerating digital transformation in the industry.

Keywords: Logistics, transportation, carg, truck, digitalization, freight matching; route.

Introduction

Digitalization in the international logistics sector is a critical transformation process that provides efficiency and cost savings. The Forwarder project is a software solution developed for logistics companies and transportation service providers, aiming to optimize business processes in the sector as part of this transformation. Through this project, the goal is to achieve time and cost savings by digitalizing logistics processes. Digitalization not only increases operational efficiency but also improves the transportation service process by providing transparent operations [1].

Research Objective

The main objective of this study is to examine the benefits that the Forwarder project will bring to the logistics sector and evaluate the impact of digital transformation on the industry. The project aims to reduce the workload of transportation service providers, accelerate transportation decisions, and make logistics operations more transparent. In this context, the necessity of a user-friendly software interface and a strong infrastructure is emphasized. As highlighted in the literature, modern software architectures play a significant role in improving operational processes in the logistics sector [2].

Omparison with Similar Systems and Differences

The Forwarder project targets significant gaps compared to existing logistics software in the industry. In current software, complex and user-unfriendly interfaces negatively impact the operational efficiency of logistics companies. Although similar platforms are found in the literature, most platforms use limited data analytics and prioritize user experience in the background [3]. Additionally, shortcomings in data security increase the risk of cyber-attacks and data breaches. Specifically, it is noted that the number of data breaches in the logistics sector increased by 97% compared to the previous year [4]. This situation can damage customer trust and lead to financial losses. Forwarder aims to provide flexible and secure responses to the needs of logistics companies with its user-friendly interface and strong security measures. In this regard, it aims to ensure data security through strategies such as data encryption, firewalls, and access control [5].

Potential of the Forwarder Project

The Forwarder project stands out as a critical step supporting digital transformation in the logistics sector. The integration of digital technologies in logistics offers significant benefits, such as increasing efficiency, making operations more transparent, and reducing costs [6]. Once completed, the project will enable the transportation services process to be managed more quickly, securely, and efficiently, thus introducing a new operational model to the sector. With its modern software architecture and technological infrastructure, Forwarder aims to provide a lasting solution not only for transportation service providers but for the entire logistics ecosystem.

Detailed Project Description

The Forwarder project is a comprehensive software solution developed to address one of the largest challenges in the logistics sector: inefficiencies in transportation processes, underutilization of transport capacity, and shipments being sent out incomplete or empty. The project aims to minimize transportation costs by using advanced techniques such as route optimization for the optimization of transportation processes.

The project will enable the digitalization of all logistics processes, allowing real-time tracking and management of transportation data. Users will be able to monitor the current status of transportation vehicles, their routes, load capacities, and transit times instantaneously. This will prevent vehicles from departing empty or with insufficient loads, enabling logistics service providers to plan transportation operations more efficiently and reduce costs.

Forwarder will also allow for transparent pricing of transportation services, ensuring that users can obtain accurate price information and make transportation decisions in a more reliable and predictable manner. As a result, logistics service providers will be able to calculate transportation prices more accurately and offer their services more reliably. Transparent pricing practices will address the trust deficit in the sector and provide users with more information about prices.

Another important aspect of the project is its goal to address the challenges in the transportation sector and prevent inefficient transportation operations by offering a platform that consolidates all vehicle-related data. Users will easily access information such as vehicle status, load capacity, and routes from this platform. This will enable transportation companies to manage load transportation processes more quickly and accurately. Additionally, this platform will minimize human errors in transportation processes, streamlining workflows and making logistics service providers' operations faster and more efficient.

Importance of Digitalization in Logistics

Digitalization plays a critical role in optimizing business processes and enhancing logistics performance in the transportation sector [7]. Route optimization not only reduces transportation costs but also increases energy efficiency, thereby reducing environmental impacts [8]. These factors are essential components that will enable the Forwarder project to achieve its sustainability goals.

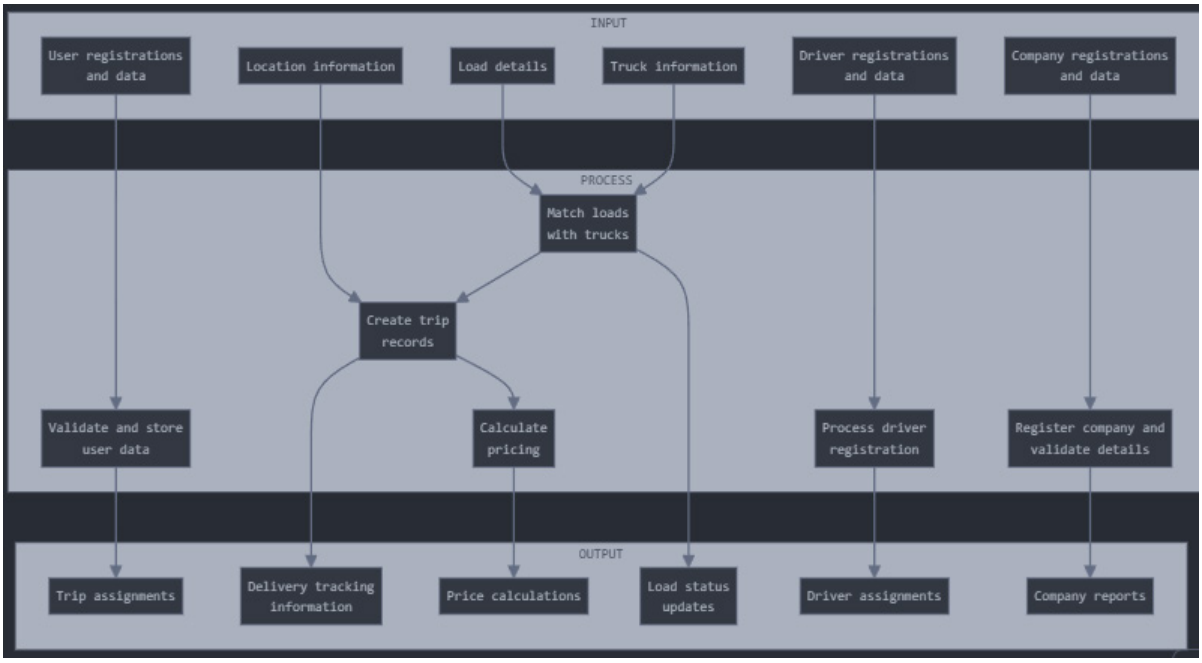


Figure 1: Forwarder Truck System Input-Output

Material and Method

This section will discuss in detail the software technologies, architectural structures, and development methodology used in the Forwarder project. The technologies selected are carefully chosen to enhance the system's performance, ensure its security, and strengthen its sustainability.

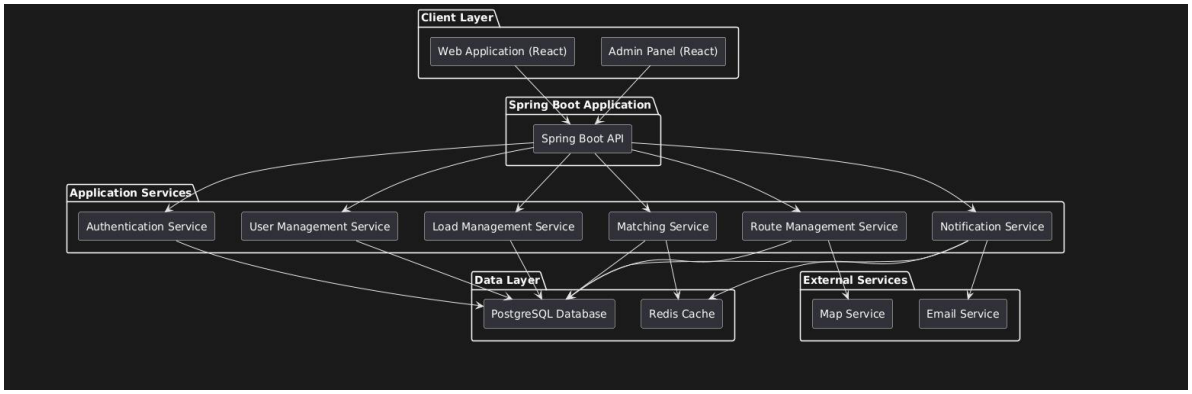


Figure 2: Forwarder Truck System Mimari

Software Infrastructure and Technologies

The architecture of the Forwarder project is carefully designed to manage and optimize transportation processes in the logistics sector in a digital environment. The project continues to be developed on a monolithic architecture, with all components working in an integrated manner within a single system. This architecture simplifies management, maintenance, and development processes, increasing the project's sustainability and enabling quick adaptation.

Backend Architecture

The backend architecture is created using Spring Boot 3 technology and is continuously optimized to ensure high performance and security in logistics processes. The monolithic architecture was chosen to manage the system through a single codebase, thus gaining speed and cost advantages during the development process. The monolithic structure ensures seamless integration between components and simplifies management processes.

Thanks to the extensive plugin support offered by Spring Boot, processes such as security, error management, and data processing have been made more efficient. API services effectively handle user requests with fast response times, and as the system is scalable, it can easily adapt to increasing data traffic.

There are three primary user roles in the system:

- **Forwarder Agent:** The main user who manages operations, coordinates between load owners and truck owners, and oversees the entire process.

- **Truck Owner:** Does not have direct access to the system and only receives information through emails sent by the Forwarder Agent.

- **Shipper:** The party that initiates load requests and communicates with the Forwarder Agent to organize the transportation process.

The backend manages all business logic in the system, connecting to the database and ensuring that data is processed, stored, and analyzed correctly. Additionally, a centralized monitoring mechanism is being integrated for log management and error debugging processes.

Frontend Architecture

The frontend is developed using ReactJS, providing a dynamic and user-friendly interface. An intuitive and modern design has been adopted to enable users to manage logistics processes in the most efficient way. Thanks to the modular structure provided by ReactJS, components have been made reusable, saving time in the development process.

User experience is prioritized in the interface design, and development continues with the consideration of non-straining color palettes, organized page structures, and accessibility standards. Predictable navigation and interactive graphics are used to allow users to manage transportation processes with ease.

Additionally, optimization techniques such as server-side rendering (SSR) and lazy loading are applied on the frontend to enhance performance. As a result, page load times have been improved, and the system now offers a smoother user experience.

Data Storage Architecture

The database layer is managed using PostgreSQL, which is a high-performance and reliable relational database management system that ensures secure storage and access to transportation and logistics data. This database centrally stores all transportation transactions, vehicles, loads, and user data in the system. Additionally, PostgreSQL provides a robust infrastructure for data analysis and reporting with its large data processing capabilities.

Data Storage for User Roles:

The system provides detailed data management based on user roles. All actions performed by users in the system, along with access levels and authorization information, are stored in the PostgreSQL database. Separate authentication processes and access control mechanisms are applied for Forwarder Agent, Truck Owner, and Shipper roles.

- **Forwarder Agent Data:** Login credentials, authorization levels, transaction history, and customer relations are stored in detail.

- **Truck Owner Data:** Vehicle information, transportation history, and financial records related to operations are securely stored.

- **Shipper Data:** Load requests, transportation process details, and payment information are stored.

These data are encrypted according to security protocols and can only be accessed by authorized personnel.

System Algorithms and Route Optimization

To efficiently manage logistics processes, the Forwarder project utilizes intelligent system algorithms such as route optimization. These algorithms aim to reduce unnecessary mileage, lower costs, and shorten transportation time. The system enhances operational efficiency by offering dynamic pricing and load-matching mechanisms, providing logistics companies with the most suitable transportation options.

Development Methodology

The Forwarder project is being developed based on a fundamental Sprint structure. The Agile methodology is adopted, and the system is updated with regular iterations, ensuring continuous improvements.

Sprint Structure

The following steps are followed during the project development process:

- **Planning:** At the beginning of each sprint, the targeted features are determined, and the workload is distributed.
- **Development:** The identified tasks are distributed among the backend, frontend, and database components for coding.
- **Testing:** Developed components are subjected to integration and unit tests to ensure the system's stability.
- **Evaluation and Feedback:** At the end of each sprint, internal team evaluations are conducted to fix any issues and plan improvements for the next sprint.

With this methodology, the Forwarder project maintains a sustainable development process and can quickly adapt to customer feedback.

Results and Recommendations

Efficient management of transportation capacity is one of the critical elements that increase operational efficiency in the logistics sector. This project aims to minimize empty trips and optimize resource usage by providing a quick and reliable match between carriers and cargo owners.

As a result of the project, a system has been developed that allows carriers to easily find loads that match their routes, while cargo owners can access reliable transportation solutions. With its user-friendly interface and real-time tracking features, the platform digitizes logistics processes, adding innovation to the sector.

Benefits Provided by the Platform:

- Reduction of unused capacity and cost reduction,
- Access for cargo owners to a wide network of carriers,
- Increased transparency and trust in transportation processes,
- Strengthened competitive advantage.

Future Development Recommendations:

- **Mobile compatibility:** Developing a mobile application to ensure users can access the platform anytime.
- **Data analysis:** Adding advanced analytics systems to improve matching algorithms.

• **International expansion:** Integrating the platform into global transportation processes.

• **Feedback mechanism:** Strengthening the feedback system to enhance user satisfaction.

• **Sustainable solutions:** Implementing measures to reduce the environmental impact of logistics operations.

Conclusion

This project aims to accelerate digital transformation in the logistics sector by optimizing transportation processes, increasing operational efficiency, and improving user experience. Today, the logistics sector faces challenges such as time losses and cost increases due to manual processes while needing to adapt to rapidly advancing technology. In this context, the developed platform, with smart algorithms and data analytics support, makes transportation processes more efficient, reduces costs, and increases the reliability of logistics operations.

The platform adds value to all stakeholders in the logistics ecosystem by offering solutions from supply chain management to delivery tracking. Real-time data flow and automated process management help minimize operational errors, while logistics companies can utilize resources more efficiently. Additionally, the AI-supported analytics provided by the system enable users to make more informed decisions and assist in pre-planning operations with predictive analysis.

With future developments, the platform's scope will be expanded to offer innovative solutions in different areas of the logistics sector. By focusing on green logistics practices, significant steps can be taken in environmental sustainability and carbon footprint reduction. Moreover, integration processes in international transportation can be improved, ensuring more effective operation of global logistics networks.

The steps taken toward modernization in the logistics sector will increase the competitive power of businesses and encourage the widespread digital transformation of the sector. Digitalized and automated logistics processes will offer significant benefits to sector players, with advantages in speed, security, and cost, while also contributing to the spread of sustainable logistics solutions.

Digitalization in International Logistics: Forwarder Project

REFERENCES

[1] M. Johnson, "Digital Transformation in Logistics: Efficiency and Cost Reduction," *Journal of Logistics Innovation*, vol. 15, no. 4, pp. 25-33, 2023.

[2] A. S. White, "The Impact of User Interface Design in Logistics Software," *International Journal of Software Engineering*, vol. 21, no. 2, pp. 112-119, 2022.

[3] MDPI, "Artificial Intelligence in Logistics Optimization with Sustainable Criteria," *Sustainability*, vol. 16, no. 21, pp. 9145-9162, Dec. 2024.

[4] Lojiyol, "Casus Yazılımlar ve Veri Güvenliği: Lojistik Sektörünü De Tehdit Ediyor," [Online]. Available: https://www.lojiyol.com/casus-yazilimlar-ve-veri-guvenligi-lojistik-sektorunu-de-tehdit-ediyor. [Accessed: 28-Feb-2025].

[5] DBT, "Lojistik Yazılımı: Taşımacılıkta Teknolojik Dönüşüm," [Online]. Available: https://blog.dbt.com.tr/2024/03/01/lojistik-yazilimi-tasimacilikta-teknolojik-donusum. [Accessed: 28-Feb-2025].

[6] D. Thomas, "The Role of Digitalization in Enhancing Transparency and Efficiency in Logistics," *Logistics Technology Review*, vol. 18, no. 5, pp. 45-50, 2024.

[7] Pusat Lojistik, "Lojistikte Rota Optimizasyonu," [Online]. Available: https://pusatlojistik.com/lojistikte-rota-optimizasyonu/. [Accessed: 28-Feb-2025].

[8] Pusat Lojistik, "Lojistikte Maliyet Optimizasyonu," [Online]. Available: https://pusatlojistik.com/lojistikte-maliyet-optimizasyonu/. [Accessed: 28-Feb-2025].

[9] "Spring Boot Documentation," Spring, [Online]. Available: https://docs.spring.io/spring-boot/docs/current/reference/htmlsingle/. [Accessed: 28-Feb-2025].

[10] "React Documentation," ReactJS, [Online]. Available: https://reactjs.org/docs/getting-started.html. [Accessed: 28-Feb-2025]

[11] "PostgreSQL Documentation," PostgreSQL, [Online]. Available: https://www.postgresql.org/docs/. [Accessed: 28-Feb-2025].

[12] "Redis Documentation," Redis, [Online]. Available: https://redis.io/documentation. [Accessed: 28-Feb-2025].

[13] "Spring Boot Email Integration," Baeldung, [Online]. Available: https://www.baeldung.com/spring-boot/email. [Accessed: 28-Feb-2025].

[14] "Server-Side Rendering in React," [Online]. Available: https://www.bairesdev.com/blog/server-side-rendering-react/?__cf_chl_tk=EmZbeKmq5eX8Ki2MJ5vsOBb2kaTawuDtwkaAtJVb-JS8-1741522872-1.0.1.1-sqkoye6..MEyaxAVO0G62Fvf9kYhDApLKp.g5LdyEso

GESTATIONAL DIABETES PREDICTION WITH ANFIS AND MACHINE LEARNING METHODS

Sevcan Emek*
Manisa Celal Bayar University, Department of Computer Engineering, Manisa, Türkiye

Mehmet Emre Yılmaz
Amasya University, Educational Sciences, Amasya, Türkiye
(Corresponding author) sevcan.emek@cbu.edu.tr

ABSTRACT

Gestational Diabetes Mellitus (GDM) poses a significant challenge in maternal health, requiring precise diagnostic tools for early detection and intervention. This study explores the potential of machine learning algorithms and hybrid intelligent systems, notably the Adaptive Neuro-Fuzzy Inference System (ANFIS), alongside traditional methods such as Logistic Regression, Random Forest, and XGBoost, for GDM prediction. Utilizing a diverse dataset, key clinical and lifestyle features including BMI, HDL levels, age, PCOS, sedentary lifestyle, prediabetes, and family diabetes history were analyzed to optimize predictive accuracy. Experimental results highlighted XGBoost's accuracy of 95.10% while ANFIS demonstrated a robust 90.24% accuracy rate. Findings underscore the value of machine learning approaches and ANFIS models in enhancing early diagnosis and personalized treatment plans for GDM, offering promising pathways for healthcare advancements.

Keywords: Gestational diabetes, ANFIS, prediction, machine learning.

Introduction

Diabetes mellitus, commonly known as diabetes, is a chronic disease caused by the pancreas's inability to produce sufficient insulin or the body's inability to effectively use insulin, resulting in elevated blood sugar levels [1]. The disruption of the body's mechanism to convert food into energy leads to uncontrolled increases in blood sugar levels. The body's energy needs are primarily met through proteins, fats, and carbohydrates. Glucose, the most essential component of carbohydrates, is the primary energy source for the body and is vital for critical organs like the brain. For glucose to be utilized by cells as energy, the hormone insulin, secreted by the pancreas, is required. However, in cases of insulin deficiency or inefficiency, glucose cannot be converted into energy, leading to dangerously high blood sugar levels [2-5]. The increasing prevalence of diabetes is largely attributed to changing living conditions, sedentary lifestyles, and unhealthy eating habits. Diabetes imposes a significant burden not only on individual health but also on global healthcare systems. According to the World Health Organization (WHO), diabetes is one of the leading causes of death worldwide, and its prevalence has been rapidly increasing in recent years. As of 2024, there are 830 million people with diabetes globally. This number, which

was 326 million in 2010, is projected to exceed 1.3 billion by 2050. No country is expected to experience a decline in diabetes rates in the next 30 years [6].

GDM (Gestational Diabetes Mellitus) is glucose intolerance detected during pregnancy [7]. With the increase in obesity and diabetes, the number of pregnant women with undiagnosed Type-2 diabetes is also increasing. GDM is estimated to occur in an average of 10% of pregnancies. This rate varies between 1% and 22% depending on the population studied and the diagnostic methods used [8, 9]. Diabetes poses a number of risks for both maternal and fetal health in later life, such as metabolic dysfunction and obesity. For this reason, early diagnosis and treatment of diabetes is important. Computer-aided artificial intelligence models are focused on diagnosing and predicting some diseases or risk factors such as diabetes.

Several studies in the literature have demonstrated the effectiveness of machine learning methods in predicting diabetes. Zhang et al. evaluate the effectiveness of machine learning (ML) models in predicting GDM. An analysis of 25 studies reveale that the average area under the receiver operating characteristic curve (AUROC) for ML models is 0.8492. Logistic regression model demonstrates satisfactory performance with an AUROC of 0.8151, whereas advanced ML methods such as LightGBM and GA-CatBoost achieve higher accuracy, with an AUROC of 0.8891. The most commonly utilized predictive factors include maternal age, body mass index (BMI), fasting blood glucose levels, and a family history of diabetes [10]. Savvidou et al. develop model that predict the risk of GDM during the first trimester by integrating maternal characteristics and laboratory measures. The study analyzes data from 124 GDM cases and 248 controls. The baseline model includes parameters such as maternal age, BMI, ethnicity, parity, prior GDM, and a family history of diabetes [11]. Cubillos et al. develop to predict the risk of GDM during early pregnancy. The research was based on data from 1611 pregnancies, with a total of 12 different models optimized. Approaches such as MLP, SVM, and LR are employed, and efficiency is enhanced through variable selection methods. Notably, an MLP model utilizing only 5 variables achieves sensitivity of 89%, specificity of 62%, and an AUC-ROC value of 0.82. More complex models employing 12 variables attain sensitivity of 82% specificity of 74% and an AUC-ROC value of 0.81. Data augmentation techniques (DA) are applied to improve the performance of the models [12]. Kang et al. evaluate machine learning algorithms to predict the risk of GDM using data from 34387 pregnancies in South Korea. The algorithms compared in the study are Light Gradient Boosting Machine (LGBM) and Extreme Gradient Boosting (XGBoost). Key parameters influencing the performance of the algorithms include maternal age, body mass index (BMI), previous GDM history, endocrine disease history, and family history of diabetes [13]. Kurt and her team improve the diagnosis of gestational diabetes (GD) through the development of an innovative clinical decision support system using advanced machine learning and deep learning algorithms. Supported by deep learning methodology, the RNN-LSTM model coupled with Bayesian optimization achieve remarkable performance metrics, including 95% sensitivity, 99% specificity, and an AUC of 98%. For performance comparison, traditional machine learning algorithms like Support Vector Machines (SVM) and Random Forest are also evaluated [14]. Watanabe et al. investigate the prediction of GDM using machine learning algorithms applied to data from the Japan Environment and Children's Study (JECS) collected during early pregnancy from

82,698 pregnant women. GDM parameters include pre-pregnancy BMI, maternal age, HbA1c levels, and maternal birthweight. In this work, they implements Gradient Boosting Decision Tree (GBDT) algorithm, Logistic Regression (LR), Random Forest (RF), and Support Vector Machine (SVM) [15]. Artzi et al. develop a model for early diagnosis of gestational diabetes mellitus (GDM) based on electronic health records (EHR) obtained from Israel's largest healthcare provider. The study includes 588622 pregnant who gave birth between 2010 and 2018. The diagnosis of GDM is based on glucose challenge tests (GCT) and oral glucose tolerance tests (OGTT) performed during the 24-28 weeks of gestation [16]. Shanmugam and his team evaluate advanced machine learning techniques including logistic regression (LR), Random Forest (RF), Extreme Gradient Boosting (XGBoost), and LightGBM. This work explores various algorithms used in GDM diagnosis to predict the adverse effects on maternal and fetal health [17]. Zaky et al. evaluate clinical measurements and machine learning (ML) techniques for the early- stage identification of GDM in Qatar. This study utilizes a dataset comprising 68 features collected during the first trimester and assesses the performance of 11 different algorithms, along with a combined "StackEnsemble" model. According to the results, the stacking method demonstrates the best performance, achieving 92.1% sensitivity and 88.83% accuracy. Gradient Boosting (80.11% accuracy) and AdaBoost (81.94% accuracy) also show strong predictive capabilities. Random Forest exhibits effective performance with an accuracy of 80.22%. Among more traditional models, Logistic Regression offers a reasonable performance with 79.22% accuracy, followed closely by the Support Vector Classifier (SVC) with 78.06%. CatBoost stands out with an accuracy of 81.56%, excelling in both prediction accuracy and overall model performance. In contrast, the Gaussian Naive Bayes (GaussianNB) model demonstrates limited performance with 77.28% accuracy. The K-Nearest Neighbors (KNN) algorithm achieves an accuracy of 78.39%, playing a notable role in capturing localized patterns [18]. Sumathi and Meganathan conduct a study on the prediction of GDM using a real-time GDM dataset. The dataset comprises 3525 instances and 15 features, with 2153 samples labeled as class 0 and 1372 as class 1. The dataset includes clinical attributes such as age, body mass index (BMI), family history, polycystic ovary syndrome (PCOS), and oral glucose tolerance test (OGTT) results. This study implements various machine learning models, including Logistic Regression (LR), k-Nearest Neighbors (KNN), Support Vector Machines (SVM), and Random Forest (RF). According to performance evaluation results, the ensemble model outperforms the individual models, achieving 94% accuracy, 94% precision, 94% recall, and an F1-score of 94%. RF ranks second with an accuracy of 92.39% and an F1-score of 92%. LR demonstrates moderate performance with 91.60% accuracy and an F1-score of 91%. In contrast, the KNN and SVM models yield the lowest performance, with accuracy rates of 84.96% and 82.49%, and F1-scores of 84% and 82%, respectively [19].

This work presents to assess the potential of various machine learning algorithms including Adaptive Neuro-Fuzzy Inference System (ANFIS), Random Forest, Support Vector Machine (SVM), Extreme Gradient Boosting (XGBoost), Decision Tree, and Logistic Regression for the diagnosis of diabetes. Presented model uses key clinical and lifestyle parameters such as age, high-density lipoprotein (HDL), presence of polycystic ovary syndrome (PCOS), body mass index (BMI), sedentary lifestyle, prediabetes status, and family history of diabetes. These features are selected due to their strong association with the devel-

opment of diabetes. The results indicate that machine learning approaches and hybrid intelligent systems such as ANFIS demonstrate strong potential as tools to assist healthcare professionals in early diagnosis and personalized treatment planning.

This study is organized as follows: Section 2 offers detailed information about methodology and implementation of this study. Section 3 presents experimental study and outcomes of the models. The conclusions of this study are discussed in Section 4. Finally, Section 5 presents discussion and future works.

Methodology and Implementation

In this work, GDM estimation results are observed by implementing machine learning methods with an ANFIS model. ANFIS is a hybrid method that combines the learning capability of neural networks with the flexibility and interpretability of fuzzy logic. These features make it an extremely effective tool for overcoming challenges such as uncertainty and variability, which are frequently encountered in medical datasets. In addition to the ANFIS model, various machine learning techniques are employed to enhance the model's performance and to provide a comparative analysis [20]. Logistic Regression, Decision Tree, Random Forest, Support Vector Machine, and XGBoost are implemented to observe model capacity. The preparation and analysis of the data are performed using Python and Pandas library provided powerful tools for data manipulation and analysis, significantly accelerating the preprocessing steps. Key data processing tasks included:

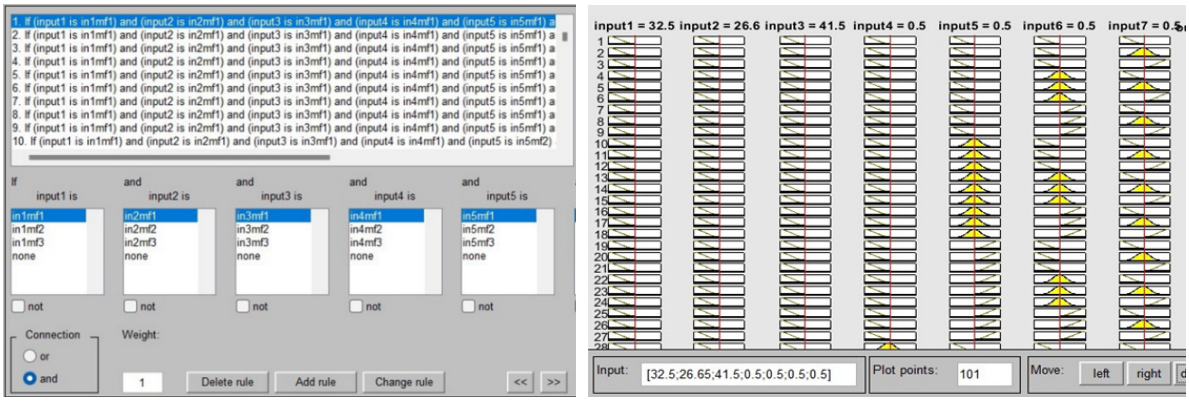
- Detection and imputation of missing values.
- Data normalization to make the dataset suitable for model training.

The design process for the fuzzy model utilizes the fuzzy logic toolbox in MATLAB based on the FIS Sugeno-type, specifically using ANFIS. Below are the steps that have been performed throughout this study:

- Step 1: Identify the inputs, output, linguistic variables, and their membership functions.
- Step 2: Fuzzification: Transform the inputs into three linguistic variables (low, normal, high) within a specific fuzzy set value, as shown in Table 1.
- Step 3: Fuzzy rules base: Construct IF-THEN rules to predict diabetes status shown in rule editor, Figure 1.
- Step 4: Inference engine: Prediction of diabetic status shown in Figure 2.

Table 1: Linguistic variables of input and output

Function	Parameter	Linguistic Variable
	PCOS	Low, Normal, High
	Family History	Low, Normal, High
	BMI	Low, Normal, High
Input	HDL	Low, Normal, High
	Age	Low, Normal, High
	Sedentary Life	Low, Normal, High
	Prediabetes	Low, Normal, High
Output	Diabetes Status	Healthy, Diabetic



(a) (b)

Figure 1: a) Rule editor, b) Rule viewer

The implementation involves a structured pipeline, including data preprocessing, resampling, model training, evaluation, and visualization. The steps followed are detailed below:

•Step 1: Data Preprocessing: The dataset is imported from an Excel file and cleaned by removing irrelevant columns. The target variable is defined as GDM or non GDM.

•Step 2: Handling Class Imbalance with SMOTE: To address the issue of class imbalance between GDM and non-GDM samples, the SMOTE (Synthetic Minority Over-sampling Technique) algorithm is applied.

•Step 3: Data Splitting: The resampled dataset is split into training and testing sets with an 80/20 ratio using stratified sampling to preserve class distribution.

•Step 4: Model Training: Five different supervised machine learning algorithms are trained on the prepared dataset: Logistic Regression, Decision Tree Classifier, Random Forest Classifier, Support Vector Machine (SVM) with RBF kernel, and XGBoost Classifier. Each model is trained using the scaled training set and evaluated on the test set.

•Step 5: Model Evaluation: The performance of each model is evaluated using accuracy, classification report metrics (precision, recall, F1-score), and error metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Normalized RMSE (NRMSE).

•Step 6: Performance Comparison and Visualization: The model with the highest accuracy score is selected as the best-performing model. A bar plot was generated to visually compare the performance of all models in terms of accuracy.

Experimental Study and Results

The dataset used for this work is obtained from Kaggle and initially consisted of 3525 records. During the preprocessing phase, the following tasks are performed:

•Removal of irrelevant columns that did not contribute to model performance.

•Elimination of outliers to ensure the quality of input values.

ANFIS is implemented using MATLAB tool. The configuration of the model is carried out as follows:

•Input Membership Functions: Each of the seven input variables is assigned Gaussian membership functions (gaussmf).

•Rules: The system comprises 2187 fuzzy rules, derives from the 3^7 combinations of input membership functions.

•Training Algorithm: A hybrid optimization method is utilized to train the fuzzy inference system, ensuring efficient convergence and improved accuracy in optimizing membership functions and fuzzy rules.

•Fuzzy Inference System: Sugeno-type inference is employed to manage the output.

In addition to the ANFIS model, several traditional machine learning models, including Logistic Regression, Decision Tree Classifier, Random Forest Classifier, Support Vector Machine (SVM), and XGBoost Classifier, are implemented and evaluated using Python's scikit-learn and XGBoost libraries.

Each model is trained on 80% of the dataset and evaluated on the remaining 20%. Model performance is assessed based on metrics such as accuracy, precision, recall, and error measurements. These machine learning models provide a comparative with ANFIS system, enabling a comprehensive performance evaluation across different algorithmic paradigms.

ANFIS model is trained on a dataset consisting of 1443 samples. The model achieves an accuracy rate of 90.24%, demonstrating its ability to effectively learn and generalize from the input data using fuzzy rules and Gaussian membership functions. This performance indicates that the system is well-optimized in capturing complex patterns in the data through its 2187 fuzzy rules.

Five machine learning models are trained and evaluated for comparative analysis. Their performances, based on accuracy, are shown in Table 2.

Table 2: Accuracy of Machine Learning Models

Model	Accuracy (%)
Logistic Regression	93.63
Decision Tree	93.63
Random Forest	94.61
SVM	93.87
XGBoost	95.10

Error measurements of models are shown in Table 3.

Table 3: Error Metrics for Machine Learning Models

Model	MSE	RMSE	MAE	NRMSE
Logistic Regression	0.0637	0.2524	0.0637	0.2524
Decision Tree	0.0637	0.2524	0.0637	0.2524
Random Forest	0.0539	0.2322	0.0539	0.2322
SVM	0.0613	0.2475	0.0613	0.2475
XGBoost	0.0490	0.2214	0.0490	–

The following results are obtained for this system evaluated with ANFIS and machine learning methods.

• XGBoost achieves the highest accuracy but also has the lowest error rates across most metrics, making it the best-performing model in this study.

• ANFIS model demonstrates strong learning capabilities, with over 90% accuracy using fuzzy inference and Gaussian membership functions.

• Traditional machine learning models, especially XGBoost and Random Forest, outperform ANFIS in terms of classification accuracy and error minimization.

• Despite the relatively small dataset size, all models yield competitive results, which can be further enhanced with additional data and fine-tuning.

• Effective feature selection and data preprocessing (including outlier removal and normalization) significantly contribute to improved performance across all models.

Conclusion

This study demonstrates the applicability of the ANFIS model in the field of diabetes diagnosis, achieving promising results in improving diagnostic accuracy using specific input parameters such as HDL, BMI, age, PCOS, sedentary life, prediabetes and family history of diabetes. A dataset comprising 3525 records is preprocessed and analyzed, and the ANFIS model is implemented using MATLAB's fuzzy logic toolbox. The system is structured with a Sugeno-type fuzzy inference system to evaluate the input parameters and generate meaningful results. This approach enables the model to produce more effective and consistent predictions based on the provided input data. The model is evaluated on training and testing datasets, achieving an overall accuracy rate of 90.24%. However, the machine learning models, especially XGBoost, perform better with an accuracy of 95.10%. This highlights the importance of exploring various modeling techniques for improving prediction accuracy. Despite this, the ANFIS model retains an advantage in terms of interpretability and the ability to handle uncertainty and imprecision in the data.

This study indicates that additional factors, such as lifestyle habits or more comprehensive datasets, should also be considered. This study provides a foundation for developing more precise and personalized approaches to diabetes diagnosis. Future studies should focus on enhancing model performance through the integration of new variables and larger, more diverse datasets.

Discussion and Future Work

ANFIS model, while achieving an accuracy of 90.24%, demonstrated solid performance for diagnosing gestational diabetes. However, certain limitations suggest areas for future research and model improvement. First, the dataset size and diversity directly influence the model's accuracy. The current dataset may be restricted in scope, potentially limiting its generalizability to other populations. Future research should focus on increasing the dataset size and diversity to include more varied demographic groups, improving the model's adaptability and applicability. In addition, while parameters such as age, BMI, HDL levels, family history of diabetes, PCOS, sedentary lifestyle, and prediabetes contributed significantly to prediction accuracy, incorporating other lifestyle factors (such as physical activity and dietary habits) or genetic markers may further enhance the model's predictive power. Lastly, the practical implementation of the model in real-world clinical settings is essential. This would require improving the model's accuracy and integrating it into healthcare systems with user-friendly interfaces for healthcare professionals. Ensuring ethical, privacy, and security compliance is crucial for the responsible deployment of such systems in clinical practice. In conclusion, future work should focus on expanding the dataset, integrating more features, conducting comprehensive model comparisons, and exploring real-world applications to further improve the effectiveness and reliability of diabetes prediction systems.

REFERENCES

[1] R. E. Pyeritz, "Diabetes Mellitus", Emery and Rimoin's Principles and Practice of Medical Genetics and Genomics (Seventh Edition), Academic Press, 529-588, 2025, ISBN 9780128125311, https://doi.org/10.1016/B978-0-12-812531-1.00001-4.

[2] E. G. Wilmot, P. Mansell, "Diabetes and pregnancy", Clin Med (Lond). 2014 Dec. 14 (6):677-80.

[3] H. D. McIntyre, "Discovery, Knowledge, and Action-Diabetes in Pregnancy Across the Translational Spectrum: The 2016 Norbert Freinkel Award Lecture", Diabetes Care. 2018 Feb. 41 (2):227-32.

[4] A. M. Stuebe, C. Mantzoros, K. Kleinman, et al.,"Gestational glucose tolerance and maternal metabolic profile at 3 years postpartum", Obstet Gynecol. 2011 Nov. 118(5):1065-73.

[5] V. A. Moyer, "Screening for Gestational Diabetes Mellitus: US Preventive Services Task Force Recommendation Statement", Annals of Internal Medicine 160(6), 2014. DOI: 10.7326/M13-2905

[6] Z. Bin, et al., "Worldwide trends in diabetes prevalence and treatment from 1990 to 2022: a pooled analysis of 1108 population-representative studies with 141 million participants", The Lancet, Volume 404, Issue 10467, 2077 - 2093

[7] R. Mittal, K. Prasad, J. R. N. Lemos, G. Arevalo, K. Hirani, "Unveiling Gestational Diabetes: An Overview of Pathophysiology and Management", International Journal of Molecular Sciences. 2025; 26(5):2320. https://doi.org/10.3390/ijms26052320

[8] Republic of Turkiye Ministry of Health General Directorate of Public Health, Available from: https://hsgm.saglik.gov.tr/tr/

[9] Ö. Çelik, N. Rustamova, "Diagnosis and Treatment of Gestational Diabetes", Klinik Tıp Bilimleri Dergisi, Cilt: 7 Sayı: 3 Nisan 2019

[10] Z. Zhang, L. Yang, W. Han, Y. Wu, L. Zhang, C. Gao, K. Jiang, Y. Liu, H. Wu, "Machine Learning Prediction Models for Gestational Diabetes Mellitus: Meta-analysis", J Med Internet Res. 2022 Mar 16;24(3):e26634. doi: 10.2196/26634. PMID: 35294369; PMCID: PMC8968560.

[11] M. Savvidou, S. M. Nelson, M. Makgoba, C. Messow, N. Sattar, and K. Nicolaides, "First-Trimester Prediction of Gestational Diabetes Mellitus: Examining the Potential of Combining Maternal Characteristics and Laboratory Measures", Diabetes. 2010 Dec;59(12):3017-22. doi: 10.2337/db10-0688. Epub 2010 Sep 28. PMID: 20876721; PMCID: PMC2992761.

[12] G. Cubillos, M. Monckeberg, A. Plaza, et. al, "Development of machine learning models to predict gestational diabetes risk in the frst half of pregnancy", BMC Pregnancy Childbirth. 2023 Jun 23;23(1):469. doi: 10.1186/s12884-023-05766-4. PMID: 37353749; PMCID: PMC10288662.

[13] B. S. Kang, S. U. Lee, S. Hong, et. al., "Prediction of gestational diabetes mellitus inAsian women using machine learning algorithms", Sci Rep 13, 13356 (2023). https://doi.org/10.1038/s41598-023-39680-8

[14] B. Kurt, B. Gürlek, S. Keskin, et al., "Prediction of gestational diabetes using deep learning and Bayesian optimization and traditional machine learning techniques", Med Biol Eng Comput 61, 1649–1660 (2023). https://doi.org/10.1007/s11517-023-02800-7

[15] M. Watanabe, A. Eguchi, K. Sakurai, et. al., "Prediction of gestational diabetes mellitus using machine learning from birth cohort data of the Japan Environment and Children's Study", Sci Rep. 2023 Oct 13;13(1):17419. doi: 10.1038/s41598-023-44313-1. PMID: 37833313; PMCID: PMC10575866.

[16] N. S. Artzi, S. Shilo, E. Hadar, et al. "Prediction of gestational diabetes based on nationwide electronic health records", Nat Med 26, 71–76 (2020). https://doi.org/10.1038/s41591-019-0724-8

[17] V. C. Shanmugam, C. Vijayalakshmi, M. Mynarani, and K. Pradeepa Veerakumari, "A study utilizing advanced machine learning techniques to analyze gestational diabetes mellitus and its implementations", JP Journal of Biostatistics 24(2) (2024), 227-237. http://dx.doi.org/10.17654/0973514324015

[18] H. Zaky, E. Fthenou, L. Srour, et. al., "Machine learning based model for the early detection of Gestational Diabetes Mellitus", BMC Med Inform Decis Mak. 2025 Mar 13;25(1):130. doi: 10.1186/s12911-025-02947-3. PMID: 40082942; PMCID: PMC11905636.

[19] A. Sumathi, and S. Meganathan, "Ensemble Classifier Technique to Predict Gestational Diabetes Mellitus (GDM)", Computer Systems Science & Engineering, 2021. DOI:10.32604/csse.2022.017484

[20] J.-S.R. Jang, "ANFIS: adaptive-network-based fuzzy inference system," in IEEE Transactions on Systems, Man, and Cybernetics, vol. 23, no. 3, pp. 665-685, 1993.

GIS-BASED APPROACH TO ASSESSING LAND USE CHANGE IMPACTS ON URBAN HEAT ISLAND PATTERNS

Rabia Bovkir*
Hacettepe University Geomatics Engineering Department, Ankara, Turkiye
(Corresponding author) rbovkir@hacettepe.edu.tr

ABSTRACT

Climate change, coupled with the increasing impacts of rapid urbanization and industrialization, has made environmental sustainability a major concern in contemporary urban planning. One of the most immediate consequences of this unplanned and rapid urban expansion is the Urban Heat Island (UHI) effect, which refers to the increase in land surface temperatures in built-up areas compared to surrounding rural areas. This phenomenon poses significant risks to urban livability, public health and ecological stability. Therefore, understanding the spatial and temporal dynamics of the UHI effect is essential for developing effective mitigation and adaptation strategies. This study investigates the relationship between land use and land cover (LULC) changes and the evolution of UHI intensity in Ankara province between 2015 and 2024. The results show that there has been a significant transformation in the urban landscape of Ankara over the last decade, characterized by an expansion of built surfaces and a corresponding reduction of green and natural land cover classes. These changes have led to a noticeable concentration and expansion of UHI zones, particularly on the urban periphery and in newly developing areas. Across the region, there has been a 57.57% decrease in green areas and a 21.62% increase in urban areas, while there has been a 14.29% increase in UHI areas.

Keywords: GIS, urban heat island (UHI), LULC, climate change.

Introduction

Since the concept of sustainability encompasses multiple dimensions such as the environment, economy, physical and technological infrastructure, and social participation, achieving this goal through singular and simplistic strategies is highly challenging [1]. The integration of sustainability with these multidimensional elements necessitates a strategic perspective, prompting governments and public institutions at all levels to develop new policies, plans, and programs aimed at creating livable, efficient, and resilient urban spaces that address the needs of both current and future generations [2, 3]. For urban centers to be genuinely sustainable, it is essential to prioritize environmental and social dimensions [2].

Within the scope of environmental sustainability, the most pressing issues include climate change and resource depletion, alongside other global environmental problems that pose significant negative impacts [4]. These adverse effects of global climate change can be mitigated through the enhancement of green infrastructure, increased reliance on renew-

able energy sources, and improved recyclability of water resources and waste, thereby minimizing ecological footprints [5]. In doing so, sustainability contributes not only to the improvement of urban life quality but also to the creation of healthier environments for citizens [2, 5]. Central to this transition is the aim to minimize the negative consequences of urbanization and to compensate for the remaining impacts through mitigation and recycling measures.

One prominent outcome of urbanization is the urban heat island (UHI) effect, which refers to the phenomenon where urban areas experience higher temperatures than surrounding natural landscapes [6]. Building materials such as rooftops, pavements, asphalt, and concrete absorb solar energy and contribute to localized climate alterations in urban areas [7]. The ongoing internal migration from rural areas to cities, and from small towns to metropolitan regions, indicates that urbanization will continue in the coming years. Consequently, it is highly likely that the UHI effect will intensify, posing adverse implications for both human and environmental health. In this regard, identifying the spatial distribution of heat island effects is crucial for the development of effective climate change adaptation strategies. The spatial analysis capabilities of Geographic Information Systems (GIS) offer a robust foundation for developing and implementing urban solutions in line with climate adaptation, thereby enabling the creation of sustainable and effective decision-support systems [8, 9]

In this context, the present study examines land use and land cover (LULC) changes in the province of Ankara, which was selected as the study area, between the years 2015 and 2024. This analysis is conducted using LULC imagery obtained from remote sensing platforms and is expressed in hectares across various land cover classes. Additionally, surface temperature variations during the same period were analyzed through the use of thermal imagery. Through this approach, the study aims to identify the extent of the urban heat island (UHI) effect in the region, investigate its spatio-temporal relationship with environmental and anthropogenic factors, and propose potential solutions to mitigate its impacts.

Methodology

Study Area

Located at the coordinates $39°55'48''N, 32°51'00''E$, study area Ankara is situated near the geographical center of Türkiye and serves as the capital and second-largest city of Türkiye (Figure 1). Geographically, Ankara lies between the Kızılırmak and Sakarya Rivers, with the Sakarya River forming its western boundary with Eskişehir. It shares borders with Bolu and Çankırı to the north, Konya to the south, and Kırıkkale to the east. In addition to being the administrative center where major governmental institutions are located, Ankara's central position in Anatolia makes it a strategic hub connecting the eastern, western, northern, and southern regions of the country. This strategic importance contributes to continuous population growth and the presence of anthropogenic factors influencing land use and land cover (LULC) changes.

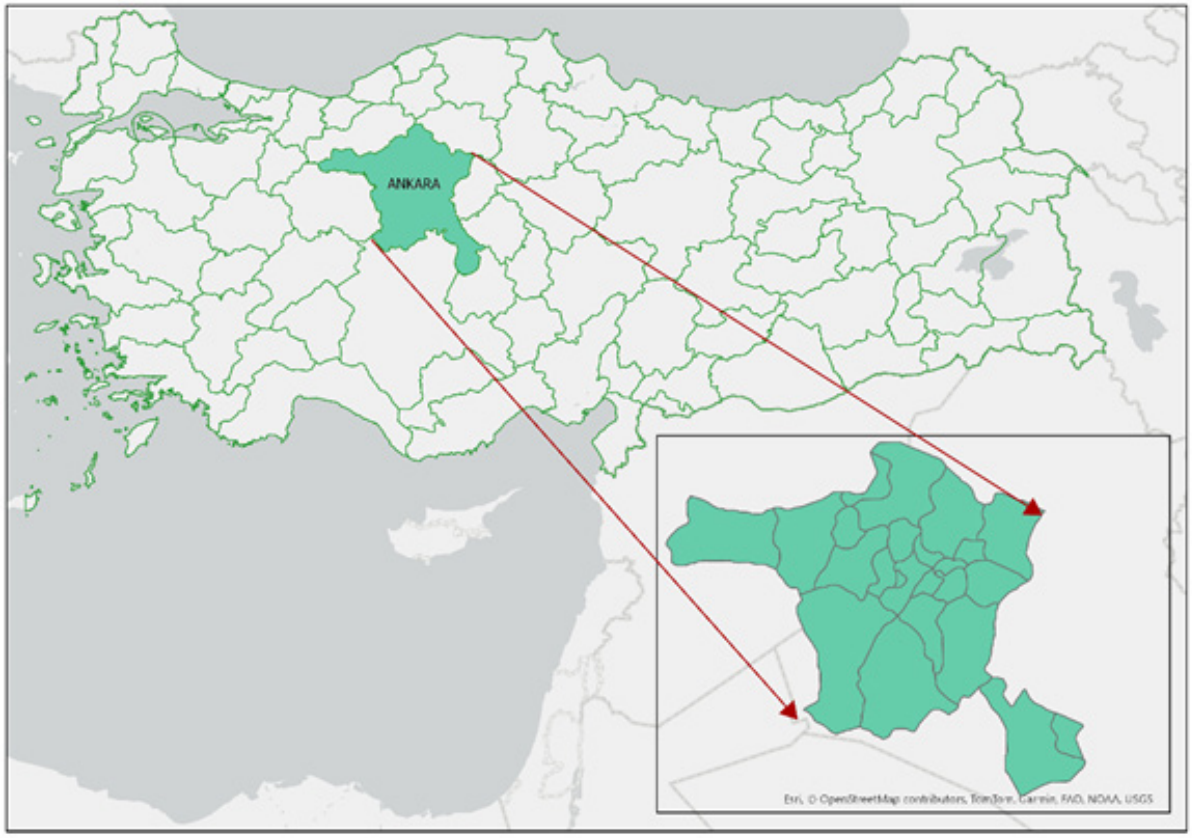


Figure 1: Study Area

According to data from the Turkish Statistical Institute (TUIK) in 2024, the city has a population of approximately 5.8 million, with nearly 5.2 million residents living in urban centers [10]. In this regard, Ankara is Türkiye's second most populous city after İstanbul. Characterized by a continental climate, Ankara experiences a cold semi-arid climate [11]. Due to its elevation and inland location, the city is marked by cold and snowy winters, and hot and dry summers.

Data and Analysis

The most widely used remote sensing platform for determining the Earth's surface temperature is the LANDSAT 8 satellite operated by NASA. The LANDSAT 8 platform includes two different sensor systems. The first is the Operational Land Imager (OLI) and the second is the Thermal Infrared Sensor (TIRS) [13]. Within the scope of the study, average Land Surface Temperature (LST) maps for Ankara were calculated with the help of the Google Earth Engine (GEE). Only images with less than 30% cloud cover were used. For the LST calculation, the Normalized Difference Vegetation Index (NDVI) is calculated using the red and near infrared bands. From the NDVI, how much vegetation is present in each pixel is determined by calculating the vegetation cover ratio (PV). Then, the brightness temperature (BT) is obtained from the satellite's thermal band by calculating the emissivity (ε), which indicates how efficiently the land surface emits thermal radiation. Finally, using a standard equation based on Planck's law, the true land surface temperature (LST) is calculated with the following equation:

$$LST = \frac{BT}{1+\left(\frac{\lambda.BT}{\rho}\right).\ln(\epsilon)} \quad (1)$$

where λ=10.8 μm is the wavelength of emitted radiance and ρ=14380 μm·K is a constant derived from Planc's Law. Figure 2 shows the average LST for Ankara for 2015 and 2024 (Figure 2). The 2015 and 2024 LST maps reveal a significant increase in both the intensity and spatial extent of surface temperatures across Ankara. In 2015, LST values ranged from approximately -8.9°C to 48.5°C, while in 2024 this range widened significantly to -9.1°C to 63.4°C.

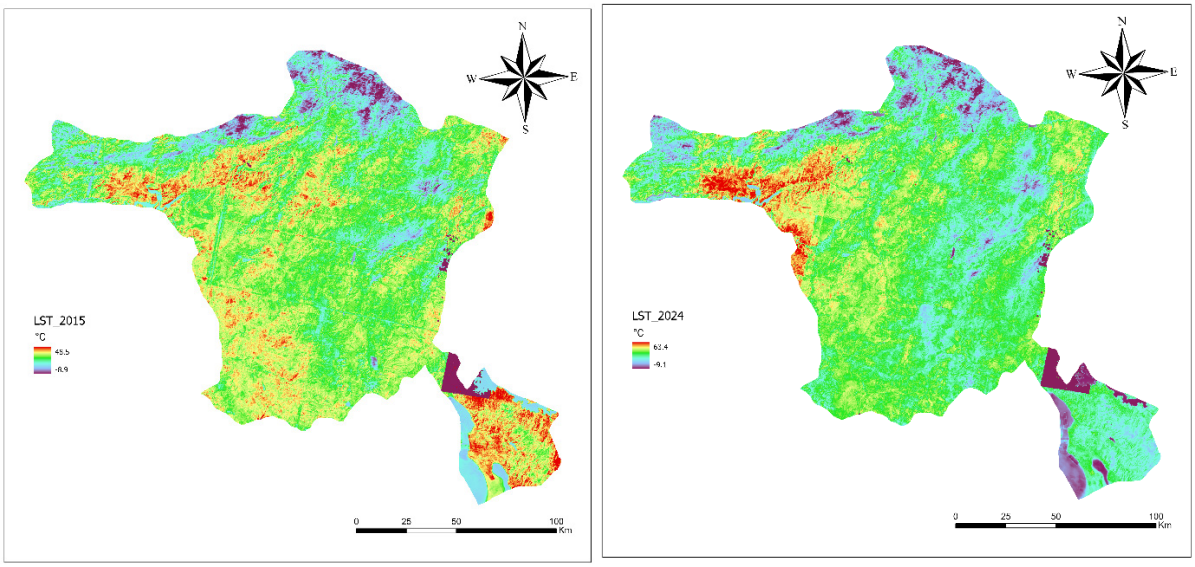


Figure 2: Land Surface Temperature (LST) Maps of Ankara for 2015 and 2024

The UHI index is also calculated by subtracting the average rural LST from the average urban LST pixels, resulting in a raster representing the heat intensity due to urbanization (Eq.2). It is then normalized between 0 and 1 using its minimum and maximum values, producing a standard UHI index for visualization and comparison (Eq.3).

$$UHI = LTS_{urban} - mean(LST_{rural}) \quad (2)$$

$$UHI_{normalized} = \frac{UHI - UHI_{min}}{UHI_{max} - UHI_{min}} \quad (3)$$

Dynamic World [12] land use land cover data was obtained from GEE data catalog. Dynamic World provides a 10 m resolution global near real-time (NRT) global land cover dataset since 2015 for a total of nine classes using deep learning [12]. Its classes are: Water, Trees, Grass, Flooded vegetation, Crops, Shrub and scrub, Built, Bare, Snow and ice. In this study, snow and ice class is included in the water class. Figure 2 represents the LULC map of Ankara for the years of 2015 and 2024 (Figure 3).

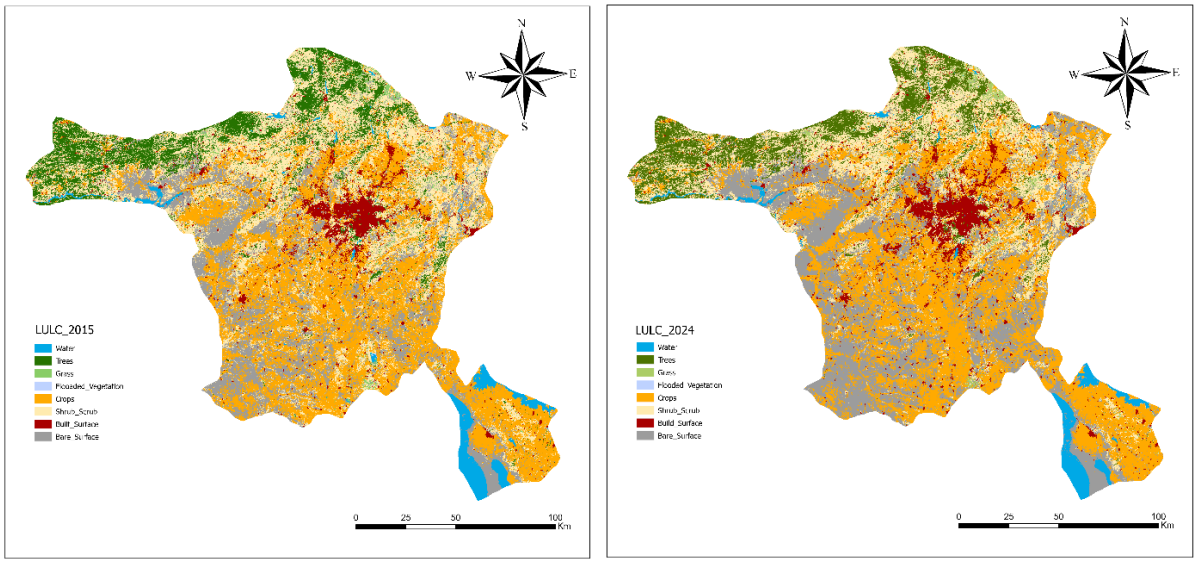


Figure 3: Land Use Land Cover (LULC) Maps of Ankara for 2015 and 2024

Results

After obtaining data for the maps representing the LULC classes for 2015 and 2024 of study area Ankara change analysis was performed. Firstly, green are classes are determined as Trees, Grass, Flooded Vegetation and Crops and their area change in through the years were calculated. The same procedure was performed for also Built surface class. Table 1 shows the land use land cover change for green area and built surface classes for Ankara between the years of 2015 and 2024.

Table 1: Land Use Land Cover (LULC) Change of Ankara for 2015 and 2024

Class Name	Area in 2015 (ha)	Area in 2024 (ha)	Change (ha)	% Change
Trees	282,435.88	264,070.12	-18,365.76	**-6.50%**
Grass	24,619.20	27,191.09	2,571.89	**10.45%**
Flooded Vegetation	1,073.05	404.31	-668.73	**-62.31%**
Crops	914,933.78	922,151.86	7,218.09	**0.79%**
Built	125,545.57	152,692.50	27,146.93	**21.62%**

According to the table, the total green area change is -9244.52 ha and the built surface area change is 27146.93 ha between 2015 and 2024. Between 2015 and 2024, Ankara has undergone significant land cover changes, reflecting a broader trend of urban expansion and environmental transformation. The most notable change is the 21.62% increase in built-up areas of over 27,000 hectares. The main reason for this increase is the rapid urbanization effect due to population growth and infrastructure development in the relevant years. This expansion has been realized at the expense of a decrease in the natural and semi-natural landscape ecology. The area covered by trees has decreased by 6.5%, indicating negative impacts on climate regulation, particularly on biodiversity and air quality, as well as heat islands. An increase of 10.45% in grassland areas can be attributed to the conversion or abandonment of former woodland or agricultural land. Worryingly, inundated vegetation, such as wetlands and marshes, experienced a dramatic decline of 62.31%, which could pose a risk to the ecosystem. Agricultural land expanded at a small rate of only 0.79%, indicating relatively stable agricultural land use.

Figure 4 represents the temporal comparison of the 2015 and 2024 Urban Heat Island (UHI) maps for Ankara (Figure 4). A normalized UHI index for the study area Ankara is calculated using Landsat 8 LST and Dynamic World land cover classification. Using the generated annual average LST and land cover surface, two masks are generated, one for urban areas (class 6: Built) and one for rural areas (classes 1, 2, 4: Trees, Grass, Crops). In this way, urban and rural LST areas are isolated from the average LST composition.

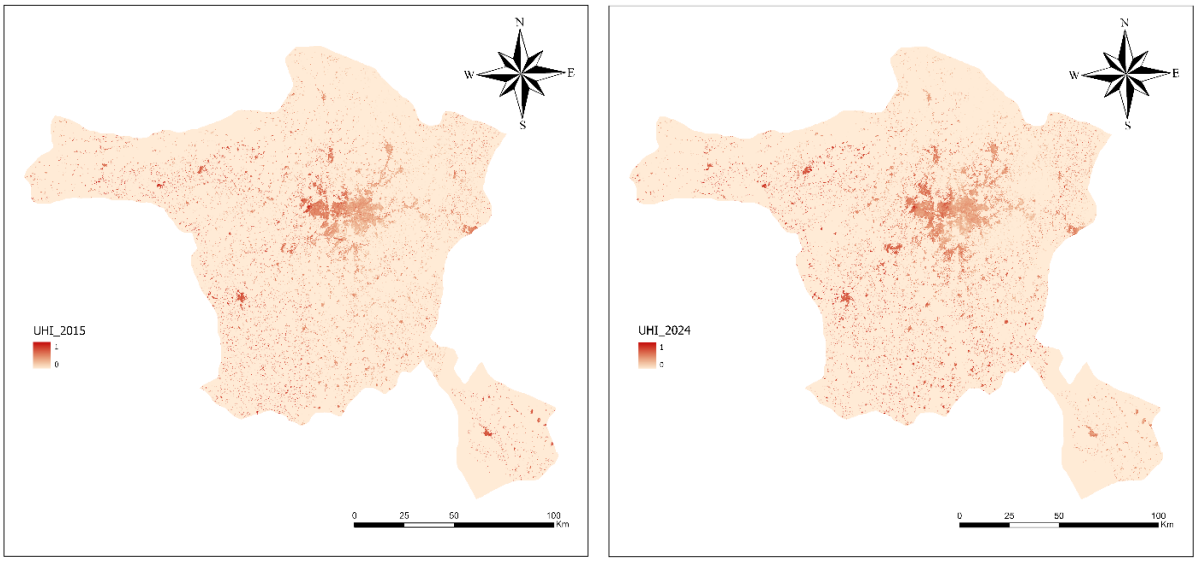


Figure 4: Urban Heat Island (UHI) Maps of Ankara for 2015 and 2024

The maps highlight a significant spatial transformation in surface temperature dynamics induced by urbanization. While the urban core remains the primary hotspot in both years, the 2024 map reveals a wider and more fragmented UHI distribution. The increasing intensity and spatial spread of UHI impacts in the peripheral districts indicates accelerating urban sprawl, intensification and land cover changes such as vegetation loss. Emerging hotspots, particularly in the southern and western parts of the city, reflect the outward growth of Ankara's built-up area, as highlighted by the LULC maps in Figure 3.

Table 2 shows the UHI change in the study area Ankara. In 2015, the total area affected by the UHI in Ankara was approximately 226,827.46 hectares, increasing to 259,233.33 hectares in 2024, a net increase of 32,405.88 hectares. This change corresponds to a 14.29% increase in UHI coverage across the region. The significant growth in UHI-affected areas is consistent with observed land use changes, particularly the conversion of vegetated and agricultural land to built and bare surfaces. This expansion highlights the intensifying thermal stress in Ankara's urban and peri-urban environments and underscores the urgent need for sustainable urban development strategies, green infrastructure and heat mitigation planning.

Table 2: UHI Change of Ankara for 2015 and 2024

UHI Area in 2015 (ha)	UHI Area in 2024 (ha)	UHI Area Change	% Change
226827.46 ha	259233.33 ha	32405.88 ha	**14.29%**

Conclusion

The continued rapid and uneven urbanization of cities around the world, particularly over the last 30 years, is a sign of economic and infrastructural development, but has also brought with it complex challenges for environmental sustainability and human well-being. Among the most critical of these challenges are urban microclimate changes through the Urban Heat Island (UHI) effect, which is evolving in parallel with rising land surface temperatures. As cities grow outward and upward, natural landscapes are replaced by impervious surfaces such as concrete, asphalt and roofs, disrupting energy balances, reducing vegetation-based cooling and increasing surface temperatures. These temperature changes have far-reaching impacts, such as increased energy demand, deteriorating air quality, exacerbating heat-related health problems and reducing biodiversity. Therefore, understanding the spatial drivers and distribution of UHI is crucial for designing cities that are both resilient to climate impacts and conducive to sustainable living.

In this context, this study investigated the spatio-temporal interaction between land use/land cover (LULC) transformation and the urban heat island effect in Ankara over a nine-year period. By integrating thermal satellite imagery and land cover classification data, the study revealed a significant change in the city's thermal landscape. Over the study period, built-up areas expanded significantly, while tree cover and other vegetation classes decreased. These land cover dynamics are directly reflected in the spatial concentration of UHI zones, particularly in the urban core and developing peripheral areas. The findings emphasize that the expansion of urban infrastructure at the expense of green spaces, agricultural land and semi-natural areas is a major contributor to increased surface temperatures and widening UHI impacts.

The shift from a concentrated UHI model to a polycentric UHI model underscores the need for integrated urban planning and climate adaptation measures, such as increasing green infrastructure, promoting sustainable urban design, and implementing heat-reducing land use policies across the metropolitan region. For Ankara and other cities experiencing similar trends, prioritizing green infrastructure and adaptive urban planning is not only a strategic response to climate change, but also a fundamental step towards ensuring environmental justice and urban livability for current and future generations.

REFERENCES

[1] M. Moshrefzadeh, T. H., Kolbe, "Smart Data Infrastructure for Smart and Sustainable Cities", Proceedings of the 13th International Conference on Design & Decision Support Systems in Architecture and Urban Planning (DDSS 2016), June 27-28, 2016, Eindhoven, The Netherlands

[2] The International Telecommunication Union Telecommunication Standardization Sector (ITU-T), "Shaping smarter and more sustainable cities: Striving for sustainable development goals". ITU-T's Technical Reports and Specifications, 2016, Geneva, Switzerland.

[3] M. Peris-Ortiz, D. R. Bennett, Yábar, D.P.B, "Sustainable smart cities. Innovation, Technology, and Knowledge Management", Cham: Springer International Publishing Switzerland, 2017.

[4] R. L., Singh, and P. K. Singh, "Global environmental problems", In Principles and applications of environmental biotechnology for a sustainable future, 13-41, Singapore: Springer Singapore, 2016.

[5] H. Ahvenniemi, A. Huovila, L. Pinto-Seppä, M. Airaksinen, "What are the differences between sustainable and smart cities?", Cities, 60, 234-245, 2017.

[6] L. Chen, R. Jiang, W. N. Xiang, "Surface heat island in Shanghai and its relationship with urban development from 1989 to 2013", Advances in Meteorology, 2016(1), 9782686, 2016.

[7] K. Deilami, M. Kamruzzaman, Y. Liu, "Urban heat island effect: A systematic review of spatio-temporal factors, data, methods, and mitigation measures", International journal of applied earth observation and geoinformation, 67, 30-42, 2018.

[8] K. Soomro, M.N.M. Bhutta, Z. Khan, M. A. Tahir, "Smart City Big Data Analytics: An Advanced Review", WIREs Data Mining and Knowledge Discovery, 9(5), e1319, 2019.

[9] C. G. Fernández, and D. Peek, "Smart and Sustainable? Positioning Adaptation to Climate Change in the European Smart City", Smart Cities, 3(2), 511-526, 2020.

[10] Turkish Statistical Institute (TUIK), "The Results of Address Based Population Registration System, 2024", Retrieved 6 April 2025.

[11] Turkish State Meteorological Service, "Resmi İstatistikler: İllerimize Ait Genel İstatistik Verileri" (in Turkish). Retrieved 6 April 2025.

[12] C.F. Brown, S.P. Brumby, B. Guzder-Williams, et al., "Dynamic World, Near real-time global 10 m land use land cover mapping", Sci Data 9, 251, 2022.

[13] U.S. Geological Survey (USGS), "Landsat 8 ", Retrieved 8 April 2025.

IMPROVING E-MAIL SECURITY WITH ZERO-SHOT LEARNING: A BART-MNLI MODEL PERSPECTIVE

Remzi Gürfidan*
Isparta University of Applied Sciences, Yalvaç Vocational School of Technical Sciences, Computer Programming, Isparta, Türkiye

Ahmet Ali Süzen
Isparta University of Applied Sciences, Faculty of Technology, Computer Engineering, Isparta, Türkiye
(Corresponding author) remzigurfidan@isparta.edu.tr

ABSTRACT

Spam email classification is a very important task in ensuring digital communication security. Traditional machine learning methods employed here rely primarily on vast amounts of labelled data, which are time-consuming and costly to obtain. In this study, zero-shot learning (ZSL) was employed to classify emails as spam or ham without the requirement of task-specific labelled data. Semantic inference between text descriptions and candidate labels was performed through the pre-trained Hugging Face Transformers library BART-large-MNLI model. The raw email messages were subject to aggressive preprocessing that eliminated hyperlinks, HTML tags, punctuation, and stop words to enhance semantic representation clarity. The cleaned text was then classified using ZSL by defining two descriptive labels: "spam email" and "normal email." The model achieved an overall accuracy of 80.99%, with a notably high recall of 92.5% in identifying spam emails. However, the precision for spam detection was moderate, indicating the model's inclination toward false positives. The results demonstrate that zero-shot learning is a viable alternative for spam classification tasks, especially in data-scarce scenarios, and offers significant flexibility and adaptability in dynamically evolving threat environments.

Keywords: Zero-shot learning, spam detection, email classification, natural language inference, BART-MNLI.

Introduction

The learning and decision-making processes of artificial intelligence algorithms are based on data that is as reliable and labelled as a correctly structured model. In cases where the data is insufficient or far from reliable labelling, even the most advanced algorithms will experience underfitting or overfitting problems. This is because artificial intelligence systems can only make sense, recognise patterns and make generalisations with data. For this reason, reliable data that has undergone quality pre-processing plays a critical role in the development of artificial intelligence. But such information in artificial intelligence applications typically must go through labelling processes that are very

time-consuming [1]. Data labelling typically requires specialized skills and an enormous amount of time and money. Especially in issues of the highest importance such as healthcare [2], military technology, security [3] or the financial industry [4], inaccurate or missing data labelling can have serious consequences. This situation negatively affects the uniformity of artificial intelligence systems and restricts the technology from achieving its true potential. Quality data supply and adequate labeling practices are therefore necessary for the effective deployment of artificial intelligence.

Zero-Shot Learning (ZSL) refers to a machine learning paradigm that attempts to provide the ability to identify instances of classes never encountered during training. It builds a semantic connection between observed and unobserved classes usually by way of intermediate representations such as word embedding, attribute vectors or text descriptions. One of the greatest advantages of ZSL is that it reduces the demand for labelled data, which is of critical relevance in use cases such as biomedical imaging, natural language processing or rare event detection where unlabelled data are often encountered. It also makes systems more overallisable and easier to convert to novel situations and eliminate class dependency problems in conventional approaches. ZSL here is at the forefront of creating data-driven and sustainable AI solutions [5], [6], [7], [8].

Li et al. present the METS framework using multimodal ECG-text self-supervised pre-training for improved electrocardiogram (ECG) classification.This approach uses a trainable ECG encoder alongside a frozen language model to independently embed paired ECG signals and automatically generated clinical reports. The self-supervised learning objective is designed to maximize the similarity between these paired embeddings while minimizing it for unpaired samples. Remarkably, METS outperforms other supervised and self-supervised baselines, achieving a performance improvement of around 10% on zero-shot classification tasks without relying on annotated data. In zero-shot classification, a segment of the ECG is used as input [9]. Gowda studied the challenge of recognizing objects from invisible classes by proposing a reinforcement learning-based approach to select synthetic features that improve model performance. He argues that traditional generative methods often produce redundant synthetic features, leading to increased training time and decreased accuracy. To overcome this, the author presents a transformer-based selector trained using proximal policy optimization (PPO) that selects useful synthetic features based on validation classification accuracy as the reward signal. Zero shot learning is both model-independent and data-independent and demonstrates improved performance in feature generation methods, showing versatility in a variety of applications involving images and video[10]Cheng et al. address the challenge of zero-shot image recognition by emphasizing that there must be an active relationship between visual features and feature vectors.They introduce a transformer-based model that allows this active connection to occur and upgrades the model to identify unseen classes more effectively. The proposed method employs a hybrid routing mechanism in the transformer architecture to facilitate effective capture and merging of semantic information. Experimental findings indicate that the method significantly improves classification performance on benchmark datasets and outperforms earlier zero-shot learning models. These findings highlight the importance of active feature-feature relationships in the construction of zero-shot learning methods [11]. Fan et al. present solutions to fault diagnosis of rotating machines in cases with few

or no labelled samples. They propose a metric-based solution where the model attains zero-shot learning through efficient measurement of similarity between unseen failure and known states. It entails building a feature space where unseen and known failure types are matched according to their features and then finding new failures without examples. The model does better than the traditional supervised learning methods, especially when there is limited labelled information. The experiment proves that the process makes the fault diagnosis system more flexible and accurate in actual industrial applications [12]. Chen et al. propose a zero-shot learning approach using the CLIP model for detecting machine operation in earthworks. This method allows the identification of excavator and loader operations with a mean accuracy of 86% and 82.5%, respectively, without training data for the operations. By leveraging the capability of CLIP to match text descriptions with visual inputs, earthwork volume calculation automation is improved by generalizing well to unseen classes. Results demonstrate that zero-shot learning works efficiently in accurately diagnosing construction equipment operations and present an accessible technique for real-time monitoring in construction management [13]. Zhang et al. focus on detecting faults unknown to industrial systems where fault information of certain types is absent for training. They propose a novel zero-shot learning (ZSL) method called SSB-ZSL-1DCNN, which consists of three steps: feature extraction through a one-dimensional convolutional neural network (1D CNN), label embedding through human-designed semantic descriptions shared between seen and unseen faults, and feature embedding through a bilinear compatibility function to determine the most likely fault type. The cosine distance measures the similarity between feature embeddings and fault attributes, enabling the model trained on only seen fault instances to effectively detect unseen faults. Evaluation on two benchmark datasets, the Tennessee-Eastman chemical process and Case Western Reserve University bearing experiments, demonstrates the method's outstanding performance in identifying previously unobserved faults and highlights the effectiveness of integrating semantic space-based ZSL with 1D CNN for intelligent fault detection [14]. Qiao introduces Zero-Shot Deconvolution Networks (ZS-DeconvNet), an unsupervised deep learning framework designed to increase the resolution of fluorescence microscopy images without requiring ground truth data or additional acquisitions. ZS-DeconvNet provides more than 1.5-fold improvement beyond the diffraction limit when operating under fluorescence intensities ten times lower than those typically used in super-resolution imaging. The model uses a zero-shot learning approach that allows it to generalize across a variety of imaging modalities, including total internal reflection fluorescence microscopy, three-dimensional wide-field microscopy, confocal microscopy, two-photon microscopy, lattice light sheet microscopy and multimodal structured illumination microscopy. This versatility enables multi-colour, long-term, super-resolution 2D/3D imaging of subcellular processes from mitotic individual cells to multicellular mouse and C. elegans embryos. The results demonstrate that ZS-DeconvNet effectively improves image resolution and quality in a variety of microscopy techniques and enables close monitoring of dynamic biological processes [15]. Cheng et al. address the issue of denoising Fluorescence Lifetime Imaging Microscopy (FLIM) data, which is challenging due to the complex noise patterns caused by the correlation between intensity and lifetime information. They introduce a new zero-shot denoising model founded on an intensity-guided learning process through the employment of a pre-trained intensity denoising network to guide lifetime component refinement across various channels. The nat-

ural intensity-lifetime correlation is preserved, and important biological information is not eliminated during the denoising process. The method can be used to a zero-shot scenario, i.e. without pre-training over any fixed collection of noise models or labeled data, so it is even more generalizable across a wide variety of FLIM datasets. Experimental results on real biological samples illustrate that the method described here has better performance in removal of noise and maintenance of lifetime information than existing best practice in both and allows for extraction of more accurate physiological and molecular information from FLIM data [16]. Sivarajkumar and Wang talk about issues with developing a clinical NLP system when one does not have access to big annotated datasets. They introduce HealthPrompt, a novel method that utilizes prompt-based learning, a zero-shot learning (ZSL) framework, to classify clinical text without viewing task-specific data in advance. Instead of fine-tuning pre-trained language models (PLMs), HealthPrompt uses task-specific templates to prompt PLMs to learn to encode the context of clinical stories. The authors evaluate HealthPrompt on six different PLMs in training data without settings and demonstrate its capability to perform clinical NLP tasks. This approach highlights the potency of prompt-based ZSL techniques in solving data scarcity issues in clinical text Processing [17].

Material and Method

The dataset utilized for the research was downloaded from the webpage Kaggle [18]. The dataset used for spam email filtering in the present work were pre-processed by applying various pre-processing steps first. Data were read as CSV first and columns required were parsed to get text content and related tag information. Each email message was vigorously cleaned up, especially to allow the model to generalise effectively and reduce the noise content of the message. Here, URL links, HTML tags and punctuations in the source texts were removed, while only the resulting words and phrases with meaning during model training were preserved. In the data cleaning processes, first, all URL addresses from the e-mail texts were removed using regular expressions. Afterwards, HTML code fragments and similar tags were removed from the text. In addition, punctuation marks in the texts were also removed so that the data could be processed more efficiently by the model. The resulting words were reduced by filtering common semantically weak words (stop words) of the English language. Thanks to this cleaning process, the quality of the text contents and the capacity of the model to make sense of them have been increased. Finally, in order to keep the size of the model inputs at a manageable level, the cleaned texts were shortened to a maximum length of 512 characters and presented to the model in their final form.

In this study, it is aimed to perform spam e-mail detection with zero-shot learning approach. Zero-shot learning method is a method that produces successful results in the detection of classes that have never seen training data before, and in this study, the pre-trained Facebook's BART Large MNLI model on the Hugging Face platform was preferred. Since this model is pre-trained in natural language inference (NLI) task, it provides high performance in text classification tasks in different contexts. In order to improve the performance of the model, the classes are explicitly labelled as "spam email" and "normal email". This clear labelling approach helped the model to interpret the classes more accurately. Thus, the predictions obtained were re-labelled as "spam" and "raw" and

evaluated with performance metrics such as accuracy rate, precision, recall and F1-score. In addition, in order to better understand the classification performance of the model, a confusion matrix graph was drawn and visualised.

Findings and Discussions

The classification results obtained are shown in Table 1. According to the results obtained, the overall accuracy of the model was determined as 80.99%. The precision value of 98.6% obtained for the raw (normal) class shows that the model makes very accurate decisions when predicting this class. However, the recall value remained at 79.2%, indicating that approximately one fifth of the real raw messages were incorrectly classified as spam. On the other hand, with a recall value as high as 92.5% in the spam class, it is seen that the model can catch spam messages correctly to a large extent. However, the precision value of 40.7% indicates that a significant portion of the samples that the model labelled as spam were not actually spam. This indicates that while the model provides high sensitivity (recall) in the spam class, it sacrifices specificity (precision). Looking at the macro average F1-score, it is seen that the balance between the classes is relatively weak (72.2%), whereas the weighted average F1-score, which also considers the data distribution, is 83.7%. These results indicate that the model has been optimised to prioritise accuracy, especially in the spam detection task, but there is room for improvement in terms of reducing false positive rates.

Table 1: ZSL Model Results

Classes	Precision	Recall	F1-Score	Support
Ham	0.986	0.792	0.878	4822
Spam	0.407	0.925	0.565	744
Accuracy	0.8099			
Macro Avg F1	0.722			
Weighted Avg F1	0.837			

The confusion matrix in Figure 1 shows the performance of a Zero-Shot Learning (ZSL) based spam email classification model. In this matrix, which is obtained by comparing the actual classes with the classes predicted by the model, it is seen that the model correctly classified 3820 “raw” (normal) e-mails, whereas 688 “raw” e-mails were incorrectly labelled as “spam”. On the other hand, 1002 “spam” e-mails were successfully detected; only 56 spam messages were incorrectly classified as “raw”.

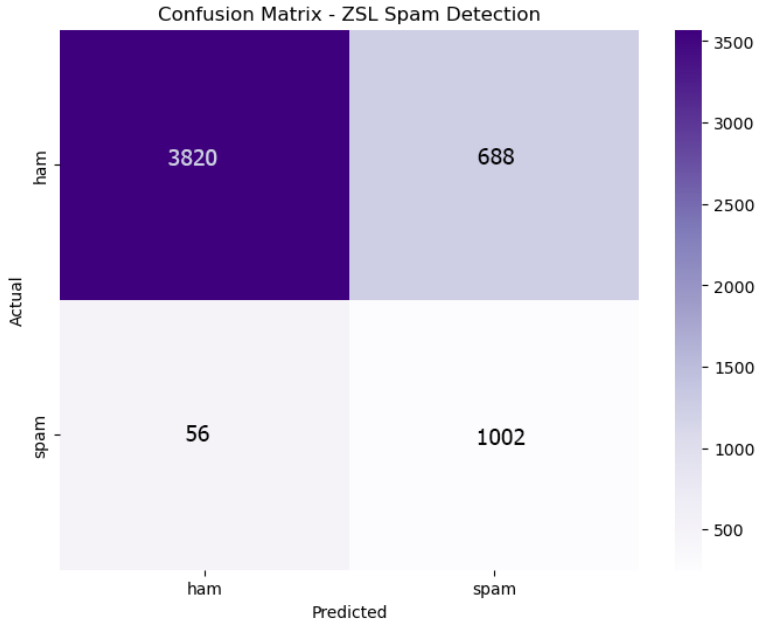


Figure 1: Confusion Matrix

This distribution shows that the success of the model in detecting messages belonging to the "spam" class is quite high and has a low false negative rate (FN). However, the relatively high number of false positives (FP) in the "raw" class suggests that the model tends to mark some innocent messages as spam in order to stay on the safe side. Such behaviour can increase the likelihood of "false alarms", for example in email filtering systems, which can negatively impact the user experience. As a result, it appears that the model works with high sensitivity (recall), but needs to be balanced in terms of specificity (precision).

Conclusions

The findings of this study reveal that zero-shot learning can be effectively applied to spam email detection tasks without the need for labelled training data. The BART-large-MNLI model, when prompted with semantically meaningful class labels, successfully differentiated between spam and ham emails with high accuracy and recall. Particularly, the model demonstrated a strong capacity for identifying spam messages, albeit with a trade-off in specificity, as reflected by the lower precision scores. This indicates a tendency toward over-detection, which, while useful for filtering malicious content, could affect user experience due to false positives. Nonetheless, the ability to generalize to previously unseen classes and operate without training data positions zero-shot learning as a powerful and scalable solution for text-based classification problems. Future work may focus on enhancing prompt engineering strategies or integrating hybrid models to further reduce false positive rates while preserving high sensitivity. This study contributes to the ongoing exploration of data-efficient and flexible AI solutions in the field of intelligent email filtering.

Declaration of Interest: The authors declare that there is no conflict of interest.

Author Contributions: The authors contributed equally to the research, writing, model development and discussion sections of the study.

REFERENCES

[1] Y. Gong, G. Liu, Y. Xue, R. Li, and L. Meng, "A survey on dataset quality in machine learning," Inf Softw Technol, vol. 162, p. 107268, Oct. 2023, doi: 10.1016/J.INFSOF.2023.107268.

[2] P. Eastman et al., "SPICE, A Dataset of Drug-like Molecules and Peptides for Training Machine Learning Potentials," Scientific Data 2022 10:1, vol. 10, no. 1, pp. 1–11, Jan. 2023, doi: 10.1038/s41597-022-01882-6.

[3] B. Kaur et al., "Internet of Things (IoT) security dataset evolution: Challenges and future directions," Internet of Things, vol. 22, p. 100780, Jul. 2023, doi: 10.1016/J.IOT.2023.100780.

[4] S. Kumar, D. Sharma, S. Rao, W. M. Lim, and S. K. Mangla, "Past, present, and future of sustainable finance: insights from big data analytics through machine learning of scholarly research," Annals of Operations Research 2021 345:2, vol. 345, no. 2, pp. 1061–1104, Jan. 2022, doi: 10.1007/S10479-021-04410-8.

[5] W. Ren, Y. Tang, Q. Sun, C. Zhao, and Q. L. Han, "Visual Semantic Segmentation Based on Few/Zero-Shot Learning: An Overview," IEEE/CAA Journal of Automatica Sinica, vol. 11, no. 5, pp. 1106–1126, May 2024, doi: 10.1109/JAS.2023.123207.

[6] W. Alhoshan, A. Ferrari, and L. Zhao, "Zero-shot learning for requirements classification: An exploratory study," Inf Softw Technol, vol. 159, p. 107202, Jul. 2023, doi: 10.1016/J.INFSOF.2023.107202.

[7] W. Cao et al., "A review on multimodal zero-shot learning," Wiley Interdiscip Rev Data Min Knowl Discov, vol. 13, no. 2, p. e1488, Mar. 2023, doi: 10.1002/WIDM.1488.

[8] V. D. Gowda, S. Suneel, P. R. Naidu, S. V. Ramanan, and S. Suneetha, "Challenges and Limitations of Few-Shot and Zero-Shot Learning," https://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/979-8-3693-1822-5.ch007, pp. 113–137, Jan. 1AD, doi: 10.4018/979-8-3693-1822-5.CH007.

[9] J. Li, C. Liu, S. Cheng, R. Arcucci, and S. Hong, "Frozen Language Model Helps ECG Zero-Shot Learning," Jan. 23, 2024, PMLR. Accessed: Apr. 07, 2025. [Online]. Available: https://proceedings.mlr.press/v227/li24a.html

[10] S. N. Gowda, "Synthetic Sample Selection for Generalized Zero-Shot Learning," 2023.

[11] D. Cheng, G. Wang, B. Wang, Q. Zhang, J. Han, and D. Zhang, "Hybrid routing transformer for zero-shot learning," Pattern Recognit, vol. 137, p. 109270, May 2023, doi: 10.1016/J.PATCOG.2022.109270.

[12] C. Fan et al., "A novel metric-based model with the ability of zero-shot learning for intelligent fault diagnosis," Eng Appl Artif Intell, vol. 129, p. 107605, Mar. 2024, doi: 10.1016/J.ENGAPPAI.2023.107605.

[13] C. Chen, B. Xiao, Y. Zhang, and Z. Zhu, "Automatic vision-based calculation of excavator earthmoving productivity using zero-shot learning activity recognition," Autom Constr, vol. 146, p. 104702, Feb. 2023, doi: 10.1016/J.AUTCON.2022.104702.

[14] S. Zhang, H. L. Wei, and J. Ding, "An effective zero-shot learning approach for intelligent fault detection using 1D CNN," Applied Intelligence, vol. 53, no. 12, pp. 16041–16058, Jun. 2023, doi: 10.1007/S10489-022-04342-1/TABLES/17.

[15] C. Qiao et al., "Zero-shot learning enables instant denoising and super-resolution in optical fluorescence microscopy," Nature Communications 2024 15:1, vol. 15, no. 1, pp. 1–15, May 2024, doi: 10.1038/s41467-024-48575-9.

[16] D. Cheng, G. Wang, N. Wang, D. Zhang, Q. Zhang, and X. Gao, "Discriminative and Robust Attribute Alignment for Zero-Shot Learning," IEEE Transactions on Circuits and Systems for Video Technology, vol. 33, no. 8, pp. 4244–4256, Aug. 2023, doi: 10.1109/TCSVT.2023.3243205.

[17] S. Sivarajkumar and Y. Wang, "HealthPrompt: A Zero-shot Learning Paradigm for Clinical Natural Language Processing," AMIA Annual Symposium Proceedings, vol. 2022, p. 972, 2023, Accessed: Apr. 07, 2025. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC10148337/

[18] "Email Spam Detection 98% Accuracy." Accessed: Apr. 08, 2025. [Online]. Available: https://www.kaggle.com/code/mfaisalqureshi/email-spam-detection-98-accuracy?select=spam.csv

IMPROVING INFORMATION ACCESS IN THE HEALTHCARE SECTOR WITH RAG: ENHANCING THE EFFECTIVENESS OF PATIENT INFORMATION FORMS

Murat Genç*

İzmir Kâtip Çelebi University, Department of Computer Engineering, İzmir, Türkiye

**(Corresponding author) Y230246002@ogr.ikc.edu.tr*

ABSTRACT

This study focuses on the use of Retrieval-Augmented Generation (RAG) technology for enhancing the effectiveness of patient information forms and communication processes in the healthcare environment. After identifying existing concerns with current patient information forms and communication processes, this study explores how RAG technology can be used to tackle these problems. In addition to a literature review exploring the structural, substantive, and usability issues concerning patient information forms, a discussion is offered on the fundamental principles of RAG technology, and potential applications of RAG technology in healthcare. From the analysis, results indicate that RAG technology could be useful for strengthening the documentation of patient information forms by improving the accuracy of patient information, expanding access to patient information, and creating patient information that is personalized and contextualized for the patient using RAG technology. The RAG model of a patient information form proposed here offers practical features that include intelligent form completion, dynamic form generation, information verification and enrichment, personalized patient communication, and multilingual support.

Keywords: Retrieval-augmented generation (RAG), patient information forms, health informatics, digital health, patient information processes

Introduction

It is extremely important to have patient information collected accurately, completely, and in a timely manner in the healthcare field for the quality of patient care, as well as to improve the operational efficiency of healthcare organizations. Patient information forms are essentially a fundamental mechanism to collect patients' data, including demographics, medical histories, current complaints, and medication allergies. They ultimately enable healthcare professionals to improve their understanding of their patients and develop treatment plans [1]. In the current landscape, the existing patient information forms and patient information processes have several disadvantages related to data quality, accessibility, timeliness, and personalization. Specifically, traditional patient information forms can easily lead to data entry mistakes, missing patient information, difficulty reading handwriting, and in some cases, the loss of information due to storage issues. Although digital patient information forms help mitigate some of these concerns, they bring their own set

of challenges related to the complexity of the forms, unclear instructions, limited capacity to customize forms, and concerns with data security [2].

While communication gaps and errors between patients and physicians can complicate patient information processes, they also pose several problems. Patients might struggle with understanding medically related terms, their health literacy may be limited, or their time is restricted when learning about their health diagnoses, getting enough information on diagnoses, or any combination of these issues. This can readily lead to non-compliance with a treatment plan, dissatisfaction with their healthcare experience, and also to medical errors [3]. In recent years, there has been a new opportunity in the healthcare field due to artificial intelligence and natural language processing; specifically due to the new approach called Retrieval-Augmented Generation (RAG) technology. RAG technology shows promise in improving the accuracy and reliability of large language models (LLMs) through external sources of information [4]. RAG is a combination of retrieval with text generation models that allows the model to access information that isn't in the training data or outdated.

RAG technology can also be used for various healthcare purposes. Based on retrieving information from sources including electronic health records, medical literature, clinical practice guidelines, and patient assessment data, RAG technology can provide both healthcare professionals and patients with accurate and current information. RAG technology can further be utilized in a range of applications including medical documentation, clinical decision support systems, patient education document personalization, and synthesizing and reviewing medical literature [5]. The purpose of this study is to offer recommendations about how RAG technology can improve the efficacy of patient information forms and improve patient information processing in the healthcare field. This study will begin with problems in patient information forms and patient information processes. The next section will describe core principles of RAG technology and its potential applicability to healthcare. Finally, a model will be designed/tested to improve the effectiveness of patient information forms for patient information processing purposes.

The results from this study may help healthcare organizations enhance patient information forms and patient information processes, which can then improve quality of healthcare services. This study may also provide healthcare researchers and practitioners with a new venue to consider for applicability with RAG technology in healthcare.

Literature Review

Patient Information Forms and Current Challenges

Patient information forms are one of the key parts of healthcare services, which are used to gather information regarding patients' health status and medical history, as well as other relevant information. The Directive on Medical Records and Archiving Services of Inpatient Treatment Institutions of the General Directorate of Treatment Services of the Ministry of Health, emphasizes, above all, "from the moment a patient applies to a healthcare institution, at every stage of treatment, medical interventions and procedures must be fully documented, archived, and preserved by the healthcare institution" [6].

A standard patient file generally contains the following forms:

- Patient Admission Form (Form 60)
- Medical Observation and Examination Form (Form 62)
- Condition Assessment Form (Form 61)
- Patient Chart (Form 51)
- Radiology Request Form and Reports (Form 64)
- Laboratory Request Form and Test Reports (Form 65)
- Surgery Form (Form 63)
- Discharge Summary (Form 67)
- Patient Consent Form [7]

Patient records are generally composed of three main components: the identification information component, the nursing component, and the medical records components. Typically, the identification information component includes the demographic information of the patient and is generally filled out by medical secretaries. The nursing component is a record of the patient's vital signs, treatment, and follow-up and is generally filled out by nurses. The medical records component is the responsibility of physicians and contains information relating to diagnosis and treatment [8].

Current patient information forms and patient information processes include the following problems:

Quality of data problems

When patient information is filled out on paper forms, problems can occur, such as data entry errors, missing information, or poor handwriting, or even lost files. In a study, it was found that there was either missing or incorrect information in 30% of patient information forms [9].

Access problems

Paper forms cannot be accessed from various locations, which has implications for patient care and continuity of care. When patients apply to different health systems, health professionals are unable to access the previous physician's information, which contributes to unnecessary and repeated tests and treatments [10].

Timeliness problems

Failure to update information in a timely manner has implications for care quality. Particularly for patients with chronic illnesses, the inability to update time-sensitive information can result in incorrect treatment [11].

Customization deficiencies

Existing patient information forms will typically have a standard structure, with lim-

itations to customize it for specific patients. This can result in a situation where patients encounter information that uses confusing medical terms and hard-to understand medical information [12].

Deficiencies in communication

A significant component of the patient information process is understanding communication deficiencies, and mistakes in patient-physician communication. Patients often feel that their medical condition is not clearly, and/or accurately, explained to them by the physician because of time constraints, patients limited health literacy, and difficulties in understanding medical terminology [13].

Data privacy and security problems

Protecting patient information is an ongoing challenge for healthcare institutions. Information stored on paper forms can be vulnerable through unauthorized access, lost or damaged paper forms; digital forms, however, can be the victim of data breaches [14].

While digital patient information forms have the potential to address some of these issues, there are also some challenges with digital forms. In a 2024 study by Bayro examining the pros and cons of internet-based patient admission forms, it was noted that the pros of digital forms are speed to collect patient data, quality of care, insights for the patient, capacity to be using a variety of devices, and positive experiences for the patient in using the digital form. The cons of digital forms are often complicated, unclear instructions, weak customizing capabilities, and data security and privacy issues [15]. In a study on the digitalization of hospitals in Turkey, Vermişli Peker et al. (2018) also pointed out that security of information is still a main concern in a digital hospital setting. Even if healthcare professionals are able to access a person's patient record, it may be disadvantageous if there was no time or location for control privacy protection concerns [16].

RAG (Retrieval-Augmented Generation) Technology

The technique called retrieval-augmented generation (RAG) can enhance the accuracy and trustworthiness of large language models by allowing them to use some knowledge from outside the language model itself. This approach was proposed by Lewis et al. (2020) in "Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks" [17]. When a model is used with the retrieval-augmented generation method, it has a retriever (a retrieval component) and a generator (a text generation model) both running simultaneously. This means the model can add information that was not present in its training or may have become outdated. Retrieval-augmented generation is especially powerful for knowledge-intensive tasks because it enhances accuracy and trust in the models [18].

The essential elements of RAG are as follows:

Knowledge Base

This includes structured or unstructured data from different external sources. This can consist of a large variety of sample topic sources, such as medical literature, clinical guidelines, electronic health records, and patient data.

Retriever

The retriever is responsible for finding relevant information in the knowledge base connected to the user's question/query. Generally, the retriever will use a vector-based search engine or information retrieval model to retrieve documents or passages that provide the most relevant information.

Generator

This is the large language model (LLM) that uses the content retrieved to generate contextually appropriate and correct answers. The generator produces a response to the user query connected to the content retrieved [19].

The following outlines the process of RAG:

•A user either makes a request or poses a question.

•The retriever acquires information relevant to the user request or question from a knowledge base.

•This information is then part of the inputs for the generator.

•The generator then uses this information to generate a contextual and accurate response. [20].

The advantages of RAG are

•More precise and trustworthy: RAG provides information from new and verified third parties who provide accurate and reliable information to add to the model's responses, especially in sensitive areas (such as health).

•Transparency: RAG can trace where the response is obtained. The user can see what information was referenced in the response of the generation model and validate the accuracy of the information.

•Easy updating: The RAG knowledge base can easily be updated. In instances where the training data of the model has become dated or used old information, RAG can always be a reference to current information.

•Personalization: RAG can be directed to create personalization responses for a certain user or institution by accessing custom information. [21].

NVIDIA (2025) noted that RAG technology takes advantage of advances with large language models with domain-specific information from credible sources of information. The benefits of RAG will be particularly enlightening in healthcare, where it is $450,000 to wait on information that is always available and up-to-date either from technical or medical perspectives. [22].

RAG Applications in the Healthcare Sector

RAG (retrieval-augmented generation) technology has promising applications in healthcare, such as enhancing patient data, enhancing clinical decision-making activities, and

augmenting patient engagement processes. A recent article by HealthTech magazine (2025) reviewing practical applications of RAG in healthcare highlights several benefits of RAG technology for healthcare organizations and institutions:

- Automation of medical documentation and coding: RAG technology can automatically summarize and code medical records. This can free healthcare providers from the burden of documentation while improving medical coding accuracy.

- Clinical decision-support systems: RAG technology can access information from clinical guidelines, the medical literature, and patient data to assist healthcare providers in diagnosis and treatment plans.

- Individualization of patient education materials: RAG technology can automatically generate education materials, individualizing the education materials based on the patient's health literacy, preferred language, and medical needs.

- Synthesis and review of medical literature: RAG technology can extract medical literature to summarize and review relevant information related to a clinical question or research domain.

- Answering patient questions: RAG technology can accurately and transparently respond to patients' questions related to health concerns, treatment options, or medication usage [23].

A case study conducted by AWS (2025) uses RAG technology to show how RAG can be used to enhance patient data in the healthcare space. In this example, RAG technology was trained to obtain patient data, augment the patient data with literature and clinical guidelines related to patients' medical problems, and then present a comprehensive patient profile to the provider [24].

To realize RAG applications in healthcare, multiple issues need to be resolved internally and externally, including data privacy and protection challenges, legal and ethical dilemmas, technology use by healthcare professionals, and integration with other systems. Protection of patient data is especially critical in ensuring compliance with laws and regulations (e.g., HIPAA [Health Insurance Portability and Accountability Act]) [25].

Overall, while the applications of RAG technology are still early and in development phases, RAG technology has capabilities to improve the quality of care for patients, reduce workloads for healthcare professionals, and promote timely, effective and efficient healthcare for patients and their communities. RAG technology has great potential for improvement in areas such as patient engagement forms and patient communication processes [26].

Methodology

This study utilized literature review and conceptual model development methodologies to investigate how RAG technology could be utilized to improve the effectiveness of patient information forms and improve patient communication processes practice and theory.

Research Design

A qualitative approach using exploratory design was adopted. This design captured a comprehensive exploration of what might result from RAG technology on a patient information form and patients communication processes in the healthcare context. The study was broken into two primary stages:

Literature Review

A systematic review of prior available literature related to a patient information form, health informatics, digital health, and RAG technology.

Conceptual Model Development

A model was designed based on literature review findings to demonstrate how RAG technology could be applied to make a patient information form more effective.

Data Collection

As part of the literature review, data was collected from the following sources:

- Academic Articles: Peer-reviewed journal articles related to health informatics, patient information forms, artificial intelligence, and RAG technology.

- Technical Reports: Technical reports and white papers published by healthcare institutions, technology companies, and research organizations.

- Industry Analyses: Industry analyses and market research related to digitalization and artificial intelligence applications in the healthcare sector.

- Legal and Regulatory Documents: Documents outlining the legal and regulatory frameworks related to patient information, data privacy, and health informatics.

The data collection process was conducted using the following keywords and their combinations: "patient information forms", "health informatics", "digital health", "artificial intelligence", "RAG", "Retrieval-Augmented Generation", "patient communication", "health data management".

Data Analysis

The present study employed literature review and conceptual model development methods to explore the ways RAG technology could enhance the efficacy of patient information forms and improve patient communication processes.

The study took a qualitative stance with an exploratory design which enabled a holistic examination of the potential implications of RAG technology on patient information forms and patient communication processes in health care. The study was framed into two stages:

- Literature Review: A systematic review of the current literature related to patient information forms, health informatics, digital health, and RAG technology

•Conceptual Model Development: Based on the evidence derived from the literature review, a model was developed to demonstrate how RAG technology could be used to enhance the efficacy of patient information forms.

The analysis revealed themes pertaining to problems with patient information forms and communications, general principles and principles of RAG technology, and the potential role of RAG technology in the healthcare domain.

Conceptual Model Development

From the material collected from the literature review, a conceptual model was created to demonstrate how RAG technology could be employed to improve the utility of patient information forms. This model includes:

•RAG Based Patient Information Form System: A system that utilizes RAG technology to collect, validate, and enrich patient information.

•Information Sources: Data source integration such as electronic health records, medical literature, clinical practice guidelines, and patient data.

•User Interfaces: User-friendly interfaces for healthcare providers and patients.

•Workflows for Data Processing: The necessary workflows to collect, validate, enrich, and present information about the patient.

•Security and Privacy: The necessary protocols for protecting patient data through appropriate security and privacy protocols.

Supported by application scenarios, the conceptual model was meant to show how RAG technology would be an effective means to utilize RAG technology to improve the quality and function of patient information forms.

Ethical Considerations

As this study utilizes literature review and conceptual model development methodologies, there are no human participants involved. However, safeguarding the privacy and confidentiality of patient data is an ethical imperative in health informatics applications. As a result, the model indicates that when the proposed model is put into operation, there is a prompt that patient information and health data must be sufficiently protected. Furthermore, ethical considerations have been examined surrounding the use of artificial intelligence and RAG technology in health care. For example, transparency, explainability, and equitable use of the technologies inform the proposed model's ethical framework.

Findings

RAG-Based Patient Information Form Model

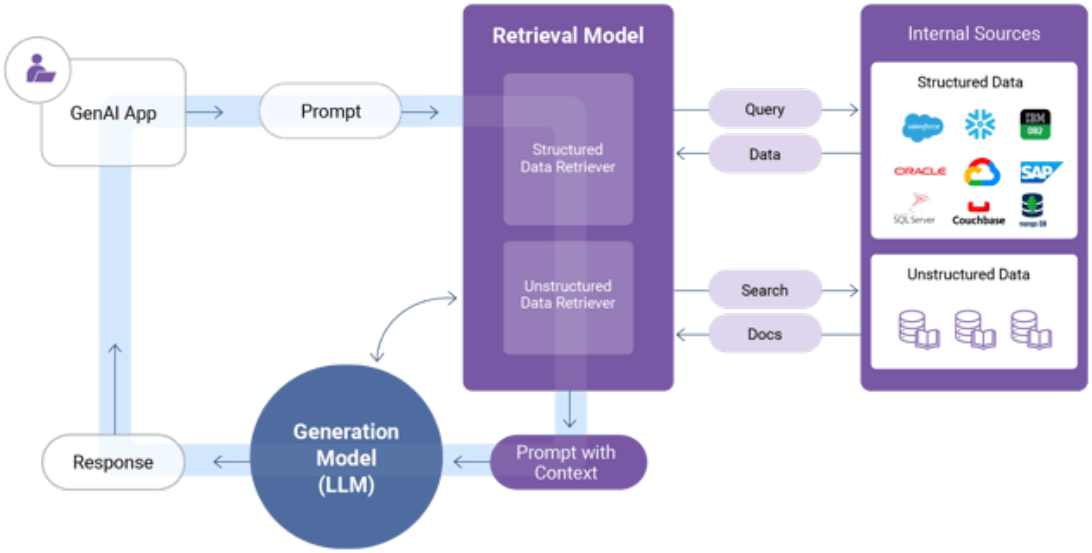


Figure 1: Prototype architecture for the proposed system of patient information form system with RAG foundations.

This workflow is illustrated here as a retriever model fetching and collating the patient's most relevant information from corporate Internal Sources (structured and unstructured data sources) through either Query/Data or Search/Docs workflows. Once the data has been retrieved, it will be sent on to the Generation Model (LLM) as a prompt with context where it will be processed. Finally, the architecture illustrates how forcing the generative model to produce a response will benefit by including outside information sources. The fundamental advantage of this architecture is that the RAG based system can generate more convincing, timely, and verified outcomes in the patient information form workflows.

Retrieval Augmented Generation (RAG) Sequence Diagram

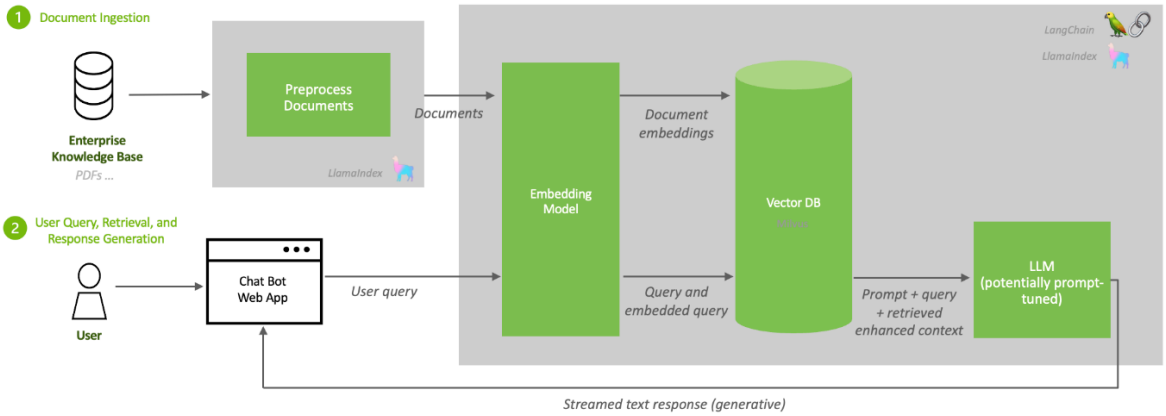


Figure 2: Process flow diagram of the RAG-based system.

This diagram demonstrates the process in two segments, from the stage of preparing documents and data to producing responses to user queries.

•Document Loading: Documents from the corporate knowledge base are pre-processed into chunks, which are then saved in a vector database (creating and storing document embeddings).

• Query and Response: When a user submits a query, relevant chunks from the vector database are retrieved, and the document chunks are sent to the LLM model as enriched context with the user's query; this model, therefore, generates a correct and contextually relevant response.

Hence, the process demonstrated in Figure 2 results in the RAG-based systems' ability to produce output based on the most current and reliable information with real-time access to data. As mentioned earlier, the work process operates fundamentally in the same manner for any application situation described in the following section. RAG technology can assist in improving and increasing the efficiency of patient information forms in a number of different ways. The basic functions and features supplied by this model are summarized below:

Intelligent Form Filling

RAG technology can promote automation of patient information forms or provide suggestions by utilizing prior records of the patient and medical literature. The advantages of this capability include the following:

•Time Savings: This capability saves healthcare professionals or patients significant time in completing forms and documentation.

•Accuracy of Data: Information is populated from prior records, which reduces the likelihood of data entry errors.

•More Detailed Information: There is a possibility of more detailed collection of information about the patient's medical history and current condition.

As an example, if the patient has been seen at the prior healthcare establishment utilizing the same RAG technology, the RAG technology could prepopulate the new patient admission information form with demographic information, medical history, or documented drug allergies that are pulled from prior records. The healthcare provider could complete the admission form by additional validation or updating of the information that had been populated into the admission form [27].

Dynamic Form Generation

RAG systems can create a form that reacts to the patient's status and needs. This allows for the following:

•Forms improved for the patient: Forms could be made to be more patient-specific in regard to the patient's medical complaint, age, sex, and so forth.

•Relevant questions: The Clinician will ask the most relevant questions based on the patient's medical complaint, etc., and will be able to skip asking irrelevant questions.

•More Detailed Information Gathered: Clinician can gain more in-depth information in certain key areas based on the patient's response to prior questions.

As an example, a patient presents with complaints of diabetes. In this example, the RAG generates a form that is going to have diabetes-related questions. After the patient has answered the question, the RAG will then ask more questions related to the treatment plan for diabetes, outpatient diabetes-related issues, and lifestyle [28].

Information Verification and Enrichment

The RAG technology is able to verify and augment the information provided by the patient by comparing it to medical records and medical literature. Put differently, it provides the following benefits:

•Data Consistency: It can assure consistency of patient information, as well as identify contradicting information.

•Completing Gaps in Data: It can complete data that the patient omits from other sources.

•Information Augmentation: It can augment patient information with medical literature and clinical guidelines.

For example, if a patient does not adequately complete their medication list, or if he/she fills it out incorrectly, the RAG system could complete or fix the patient's medications by accessing information from the patient's electronic health records. The system could also alert the physician by accessing information on drug interactions or side effects from the medical literature [29].

Personalized Patient Information

The RAG technology can adapt informative material according to the patient's health literacy level, language preference, and specific medical condition. This benefit can be of use in the following ways:

•Comprehensibility: Presenting information using language and words that the patient understands.

•Personalized Content: Tailored informative materials by the patient's medical state and needs.

•Media Communication: Using media communication based on the individual patient pREFERENCES (written, visual, auditory).

For example, when a patient is being advised on a medical procedure, the RAG system is able to customize the informative material regarding the patient's health literacy level. For example, for a patient that has lower health literacy levels, the RAG can implement easy-to-read language and visual cues with the support of technology. The adaptation of

personalized information lay in adding a few additional health factors that represent the patient-specific conditions [30].

Multilingual Support

The RAG technology is capable of creating patient information forms and educational materials in every language. This academic technology offers the following features:

•Breaking Language Barriers: Accessing healthcare services for patients who speak alternate languages.

•Accurate Translation: Accurate translation of medical terms, preventing the loss of meaning.

•Cultural Adaptation: Adapting materials according to cultural context.

For example, for a patient who is not able to speak Turkish, the RAG system has the capability to create patient education forms or obtain patient education materials in the preferred language of all patients. The system can prevent medical errors resulting from communication and misunderstandings caused by the inaccurate translation of medical terms.

Application Scenarios

Here are the detailed potential application scenarios of the RAG-based patient information form model:

Patient Admission Process

In the patient admission process, RAG technology can facilitate automatic completion of the patient admission form using the patient's previous health records and demographic information. This scenario has the following steps:

•The patient applies to the healthcare institution and provides identification information.

•The RAG system uses the patient's identification information to retrieve the patient's demographic information, medical history, and drug allergies from electronic health records. - The system automatically populates the patient admission form with this information. - The healthcare professional shows the completed form to the patient and asks for confirmation that the information is correct.

•The patient confirms the information, or a revision is made by the patient.

•The system saves the confirmed information and sends it to the appropriate departments. This scenario speeds up the patient admission process, reduces data entry errors, and improves the patient experience [32].

Medical History Taking

The physician can utilize a RAG-supported system to interactively generate pertinent medical questions based on the patient's complaints and generate a more complete medical history of the patient. This situation involves the following steps: The patient provides

their complaints to the physician. The physician types the patient's complaints into the RAG system. The system generates a dynamic medical history form that contains relevant medical questions based on the complaints. The physician utilizes this form to elicit relevant detailed information from the patient. Then the system suggests additional questions and supplies information from relevant medical literature, based on the patient's responses. Finally, the physician develops a diagnosis and treatment plan based on the information they have collected. This situation allows the physician to obtain a more thorough and focused medical history, minimizes the chances of missing important questions, and increases the likelihood of making an accurate diagnosis [33].

Patient Information and Consent Forms

The RAG system can generate tailored patient information and consent forms based on a patient's medical condition, health literacy, and linguistic pREFERENCES. The outline of the scenario is as follows:

- The physician selects a medical procedure to be conducted on the patient.
- The RAG system retrieves existing literature on medical processes relevant to the procedure.
- The RAG system prepares a tailored patient information and consent form appropriate to the patient's health literacy level, linguistic pREFERENCES, and medical condition.
- The physician then uses this form to inform the patient and to attain consent for the procedure.
- The patient can ask questions about anything they don't understand, and the RAG system will answer these questions.
- The patient provides informed consent, and the form is recorded electronically.

This scenario will allow the patient to be better informed about the medical procedures performed, support communication between the physician and the patient, and improve the quality of the informed consent process [34].

Post-Discharge Care Instructions

Personalized care instructions for after the hospital stay can be adapted for each patient based on their medical condition, medications, and living situation. We envision the following process taking place:

- The physician decides to discharge the patient.
- The RAG system analyses the patient's current medical condition, a summary of medications, treatment plan, and living condition.
- The RAG system generates personalized care instructions for them based on these factors.

•The care instructions are presented in a language and format that the patient can understand (i.e., written document, mobile app, etc.).

•The physician discusses the instructions with the patient and answers questions.

•The patient receives the discharge instructions and goes home, continues care according to the instructions.

This scenario provides patients a better mechanism to understand and follow care instructions after discharge from the hospital, improves adherence to treatment, and reduces hospital readmissions [35].

Table 1: Key Problems with Patient Information Forms and RAG-Based Intervention

Problems	**RAG-Based Intervention**
Issues of data quality (missing/inaccurate information)	Smart form auto-fill & information verification: RAG auto-fills forms based on previous patient records entered into the system while also verifying the data entered on new patients to minimize errors and omissions.
Issues of acceptability (being unable to access records from multiple sites)	Digital interoperability: RAG-based will be a digital system which integrates digital data that's part of patients' electronic health records and is not subject to the historical (or existing) barriers of time and place for provider access.
Issues of currency (not being revised in a timely manner)	Dynamic information maintenance: RAG will continuously stay connected to current clinical literature and clinical databases so that information in the forms and presented content will be in current form for the forms.
Issues of personalization (being unable to personalize forms)	Personalized forms & content: Dynamic forms created by RAG will allow patient-specific information to become part of the continuous updating process of the forms based on outcomes to patient profiles so that special questions and special explanations are generated for the specific patient.
Issues of communication (between patient and clinician and vice versa)	Multilingual and usable information: RAG creates information, including consent documents, that can be presented to the patient in their language and at a personal level of education or understanding that equalizes communication barriers; patient questions would be able to be answered in real-time without deficit of currency.
Issues of data security / privacy (patient-sensitive data)	Strong security integration: A RAG-based solution will be built with security protocols around access control, data encryption and anonymization to protect patient privacy.

Technical Requirements and Implementation Challenges

Technical Requirements

Data Infrastructure

A cohesive data infrastructure will be necessary to bring together different data sources (such as electronic health records, medical literature, and clinical guidelines). Interoperability will be needed to relate data from different sources and systems.

Natural Language Processing (NLP) Capabilities

High-level NLP capabilities will be necessary both to interpret the subtleties of patient and clinical texts and to provide contextually appropriate replies. Familiarity with medi-

cal-specific terminology and multilingual processing capabilities are also important components.

Large Language Models (LLM)

Robust, medically sophisticated LLMs may need to be leveraged for the generator component of RAG. It will also be very helpful to be able to adapt or train these models for working with health-related data.

Vector Databases

A capable and efficient vector database will be essential for the retrieval activities. Additionally, a scalable vector data store should be capable of retaining vectorial representations of documents and records and able to conduct rapid similarity searches. - Security and Privacy: Strong security and privacy elements need to be integrated into the framework to protect patient data. Access controls, encryption of data, and anonymization are some of the ways to build security and privacy into the RAG environment.

User Interface: User interfaces will need to be built that are user-friendly, simple, and straightforward for both patients and clinicians to allow for use with minimal concerns or anxiety. Clinicians, nurses, and patients should be able to use the system without any regard for complex technologies, and the complexities need to be hidden from the user [36].

Implementation Challenges

Data Quality and Normalization

From the standpoint of data quality from several sources and normalization in a specific format, discrepancies could affect the performance of the RAG system, as data sanitization and normalization are a real struggle, and any inconsistency or bad data could generate invalid outputs.

Language and Cultural Competencies

The RAG system may need to accommodate multiple languages and cultural contexts, and an awareness of the cultural context specific to each medical condition is an important consideration for effective communication in patient care. It would be important to transfer medical terminology and phrases accurately across languages and be aware of cultural context. Any deficiencies in this competence could limit its utility for certain populations of patients.

Legal and Ethical Considerations

Legal and ethical considerations regarding patient data collection and usage of artificial intelligence applications may have a bearing on the utilization of the RAG system. It will have to be established based on data privacy legislation (e.g., KVKK, HIPAA) and ethics. Violation of these guidelines can result in legal implications, as well as trust issues.

Healthcare Provider Adaptation

The healthcare provider's ability to adjust and incorporate new technology will be crucial to the successful implementation of the RAG system. They will need to know how to use and adequately employ the technology in clinical practice in their routines. Resistance to change or lack of training can restrict the uptake of the system.

Resource and Financial Considerations

There will be considerable resources and expense associated with the RAG system development and implementation. Particularly for smaller healthcare systems, the cost of infrastructure, the promotion of the best LLMs and hardware, and upkeep will be significant [37].

To mitigate these challenges, collaboration between healthcare systems, developers, vendors, regulatory bodies, and researchers will be necessary. Additionally, the successful implementation of the RAG will be gradual, with iterative improvements. [38].

Table 2: Technical Specifications of RAG-based System and Challenges of Implementation

Technical Specifications	Implementation Challenges
Data infrastructure and interoperability (integration of EHR, literature, etc.)	Data quality and standardization - The need to standardize data from multiple sources and convert data into common formats.
The need for advanced NLP capabilities (medical language processing, multilingual support)	Language and cultural barriers - The issue of ensuring the effectiveness of the system works in other languages and cultures.
The requirement for quality LLMs (specialized or adapted models)	Cost and resource needs - Limited fiscal and infrastructure resources for training/integration of large language models.
User-friendly interface (for the HS and patients)	User adaptation - Resistance or training needs in the process of HS getting use to the new system and integrating their work processes with the new system.
Strong security and privacy (access control, data encryption)	Legal and ethical compliance - The need to ensure full compliance with ethical standards and legal regulations regarding the use of patient data.

Discussion and Conclusion

The study examined RAG technology's role in enhancing hospitals' patients' information forms and patient information management. According to the findings of the study, RAG technology has the potential to improve the quality of health care services and have a positive impact on patient information forms due to the accuracy, access, and tailoring of patient information forms.

Key Findings

To summarize the main points from the study:

Inadequate Patient Information Forms

Patient information forms face a number of inadequacies including data quality, acces-

sibility, timing, and personalizing the form. Digital patient forms overcome some of these inadequacies, however, there are other challenges. Digital patient forms can have complicated forms, non-informative prompts, lack of personalization, and privacy risk.

Potential with RAG Technology

Retrieval-Augmented Generation (RAG) technology has the potential to improve patient information forms with the use of external knowledge sources to support the trustworthiness and credibility of large language models. Intelligent form fill, dynamic forms, confirm and enrich with information, patient education, and multilingual can augment the experience and trustworthiness involved with a patient information form.

Contexts with RAG can be

The RAG-based patient information form can be used in contexts such as the patient admission process, patient medical history, providing education and seeking consent from patients, and patient discharge teaching. The contexts can provide the foundation of a proof-of-concept for RAG model improving high-quality patient care experience.

Technical Considerations and Challenges

There are many technical considerations to implement a RAG-based patient information form, as well as challenges. The technical considerations are data infrastructure, natural language processing capabilities, quality large language model selection, efficient vector database selection for embedding capabilities, privacy and security issues, and user experience considerations. Additional items that require consideration include technical challenges such as data quality and standardization, linguistic and cultural considerations, legal and ethical considerations, uptake by HC PAS, and cost/resource.

Contributions of the Study

This article contributes to the literature in the following areas:

- Conceptual Framework: It provides a conceptual framework for how Retrieval-Augmented Generation (RAG) technology may enhance the effectiveness of patient information forms.

- Application Model: It provides specific information about the design, components, and use cases of a RAG-based patient information form model.

- Challenges and Solutions: It identifies potential barriers to the mobilization of RAG technology in the healthcare context, as well as potential solutions to overcome barriers.

- Directions for Future Research: It provides directions for future research to understand future directions for RAG technologies relative to patient information forms and communication processes in healthcare.

Limitations of the Study

There are a number of limitations to the present study:

•Focus on Theory: In general, the study presented a theoretical framework of RAG technology THAT could enhance the effectiveness of patient information forms. Exploring the practical application and impact of the proposed framework of study was deemed appropriate for future research.

•Technology Evolving: Given the rapid development of RAG technology and artificial intelligence more generally, it is possible that some of the findings and recommendations may be dated in their effectiveness. Regular updates may be needed over time.

•Cultural and Contextual Considerations: The study did not examine the implementation and effectiveness of RAG technology to study in other cultural and contextual scenarios. The context and implementation of RAG technology in the effectiveness cohort may differ between countries and healthcare systems.

Future Research Directions

The results of this study will contribute to the developments in the following research areas: Prototype Development and Testing: Develop prototype applications of the RAG model for patient information forms to be tested in the clinical setting. Impact Analysis: Explore the impact of RAG technology on patient outcomes, patient satisfaction, and efficiency of health service delivery. Cost-Effectiveness Analysis: Analyze the cost-effectiveness of the RAG model for producing patient information forms compared to traditional information methods. Ethical Legal Analysis: Develop an ethical and legal analysis of RAG technology for use in the health sector. Cross-Cultural Applications: Develop understanding of the accessibility and effectiveness of RAG technology across cultural contexts and health systems.

Conclusion

RAG technology has great promise to improve patient information forms and advance patient communication processes in health care. In adjusting the accuracy, access, and personalization of patient information, RAG technology can support overall improvement in quality patient care. This provides tremendous opportunity for health care. However, for RAG technology to be successfully used in health care, it will be successfully dependent on addressing several challenges to improve patient information forms and the patient educational processes in health care systems, including privacy and security of patient information, legal and ethical implications, adjustment of health care providers to the technology, and use of the technology in the health care system. To consider ways to resolve and address these issues will take ongoing collaboration between health care systems, technology developer(s), regulatory bodies, and researcher(s).

Ultimately, though, the use of RAG technology provides a viable opportunity to more effectively and efficiently implement patient informational forms and improve patient educational processes in the health care sector, though further research and development would be warranted to achieve effect and implementation ability.

Declaration of Interest: The author declares that there is no conflict of interest.

Author Contributions: Murat Genç: Conceptualization, Methodology, Literature Re-

view, Formal Analysis, Writing – Original Draft, Writing – Review & Editing, Visualization, Project Administration.

REFERENCES

[1] Sağlık Bakanlığı Tedavi Hizmetleri Genel Müdürlüğü, "Yataklı Tedavi Kurumları Tıbbi Kayıt ve Arşiv Hizmetleri Yönergesi," 2017.

[2] E. Bayro, "Hasta kabulünde çevrim içi formların kullanılması: Artı ve eksi yönleri," forms.app Blog, 2024.

[3] Dergipark, "Hasta - Hekim İletişimindeki Eksiklik ve Hataların Değerlendirilmesi," Dergipark.org.tr, 2022.

[4] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, et al., "Retrieval-augmented generation for knowledge-intensive NLP tasks," Advances in Neural Information Processing Systems, vol. 33, pp. 9459–9474, 2020.

[5] HealthTech Magazine, "How Does Retrieval-Augmented Generation (RAG) Support Healthcare AI Initiatives?" 2025.

[6] Sağlık Bakanlığı Tedavi Hizmetleri Genel Müdürlüğü, "Yataklı Tedavi Kurumları Tıbbi Kayıt ve Arşiv Hizmetleri Yönergesi," 2017.

[7] A. Arslan, "Hasta dosyası ve içerisinde bulunması gereken formlar, özellikleri," Ondokuz Mayıs Üniversitesi, 2023.

[8] A. Arslan, "Hasta dosyasının içeriği, sıralanması ve düzeni," Ondokuz Mayıs Üniversitesi, 2023.

[9] Dergipark, "Hastanelerin Tıbbi Kayıtlardan Doğan Sorumlulukları," Dergipark.org.tr, 2022.

[10] Dergipark, "Sağlık Kayıtlarının Yapısı," Dergipark.org.tr, 2022.

[11] Ankara Üniversitesi, "Hasta Dosyası Kaynak," Ankara Üniversitesi Açık Ders Malzemeleri, 2023.

[12] Dergipark, "Hastanelerde Elektronik Bilgi Sistemine Geçişte Mevcut Sorunlar," Dergipark.org.tr, 2022.

[13] Dergipark, "Hasta - Hekim İletişimindeki Eksiklik ve Hataların Değerlendirilmesi," Dergipark.org.tr, 2022.

[14] Dergipark, "Sağlık Bilişimi ve Türkiye'de Hastanelerin Dijitalleşmesi," Dergipark.org.tr, 2022.

[15] E. Bayro, "Hasta kabulünde çevrim içi formların kullanılması: Artı ve eksi yönleri," forms.app Blog, 2024.

[16] S. Vermişli Peker, M. Yavuz Van Giersbergen, and G. Biçersoy, "Sağlık Bilişimi ve Türkiye'de Hastanelerin Dijitalleşmesi," Sağlık Akademisi Kastamonu, vol. 3, no. 3, pp. 228–267, 2018.

[17] P. Lewis, E. Perez, A. Piktus, F. Petroni, V. Karpukhin, N. Goyal, et al., "Retrieval-augmented generation for knowledge-intensive NLP tasks," Advances in Neural Information Processing Systems, vol. 33, pp. 9459–9474, 2020.

[18] NVIDIA, "What Is Retrieval-Augmented Generation aka RAG?" NVIDIA Blog, 2025.

[19] HealthTech Magazine, "How Does Retrieval-Augmented Generation (RAG) Support Healthcare AI Initiatives?" 2025.

[20] AWS, "Building a medical intelligence application with augmented patient data," AWS Prescriptive Guidance, 2025.

[21] NVIDIA, "What Is Retrieval-Augmented Generation aka RAG?" NVIDIA Blog, 2025.

[22] NVIDIA, "What Is Retrieval-Augmented Generation aka RAG?" NVIDIA Blog, 2025.

[23] HealthTech Magazine, "How Does Retrieval-Augmented Generation (RAG) Support Healthcare AI Initiatives?" 2025.

[24] AWS, "Building a medical intelligence application with augmented patient data," AWS Prescriptive Guidance, 2025.

[25] Hatchworks, "RAG for Healthcare," Hatchworks Blog, 2025.

[26] NEJM AI, "RAG in Health Care: A Novel Framework for Improving Information Exchange," 2024.

[27] HealthTech Magazine, "How Does Retrieval-Augmented Generation (RAG) Support Healthcare AI Initiatives?" 2025.

[28] AWS, "Building a medical intelligence application with augmented patient data," AWS Prescriptive Guidance, 2025.

Hatchworks, "RAG for Healthcare," Hatchworks Blog, 2025.

[29] NEJM AI, "RAG in Health Care: A Novel Framework for Improving Information Exchange," 2024.

[30] Techsign, "Sağlık Sektöründe Kimlik Doğrulama - İmza ve Hasta Kayıt Süreçlerinde Dijital Teknoloji Kullanımı," 2024.

[31] E. Bayro, "Hasta kabulünde çevrim içi formların kullanılması: Artı ve eksi yönleri," forms.app Blog, 2024.

[32] HealthTech Magazine, "How Does Retrieval-Augmented Generation (RAG) Support Healthcare AI Initiatives?" 2025.

[33] Selçuk Üniversitesi, "Hasta Bilgilendirilmesi ve Rızasının Alınması Prosedürü," 2022.

[34] Çevre Hastanesi, "Hastanın ve Ailesinin Bilgilendirilmesi ve Onayının Alınması Prosedürü," 2022.

[35] EY, "Dijital Sağlıkta Dünya ve Türkiye Perspektifi Raporu," 2024.

[36] Logo, "Sağlık sektöründe dijital dönüşüm hakkında her şey," Logo Blog, 2020.

[37] Elra Sağlık, "Dijital Sağlık Hizmetlerinde Mahremiyet ve Güvenlik İpuçları," 2024.

INVESTIGATION OF NITI-BASED SHAPE MEMORY ALLOYS DESIGN AND PERFORMANCE USING ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

Ümit Zeybek*

Bergama Vocational School, Dokuz Eylül University, Türkiye

**(Corresponding author) umit.zeybek@deu.edu.tr*

ABSTRACT

In this study, the use of artificial intelligence (AI) and data science in the production, characterization, and applications of NiTi alloys is investigated. In recent years, significant contributions have been made by machine learning and big data analytics to the understanding of phase transformations, prediction of mechanical properties, and analysis of microstructures in NiTi alloys. In this paper, existing work in the literature is reviewed, and the application of artificial intelligence and data science to NiTi alloys is discussed. The potential of ML to accelerate the discovery of new alloy compositions and processing parameters, reduce experimental effort, and enable the development of shape memory alloys (SMA) with desired properties is highlighted by the studies. Furthermore, future research areas and potential developments are discussed.

The difficulty and high cost of identifying the optimum composition through experimental methods have made the use of machine learning (ML) approaches inevitable. By analyzing large data sets, the effects of specific properties and compositions can be predicted through ML methods, thereby allowing the design of new alloys to be accelerated.

Keywords: Shape memory alloys (SMA), NiTi, artificial intelligence (AI), machine learning (ML).

Introduction

Today, new materials are emerging rapidly and their properties have yet to be fully explored or optimized. This constitutes an important area of research. However, despite the advances in materials science, various factors such as the increasing complexity of experimental studies, lengthy characterization processes, and costly experimental methods pose significant constraints to research [1]. Materials discovery through the experimental process can be slow as it often depends on the intuition of individual scientists. However, materials informatics, which seeks to explain structure-property relationships in a physically meaningful and statistically robust way, has become a more data-demanding field with the advent of high-throughput detectors and advanced models. Therefore, it is necessary to speed up the discovery process, identify important material samples and make the process more efficient by reducing the number of candidates in a large combination [2].

Shape memory alloys (SMAs) are smart materials that can undergo phase transformation under the influence of temperature and mechanical loading and are used in many fields [3]. However, understanding and optimizing the complex thermomechanical behavior of these alloys is time consuming and costly using conventional methods [4]. Artificial intelligence and machine learning accelerate this process by modeling the phase transformation mechanisms, mechanical properties and microstructure relationships of SMAs. Artificial intelligence (AI) is a tool available to address the challenges of modeling SMA behavior. Among machine learning techniques, artificial neural networks (ANNs) have been widely used to model and predict the complex material behavior of SMAs [5]. Deep learning (DL) based models have been used to make more accurate predictions by analyzing phase transformations and elastic-plastic deformation processes [6]. Methods such as support vector machines (SVM) and support vector regression (SVR) have been used to optimize the mechanical properties of SMAs and determine temperature-mechanical loading relationships [7].

Purpose and Scope of the Study

The aim of this study is to investigate how artificial intelligence (AI) and data science are used in the fabrication, characterization and applications of NiTi alloys. In recent years, the impact of AI-based methods in the field of materials science has increased and these technologies have been used as a supportive and predictive tool for experimental processes. Thus, in this study, the performance of approaches such as machine learning, deep learning and big data analytics in phase transformations, mechanical property prediction and microstructure analysis of NiTi alloys is evaluated and analyzed in the light of the existing literature. In addition, the application of different AI and data science techniques to NiTi alloys is discussed and current developments and future research potential in this field are discussed. By presenting the advantages and limitations of AI-assisted materials engineering approaches, the study aims to provide a comprehensive perspective towards understanding the integration of these technologies into materials development processes.

General Properties of NiTi Alloys

Shape memory alloys are one of the most important functional materials. They have specific and unique properties such as shape memory effect and superelasticity. These properties are related to the crystallographically reversible nature of the martensitic transformation observed in shape memory alloys [8]. Shape memory effects are observed in many metallic alloys, but only some of them are of commercial interest. These alloys can be classified into three main groups: NiTi based, Cu based and Fe based alloys [9]. Shape memory alloys consist of two stable phases with different crystal structures and properties. The stable phase at high temperature has a face-centered cubic crystal structure and is called austenite or the main phase. At low temperature, the martensite phase with low symmetry appears and may have tetragonal, orthorhombic or monoclinic crystal structures.

Although the martensitic transformation is a first-order phase change, it does not occur at a single temperature, but over a temperature range that differs for each alloy system. This transformation mechanism is characterized by four different temperature parameters:

martensite starting temperature (Ms), martensite ending temperature (Mf), austenite starting temperature (As) and austenite ending temperature (Af). Phase transformations can occur due to stress as well as temperature changes. The martensite phase can also be obtained by applying stress to the sample at temperatures above the Af temperature [10-11].

An important property for shape memory alloys is the difference between the temperatures of phase transformations during heating and cooling of the alloy, called hysteresis [12]. Hysteresis is particularly important in actuator applications. The hysteresis behavior of shape memory alloy actuators, which is related to the shape memory effect, is characterized by multiple outputs (e.g. strain or elongation) for a given input (e.g. temperature) [13].

The addition of Hf, Zr, Pd, Pt and Au to NiTi alloy is known to increase the transformation temperatures of the alloy. These are called high temperature shape memory alloys [14]. Ternary Ti-Ni-X alloys have been the most promising commercial high-temperature shape memory alloys, which not only exhibit higher transformation temperatures but also exhibit good mechanical properties and retain strain recovery properties [15].

Alignment of Artificial Intelligence and Data Science with Materials Science

Artificial intelligence and data science are enabling significant innovations in materials science through big data analysis and modeling approaches. This integration is applied in a wide range of areas, from material design to the manufacturing process. We will list them as follows;

- Material Discovery and Design: Machine learning (ML) algorithms are used for prediction of material compositions and new material discovery. New NiTi alloy compositions can be rapidly designed with AI-supported high-pass screening methods.

- Phase Transformations and Microstructure Analysis: Machine learning models can predict phase transformations by analyzing experimental data such as DSC, XRD. Image processing techniques can automatically identify grain structure and precipitation zones from SEM and TEM images.

- Mechanical Property Prediction and Modeling: Regression models and neural networks can predict mechanical properties of NiTi alloys such as hardness, flexibility, fatigue strength. Data-driven approaches help to optimize shape memory effect and superelasticity properties.

- Production Process Optimization: Genetic algorithms and optimization techniques are used to optimize manufacturing processes such as powder metallurgy, heat treatment and sintering. Analysis of process data helps to reduce error rates in production stages.

Review of Existing Studies in the Literature

Sources continue to show that artificial intelligence and data science are playing a variety of important roles in materials science, enabling significant advances in areas such as materials design, discovery and prediction of properties. Recent studies highlight the diversity and effectiveness of machine learning applications, especially in shape memory alloys (SMAs).

Eyercioglu et al. [16] used artificial neural networks (ANNs) to predict the martensite (Ms) and austenite (As) initial temperatures of Fe-based shape memory alloys (SMAs). In the study, an ANN model using a back-propagation algorithm predicted the transformation temperatures using experimental data of 85 different Fe-based SMAs available in the literature. The input parameters include the weight percentages of seven different elements (Fe, Mn, Si, Ni, Ni, Cr, Cu, and Al) along with various processing conditions (hot rolling, homogenization temperature, and quenching). The ANN model exhibited a high prediction accuracy, achieving high correlation coefficients of 0.9996 on both training and test data sets. Furthermore, the accuracy of the model was evaluated using statistical metrics such as mean percentage error and very low error rates were recorded. This demonstrates the effectiveness of ANNs in temperature prediction and allows the determination of the required compositions prior to laboratory experiments. In conclusion, artificial neural networks have emerged as a useful tool to speed up SMA production processes and reduce costs. These findings open the door for future studies to investigate the applicability of such methods for different alloy types.

In another study [17], artificial neural network (ANN) models were developed to predict the behavior of 55NiTi (%wt 55% Ni and 45% Ti) shape memory alloy under stress and thermal cycling. The authors first identified the key factors affecting the cyclic behavior of the alloy and used these parameters as input variables. Experimental data of 55NiTi alloy under isobaric loading conditions, various stress levels and thermal cycles were collected. In order to improve the accuracy of the ANN model, about 37% of the available data was used to train the model, while the rest was used to validate and test the model. The predictions of the obtained ANN models showed a high accuracy rate compared to the experimental data, which proves the validity and application potential of the developed models. The study also highlights that the cyclic behavior of the alloy is highly sensitive to different stress magnitudes (such as 50, 80, 150 and 300 MPa). It was observed that ANN models are more effective than traditional regression methods, especially in capturing complex relationships. It is emphasized that these ANN structures developed are not only limited to the current data, but may guide future research under different thermomechanical conditions or with other alloy types. It is also emphasized that by showing how predictions can be made based on experimental data to determine the optimum behavior of shape memory alloys, research in this field can progress faster and more effectively.

The martensitic transformation temperature of shape memory alloys is one of the most important parameters determining the application areas of these materials and its accurate prediction is of great importance. In a study conducted in this context, machine learning (ML) methods were used to quickly and effectively predict the specific phase transition temperatures of Cu-based shape memory alloys (SMA) [18]. Alloy composition, heat treatment process, phase transition temperatures test conditions and martensitic phase transition temperatures were used as inputs. Inverse design method was applied for the design of target features. Phase transition temperatures were predicted by component modeling and feature modeling methods. Support vector regression (SVR) model, one of the machine learning models, was used and verified with experimental results. The results show that the target materials can be obtained more efficiently and accurately.

Liu et al. [19] investigated the modeling of composition-process-property relationships between thermal hysteresis and average transformation temperatures of shape memory alloys (SMAs) undergoing different production stages (production method, heat treatments and solution type). Support Vector Regression (SVR), Random Forest (RF), Gaussian Process Regression (GPR) models were trained. They used Cross-validation (CV) to reduce the margin of error. As a result, SVR, RF and GPR models were tested within the scope of the study and Gaussian Process Regression (GPR) was found to be more suitable due to its prediction accuracy and uncertainty estimation capability.

Thermal hysteresis and wide transformation temperature ranges of NiTi alloys are among the major issues limiting the efficiency and functionality of these materials. This makes the design process complex and inefficient, leading to more experimentation, time and resources spent to explore the desired material properties. In order to facilitate the studies, Trehern et al. [20] developed an AI-driven material discovery framework called AIMS to help identify new material compositions and processing methods by processing large amounts of material data. Ni32Ti47Cu21 was developed based on NiTi, defined as having a narrower thermal hysteresis and transformation temperature range. They fabricated and experimentally characterized this alloy. As a result, various AI models such as XGBoost, Random Forest and Deep Neural Network were used in the study. They reported that the Deep Neural Network is the best model in the field, showing the highest accuracy and performance.

Hysteresis plays a very important role in high temperature shape memory alloys because the performance of these materials depends on how stable the transformation temperatures are and how efficiently the transformation processes take place, depending on temperature changes. In the design of shape memory alloys, the hysteresis value is considered a critical parameter for performance optimization [21]. At this point, Machaka et al. [22] used the technique of explainable machine learning (XML) for predicting the thermal hysteresis behavior of high temperature shape memory alloys (HT-SMAs) (which is considered to have higher prediction accuracy). The alloy composition (%at) X = Ti, Ni, Pd, Pt, V, Hf, Zr, Cu, Co and Au was used as the number of elements. The main objective of the research is to develop a model that predicts the narrow thermal hysteresis behavior for Ti-Ni based HT-SMAs. Various machine learning (ML) and artificial intelligence (AI) tools were used in the study. Although the aim was not to compare the performance of these different regressors, it was reported that the highest accuracy (89.3%) was obtained with the XGBoost (Extreme Gradient Boosting) model. In conclusion, the research has proven the effectiveness of machine learning methods for predicting narrow thermal hysteresis behavior.

Recently, as a result of the developments in aerospace science, studies on NiTiHf alloys, which are popular in this field, have intensified.

Abedi et al. [23] used machine learning models (SVM, DT, RF, NN) and optimization methods such as GridSearchCV to analyze different factors affecting actuation strain (AS) in various NiTiHfX alloys. They considered NiTiHf, NiTiHfCu, NiTiHfZr, NiTiHfNb, and NiTiHfPd alloys, which constitute the majority of the literature data. The researchers collected a

large data set to predict the AS of the NiTiHfX alloy. Input parameters include the elemental composition of the alloy, fabrication methods used, heat treatments, mechanical stress treatments and post-treatments. This study involved the application of various machine learning models. The researchers used many different ML techniques, including linear models (linear regression (LR), Lasso regression (LAR), Ridge regression (RR)) and non-linear models (random forest (RF), decision trees (DT) (support vector machines (SVM), artificial neural networks (NN)). The study shows that actuation stress (AS) can be used as the main design parameter for shape memory alloys and estimating this parameter can make alloy design processes more efficient. The results obtained in the paper have proven to provide a high rapidity (R^2=0.96) and reliability in predicting the AS of NiTiHfX alloy. They reported that linear models such as linear regression, Ridge regression and Losso Regression are not suitable for such complex data, while NN, RF, SVM and DT models show relatively similar performances. They explained that the developed models provide an effective tool to optimize the design of NiTiHfX alloys and predict actuation stress more accurately.

Again Abedi et al. [24] created a comprehensive dataset to understand how the properties of NiTiHf alloy are affected by various factors (chemical composition, heat treatments and manufacturing methods). They reported that the Neural Network (NN) model they developed was validated by comparing the transformation temperatures (Mf, Ms, As, Af, Mp, Ap) of NiTiHf alloy with previous experimental studies. They also reported that the model was consistent with data obtained from different laboratories with similar chemical and processing parameters.

Raji et al. [25] used three types of ML algorithms, namely Random Forest (RF), Extreme Gradient Boost (XGB), K-Nearest Neighbors (KNN), Random Forest (RF), Extreme Gradient Boost (XGB) and K-Nearest Neighbors (KNN) to predict the austenite end temperature (Af) of NiTiHfPd shape memory alloys that perform well at high temperature and in damping systems. The XGB algorithm showed the best performance because it emphasized chemical composition and heat treatment data as important variables. The Random Forest algorithm provided an effective prediction by emphasizing the heat treatment characteristics. The KNN algorithm did not achieve the success of the other two algorithms as it showed a lower success for the data set obtained. In order to test the model, the Af temperatures of ten samples, which were not used at any point, were measured using Differential Scanning Calorimetry (DSC) after various aging heat treatments. The results were compared with the Af temperatures predicted by RF and XGB. This comparison helped to evaluate how well the predicted values agree with the experimental results.

Udesh et al. [26] aimed to identify new alloy compositions by using machine learning (ML) techniques in the design of NiTiHf high temperature shape memory alloys. While collecting data, only alloys with similar processing conditions were considered. In the study, seven different ML models were evaluated for important properties such as martensite start temperature (Ms), thermal hysteresis (TH) and work output (WO). Through these models, they aimed to understand the relationship between the properties of NiTiHf alloys and alloy compositions and to propose new compositions. Linear regression model, Polynomial regression model, Support vector regression with linear kernel, Support vector regression with polynomial kernel, Support vector regression with rbf kernel, K-nearest neighboring,

Artificial neural network models were trained. The accuracy of the models was tested with experimental data not used during training. As a result, they reported that the K-nearest neighboring model performed better than the other models.

Results

This study investigates the use of Artificial Intelligence (AI) and Machine Learning (ML) techniques in the design and performance prediction of NiTi based shape memory alloys. The studies have shown that AI and ML methods can improve the accuracy of alloy design and obtain results more efficiently compared to traditional experimental methods. In addition, these methods offer significant advantages in terms of time and cost in predicting alloy properties. The effectiveness of AI and MO methods in predicting certain properties (e.g. transformation temperatures, mechanical properties, etc.) of NiTi based alloys has been emphasized. This highlights the importance of innovative approaches in the design of NiTi alloys.

While AI and ML-based approaches can significantly improve the design process of NiTi alloys, they also face some challenges. The accuracy of these methods depends on the quality of the data used and the accuracy of the model. Therefore, it is critical to obtain data from accurate and reliable sources. These new methods in the design of shape memory alloys have the potential to produce more specialized and high-performance alloys by taking into account parameters that were previously ignored. However, with more data and improved modeling techniques, the potential benefits of research in this area are considerable.

Declaration of Interest: The authors declare that there is no conflict of interest.

Author Contributions: This paper was written entirely by Ümit Zeybek.

REFERENCES

[1] K. Guo, Z. Yang, C.H. Yu, M.J. Buehler, "Artificial intelligence and machine learning in design of mechanical materials", Mater. Horiz., 8, 1153–1172, 2021.

[2] L. Pouchard, Y. Lin, H. Van Dam, "Replicating Machine Learning Experiments in Materials Science," in Parallel Computing: Technology Trends, I. Foster et al. (Eds.), IOS Press, 2020.

[3] K. Otsuka and C.M. Wayman, "Shape memory materials", Cambridge University Press, 1998.

[4] Z. Deng, K. Huang, H. Yin, et al., "Temperature-dependent mechanical properties and elastocaloric effects of multiphase nanocrystalline NiTi alloys", Journal of Alloys and Compounds, 938, 168547, 2023.

[5] S. Chowdhury, A. Ghosh, P. Dey, "Machine learning in shape memory alloy research: A review", Materials & Design, 180, 107950, 2019.

[6] Y. Liu, Y. Zhang, X. Zhang, "Deep learning-based phase transformation modeling in shape memory alloys", Computational Materials Science, 174, 109460, 2020.

[7] W. Zhang, C. Lu, J. Wang, "Application of SVM and SVR in predicting mechanical properties of NiTi-based shape memory alloys", Materials Science and Engineering: A, 712, 210-218, 2018.

[8] L. Öncel, "Şekil hafızalı implant malzemelerinin vücut içi uygulamalara yönelik olarak geliştirilmesi", Doktora Tezi, İstanbul Teknik Üniversitesi, 148s, 2017.

[9] A. Ziolkowski, "Pseudoelasticity of Shape Memory Alloys: Theory and Experimental Studies," Elsevier, 246 p., 2015.

[10] Ü. Zeybek, "Zr ilave edilmiş NiTi şekil hafızalı alaşımlarının sıcak izostatik presleme ile üretilmesi ve özelliklerinin incelenmesi," Doktora Tezi, Ege Üniversitesi, 2023.

[11] Ü. Zeybek, S. Karaoğlu, "Effects of Mechanical Alloying on Ni-Rich NiTi Shape Memory Alloys Produced by Hot Isostatic Pressing," Russian Journal of Non-Ferrous Metals, 63(3), 336–343, 2022.

[12] D. Xue, R. Yuan, Y. Zhou, D. Xue, T. Lookman, G. Zhang, X. Ding, J. Sun, "Design of high temperature Ti-Pd-Cr shape memory alloys with small thermal hysteresis," Scientific Reports, 6, 1–7, 2016.

[13] H. Luo, Y. Liao, E. Abel, Z. Wang, X. Liu, "Hysteresis behaviour and modeling of SMA actuators," Shape Memory Alloys, 61–80, 2010.

[14] S.M. Saghaian, "Shape memory behavior of single crystal and polycrystalline Ni-Rich NiTiHf high temperature shape," Ph.D. Thesis, University of Kentucky, 179 p., 2015.

[15] X. Yi, H. Wang, B. Sun, K. Sun, W. Gao, J. Liu, X. Meng, W. Cai, L. Zhao, "The crystallization process, microstructure, martensitic transformation and mechanical properties of Ti-Ni-Zr alloy ribbons," Journal of Alloys and Compounds, 778, 542–553, 2019.

[16] Ö. Eyercioglu, E. Kanca, M. Pala, E. Ozbay, "Prediction of martensite and austenite start temperatures of the Fe-based shape memory alloys by artificial neural networks," Journal of Materials Processing Technology, 200, 146–152, 2008.

[17] J.S. Owusu-Danquah, A. Bseiso, S. Allena, "Artificial neural network models to predict the response of 55NiTi shape memory alloy under stress and thermal cycles," Neural Computing and Applications, 34, 3829–3842, 2022.

[18] M. Wu, W. Yong, C. Fu, C. Ma, R. Liu, "Machine learning-assisted efficient design of Cu-based shape memory alloy with specific phase transition temperature," International Journal of Minerals, Metallurgy and Materials, 31(4), 773, 2024.

[19] S. Liu, B.B. Kappes, B. Amin-ahmadi, O. Benafan, X. Zhang, A.P. Stebner, "Physics-informed machine learning for composition–process–property design: Shape memory alloy demonstration," Applied Materials Today, 22, 100898, 2021.

[20] W. Trehern, R. Ortiz-Ayala, K.C. Atli, R. Arroyave, I. Karaman, "Data-driven shape memory alloy discovery using Artificial Intelligence Materials Selection (AIMS) framework," Acta Materialia, 228, 117751, 2022.

[21] Y. Liu and Y. Li, "Influence of thermal hysteresis on the functional properties of NiTi shape memory alloys," Smart Materials and Structures, 22(3), 035003, 2013.

[22] R. Machaka and P.M. Radingoana, "Prediction of narrow HT-SMA thermal hysteresis behaviour using explainable machine learning," Materials Today Communications, 35, 105806, 2023.

[23] H. Abedi, M.J. Abdollahzadeh, T. Bush, O. Benafan, A. Qattawi, M. Elahinia, "Predicting actuation strain in quaternary shape memory alloy NiTiHfX using machine learning," Computational Materials Science, 246, 113345, 2025.

[24] H. Abedi, K.S. Baghbaderani, A. Alafaghani, M. Nematollahi, F. Kordizadeh, M.M. Attallah, A. Qattawi, M. Elahinia, "Neural network modeling of NiTiHf shape memory alloy transformation temperatures," JMEPEG, 31, 10258–10270, 2022.

[25] H. Raji, M. Rad, E. Acar, H. Karaca, S. Saedi, "A machine learning approach to predict austenite finish temperature in quaternary NiTiHfPd SMAs," Materials Today Communications, 38, 107847, 2024.

[26] U.M.H.U. Kankanamge, J. Reiner, X. Ma, S. Corujeira Gallo, W. Xu, "Machine learning guided alloy design of high-temperature NiTiHf shape memory alloys," Journal of Materials Science, 57, 19447–19465, 2022.

LANGUAGE MODEL APPROACH IN BLOCKCHAIN-BASED FORENSIC UNIT MANAGEMENT SYSTEM

Güncel Sarıman*
Muğla Sıtkı Koçman University, Information System Engineering, Muğla, Türkiye
Tuğçe Güzle
Muğla Sıtkı Koçman University, Information System Engineering, Muğla, Türkiye
(Corresponding author) guncelsariman@mu.edu.tr

ABSTRACT

This study aims to provide an innovative solution to the problems of security, transparency and efficiency in justice processes. In traditional justice systems, there are various difficulties in data manipulation, delays in access to information and evidence storage processes.

In this context, the study aims to digitize judicial processes with a more reliable and innovative approach by combining blockchain and artificial intelligence technologies. The developed model addresses processes such as recording all kinds of statements, evidence, defence, etc. of people in the roles of security forces, judges, lawyers, defendants, etc. with a web application in a computer environment and storing these records in a transparent, unalterable way, as well as translating verbal expressions into text with natural language processing methods and summarizing long texts. With smart contracts running on the Sepolia test network, forensic transactions are managed in a decentralized and unchangeable structure, and user verification is provided with Metamask integration. With the support of artificial intelligence, defence and indictments were summarized using big language models, audio recordings were transcribed and stored on the decentralized file storage system (IPFS) and hash information was recorded on the blockchain. The Big Language Models module used in the study was developed to create context-appropriate summaries of long and complex texts. Various language models and summarization approaches were examined in detail and the most successful model was integrated into the software via API. Open source language models Pegasus, MT5, BART and Google's gemini-1.5-flash are compared. In the analyses with 2 different datasets, the results obtained in terms of context compatibility, semantic integrity and accuracy were evaluated. In the text summarization process, the rouge-1 metric in the Gemini datasets was up to 0.77%, while the BertScore metric resulted in 88%.

The study aims to develop an impartial and transparent justice mechanism by providing effective solutions to problems such as data integrity, reliability, fast access to information and evidence management in modern justice systems, and to ensure transparency and decentralization with Blockchain, summarization of long text and audio files with artificial intelligence, and authorization of the two models with a web interface.

Keywords: Blockchain, artificial intelligence, language model, smart contracts.

Introduction

Justice systems are a fundamental building block in ensuring the order and security of societies. Today, an effective and reliable process management in judicial units is of vital importance in solving crimes. The speed of technological developments requires judicial processes to operate faster and more reliably, which necessitates the adaptation of modern technologies. However, traditional justice systems face significant problems in issues such as data security, transparency of processes, data manipulation, evidence storage, fast access to information and management of large files [1]. Existing systems based on centralized databases may be insufficient in ensuring justice due to the risk of manipulation, data loss and complexity of the processes. Witness testimonies are generally considered by legal experts to be of central importance in determining guilt or responsibility in court. Statements taken in forensic units are another important element used as evidence in criminal proceedings. The content and accuracy of the testimony is crucial for the fair and effective conduct of legal proceedings. However, studies have shown that many of the witness documents are later altered to some extent [2]. These problems necessitate solutions based on modern technologies for the reliable administration of justice.

In this study, a web interface has been developed for use in judicial units by using blockchain and artificial intelligence technologies together in order to provide an innovative solution to the problems encountered in justice systems. In this interface, users with six different authorizations (admin, police/gendarmerie, prosecutor, judge, lawyer, defendant) can use the system and the transparency problems at different steps of the judicial units are solved by summarizing the data accumulated with blockchain and converting audio files into text using natural language processing (NLP).

Blockchain technology works as a decentralized, transparent digital ledger that provides the means to store data records. As these records are added to the ledger, they become immutable and fully transparent, guaranteeing integrity and accountability given a biased trust mechanism [3]. Blockchain offers a powerful technological infrastructure to increase the security, transparency and verifiability of justice processes. Thanks to its decentralized structure, data is not stored on a single server, but on many nodes simultaneously. This structure makes it almost impossible to manipulate the data and prevents the system from being dependent on a single point of failure. When the evidence documents in forensic units are analyzed, records should be kept to track which people accessed the evidence and when [4]. In this sense, the immutability of blockchain technology is highly advantageous for recording and storing immutable records related to the examination of evidence. Any modification or access to the records can be quickly detected through blockchain tamper-resistant logs using public-key cryptography [5]. In blockchain, each new transaction is added to a chain of previous transactions and each block is cryptographically linked to the previous block. This feature ensures both the security and immutability of the system. Various examples illustrate how blockchain can be applied to reduce evidence tampering or loss. Common concerns regarding forensic evidence management have been analyzed, along with some considerations to be taken into account when integrating blockchain into the existing workflow of forensic units [6].

Our study was developed using smart contracts running on the Sepolia test network [7]. Smart contracts ensure that the operations defined in forensic processes are performed automatically and securely. IPFS (InterPlanetary File System) is used for decentralized storage of large files [8]. This method provides a secure and sustainable storage of evidence or important documents related to litigation processes.

Within the scope of the study, Big Language Models (LLM) are used to summarize the defence and indictments, so that users can quickly access critical information. Up to 90% accuracy rates were obtained with pegasus [9], mt5 [10], mbart [11] and gemini-1.5-flash [12] language models used in the summarization phase. The transcription of the audio recordings taken during the litigation process with artificial intelligence contributes to a more effective and transparent management of the litigation processes. The reliability and uninterrupted operation of the system is ensured by MongoDB, which is used as a backup mechanism against blockchain interruptions. In case of access problems that may occur in the blockchain network, transactions are temporarily stored on MongoDB and synchronized when blockchain access is restored. This approach prevents the disruption of the process by providing an uninterrupted data flow in the justice system.

The focus of the study is to provide concrete solutions for more reliable, transparent and effective operation of justice mechanisms. In particular, the elimination of critical problems such as data security, speed of access to information and management of large-sized files contributes to the strengthening of impartiality and reliability in justice systems.

In the second part of the study, Literature Review, in the third part, Material and Method, in the fourth part, the findings obtained, in the fifth part, Conclusion and Recommendations are given.

Literature Review

The problem of privacy and integrity of data in forensic units has become a very important issue in recent years. It is also stated in the literature that 40% of the audio and text data changes during the interrogation process [13]. However, although technological infrastructures are used, these systems are still managed by humans and privacy can be abused. In this part of the study, a sectoral analysis of blockchain technology and research on the use of natural language processing has been carried out.

The adoption of blockchain technology has been increasing exponentially in the last 10 years. Since the 2020s, it has been generally recognized that there are many beneficial aspects of blockchain technology, such as increasing the integrity and security of stored data, recording money transactions and provenance for food, art and documents [14,15]. The study on the economic effects of blockchain has also revealed its advantages such as eliminating intermediaries, reducing transaction costs and providing data security. These effects become especially important in financial applications such as international payments and smart contracts [16].

A study examining the adoption of blockchain technology in public services emphasizes that this process is shaped by institutional, market and technical factors. The study states that regulations are important in increasing the applicability of blockchain [17].

A study on the opportunities and challenges of blockchain technology in the energy sector evaluated the potential benefits in areas such as monitoring renewable energy resources and microgrid management, and addressed scalability and regulatory issues [18].

Öztürk and Sülüngür examined the potential of blockchain to provide transparency, security and speed in customs transactions and evaluated the advantages of preventing informality and increasing cost efficiency in tax collection processes. In the context of Turkey, it was stated that these systems can increase transaction security in public institutions [19].

DenizBank developed blockchain-based solutions for digital payment systems and data security in cooperation with BlockchainIST Centre within Bahçeşehir University. With the Blockchain Lab established at Deniz Aquarium, security and speed in banking transactions have been increased [20].

Lusetti, Saslsi and Dallatana propose to solve the problem through an operational hybrid platform that uses a consensus mechanism to transparently record the access history of digital forensic records and prevent unauthorized users from changing it. Blockchain nodes are used in the study. A hybrid approach for secure access to digital evidence is proposed by combining cryptography with blockchain consensus mechanism [21].

Salisu, Filipov and Pene established a framework to explore how OSINT (OSINT, 2022) can be reduced by using the investigation process of data collection, data protection, data processing and data presentation to detect financial crimes on the blockchain. A community-based approach to financial crime investigations is presented in the blockchain, which involves the public and victims contributing to intelligence that can be used by authorities and security agencies to track and trace criminals in the global fight against cybercrime [22].

NLP is an area of technology and process that involves the capacity for human language to be read, analyzed and interpreted by computers. This field of study has a transformative potential to significantly improve forensic investigations [23]. In forensic processes, text and audio-based recordings or transcripts will be involved in most testimonies, regardless of the media in which they were originally recorded. In such recordings, NLP can identify patterns that are not easily detectable by the naked eye [24].

In recent years, forensic approaches have been applied to linguistic analyses of legal issues and the analysis of legal texts [25]. It has been stated that the integration of language models into the deposition recording system will ensure that the documents created during the deposition are more accurate and thus systems such as automatic summarization will behave consistently [26].

In general, literature studies draw attention to the fact that there are some reservations about data security and confidentiality in statement recording systems and that the processes of generating meaningful data from the data accumulated with the development of digital systems are issues that have been addressed in recent years. It has been observed that block chain applications have started to be used in many different sectors.

Material and Methods

In this section, the infrastructure of the system developed for data consistency, confidentiality and large data volume, which is a major problem for judicial units, is described. The developed system combines blockchain and artificial intelligence technologies with an interface. Information about the blockchain infrastructure, decentralized file management systems (IPFS), language model and interface is presented below.

Blockchain

Blockchain is a decentralized data storage and transaction processing technology. A blockchain is defined as a distributed, peer-to-peer data transfer (P2P) orientated, join-only data structure that enumerates collected transactions into 'blocks' and is verified by the network's participants through consensus mechanisms [27]. Each block in the blockchain contains the hash value of the preceding block, and this structure not only ensures the integrity of the chain but also increases its reliability. The decentralized structure of blockchain prevents any authority or central server from having full control over the data. In this way, data becomes highly resistant to manipulation and transparency can be maximized. In traditional data storage systems, data is controlled by a central authority and this creates an environment open to data manipulation. However, in blockchain technology, all data and transactions are distributed equally among all participants connected to the network. This structure allows data and transactions to be verified by nodes outside the control of a central authority [27].

Smart Contracts

Smart contracts ensure that transactions are carried out in a secure, fast and automated manner [28]. They operate on the Ethereum blockchain and can be defined as digital protocols that perform transactions in accordance with certain rules and conditions [29].

MetaMask

A digital wallet tool called Metamask is used to securely connect to blockchain networks and manage transactions. Functioning as both a cryptocurrency wallet and a browser extension, MetaMask simplifies user interaction with Ethereum-based, decentralized applications (dApps) [30]. MetaMask has played an important role in the smooth user adoption of Web3 by acting as a bridge for interacting with blockchain in the browser. MetaMask is becoming more standard in integration toolkits by dApp developers for smart contract activation, consultation and privacy protection via decentralized file management (IPFS) [31].

File Management System (IPFS)

Despite the high security features of blockchain, storing large files directly on the blockchain is costly and inefficient. Therefore, the InterPlanetary File System (IPFS) is used as a decentralized file storage protocol [8]. Unlike the HTTP protocol, IPFS uses content addressing and peer-to-peer file sharing. Files are fragmented, converted to a unique ID and exchanged using other nodes in a network [32]. Figure-1 shows the decentralized network structure.

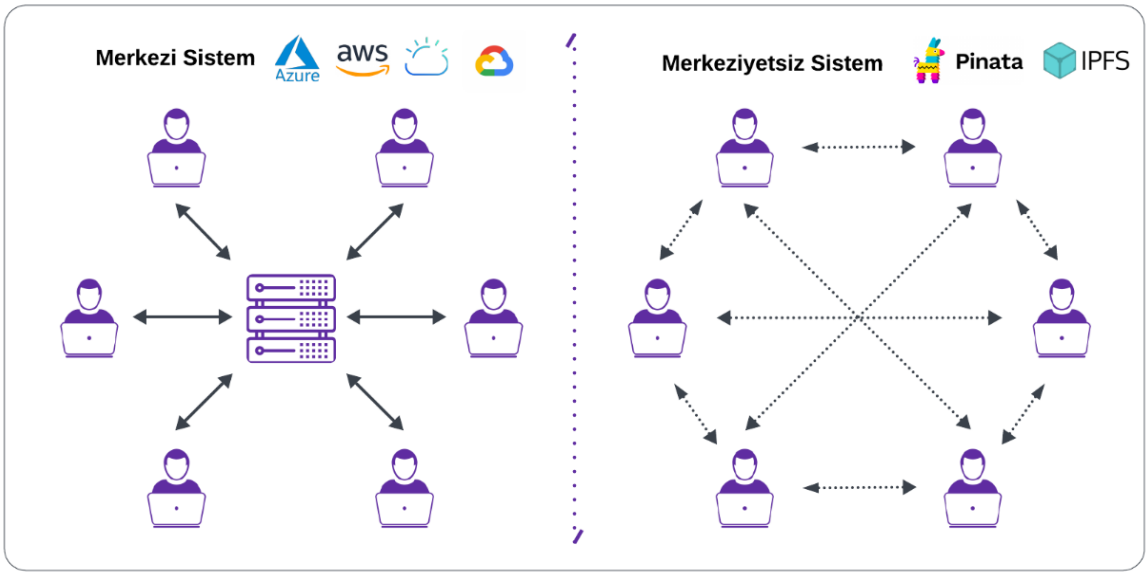


Figure 1: Decentralized Network Structure

Artificial Intelligence Module

While recording various information in forensic units, NLP processes play an important role in ensuring decentralization as well as interpreting the accumulated data in a meaningful way and generating information.

Automatic Speech Module

Automatic Speech Recognition is an artificial intelligence technology developed to convert spoken language into text [33]. Sound waves are first converted into mathematical representations. The process is completed by translating these representations into words and sentences using language models.

Large Language Model

Large Language Models (LLM) are artificial intelligence models designed to understand and interpret human language and produce text using NLP and deep learning technologies [34]. Language models are trained on large amounts of data to learn the context, grammar and semantic relationships of the language. These models have a large vocabulary and understanding of grammar rules. Thanks to these features, they can perform a wide range of NLP tasks such as text summarization, text completion, translation, question answering and sentiment analysis. The main function of language models is to learn the meaning and context relations of human language. These models extract the rules of the language from the texts used as a data set and generate new texts using this information. The basic structure behind the model is usually based on an architecture called 'transformer'. The transformer architecture allows neural networks to understand the relationships between words in texts using attention mechanisms [35].

Developed Model

The project is designed with modern technologies and innovative methods to increase

the reliability of judicial processes and optimize the user experience. The main components used in the development of the system consist of the creation of smart contracts, design of the user interface, blockchain, file management system, backup mechanism and integration of artificial intelligence modules. The general diagram of the model is given in Figure-2.

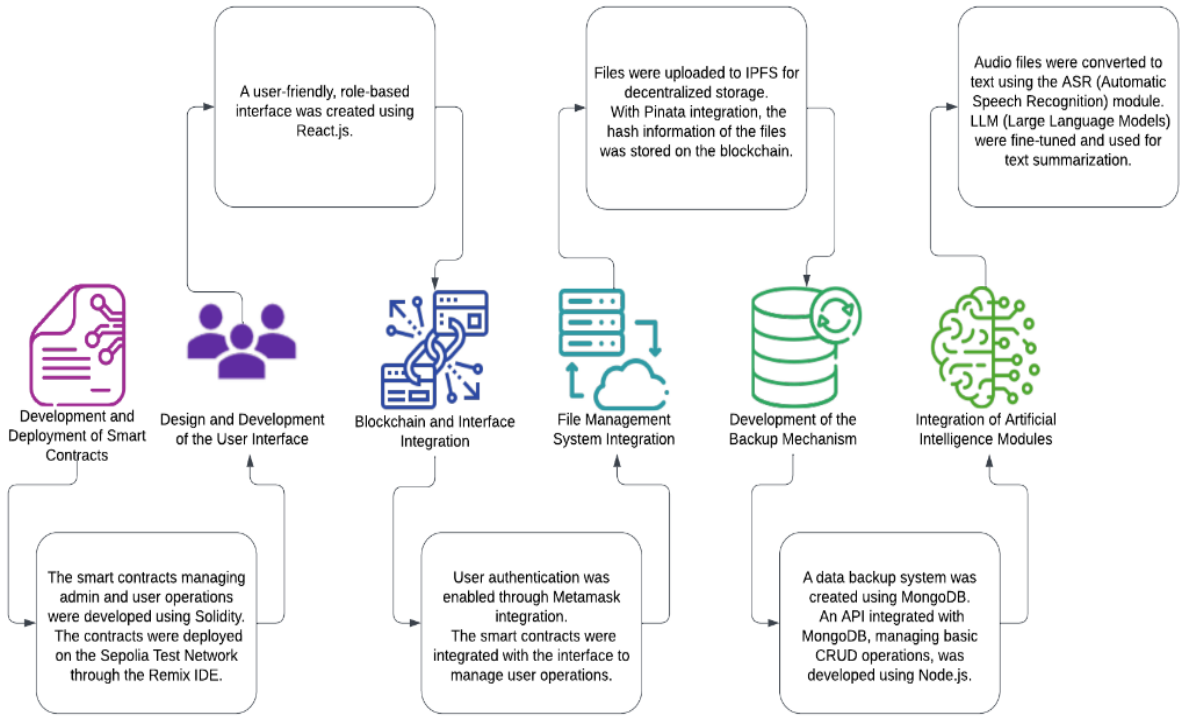


Figure 2: Model Diagram

Web technology was preferred for the user interface. React.js [36], a modern JavaScript library, TailwindCSS [37] for a modern look of the interface, React Charts for graphics and Axios library for communication between backend and frontend were used in the web interface. In order to ensure the security and continuity of user data in case of interruptions in the blockchain network, a MongoDB-based structure was designed for the backup mechanism and a structure for temporary storage of user data. This backup mechanism was used to prevent data loss and ensure the continuity of transactions.

While smart contracts ensure reliable recording of judicial transactions, the interface aims to provide a user-friendly experience. Blockchain technology guarantees that the system operates in a decentralized structure, while the IPFS-based file management system and MongoDB backup mechanism ensure data security and accessibility. With artificial intelligence modules, processes such as converting audio files into text and summarizing long texts have been accelerated and facilitated. Solidity [38] programming language was preferred for the development of smart contracts. Solidity is a programming language used to develop smart contracts running on the Ethereum blockchain. Within the scope of the project, two basic contracts were created. One of them is the administrator contract and the other is the user transaction contract. Figure-3 shows a diagram of the smart contract process.

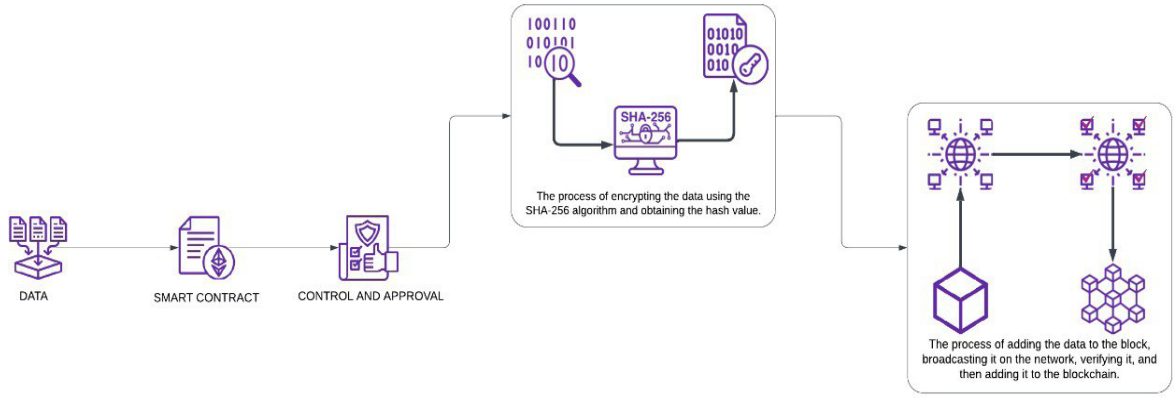


Figure 3: Smart Contract Process

Within the scope of the project, IPFS was used with the decentralised structure of blockchain technology to ensure the secure storage and accessibility of evidential files (e.g. camera recordings, audio files, documents) uploaded by users to the system. Figure-4 shows the IPFS schema of the project.

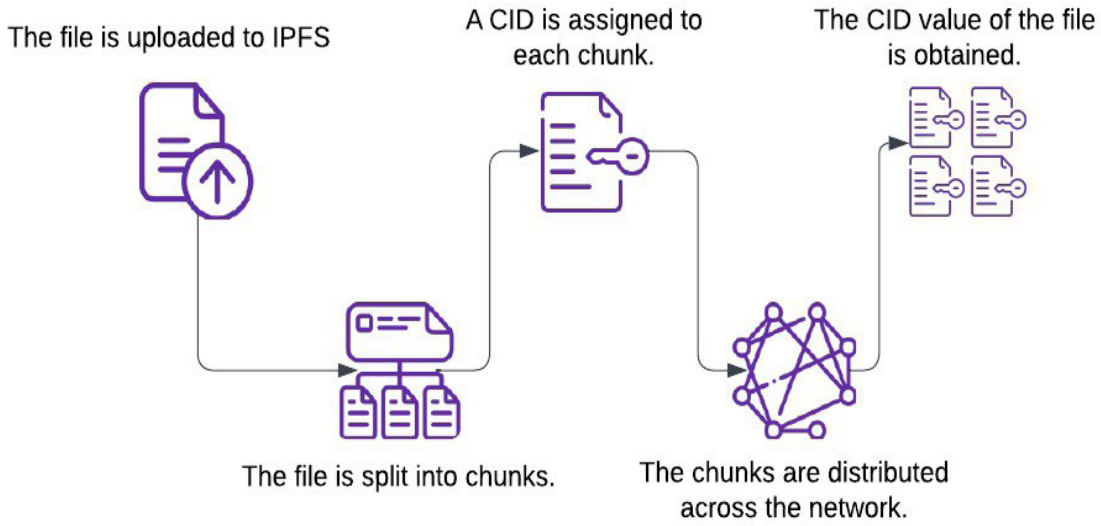


Figure 4: IPFS Diagram of the Project

For the artificial intelligence modules of the project, voice-to-text and Text Summarization methods were run. Google Speech-to-Text API [39] and Whisper [40] models were used to convert audio files into text and WER (Word Error Rate) and CER (Character Error Rate) were used to compare the performance of these models. The language models used for text summarization aim to make documents such as indictment texts used in judicial processes more user-friendly. Different language models and summarization approaches have been considered to enable users to access critical information faster and more accurately. The models used in the development of the text summarization module were evaluated in detail with Rouge, Bleu and BertScore metrics. Figures 5 and 6 show a diagram of the model for transcription and text summarization.

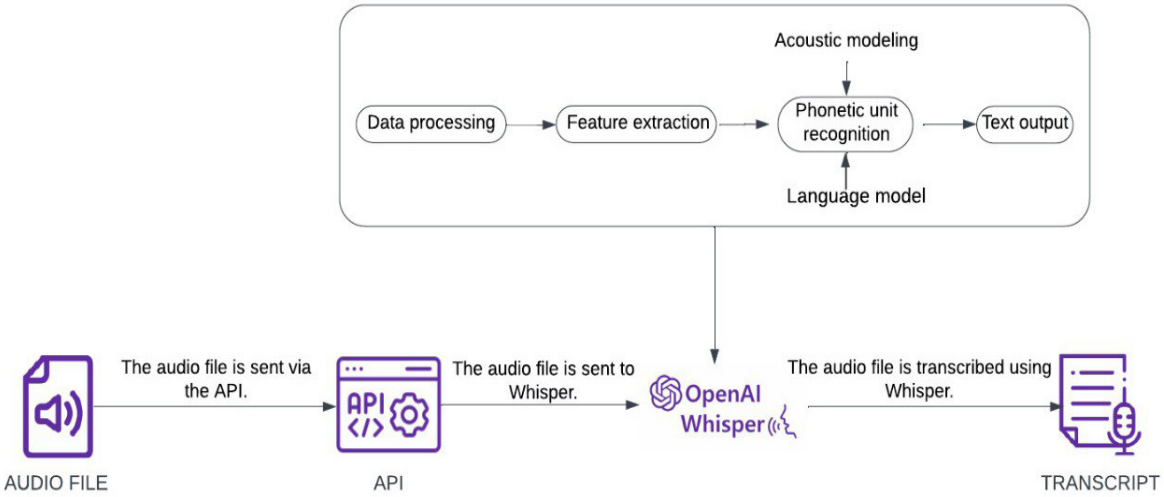


Figure 5: Audio-Text Translator Diagram

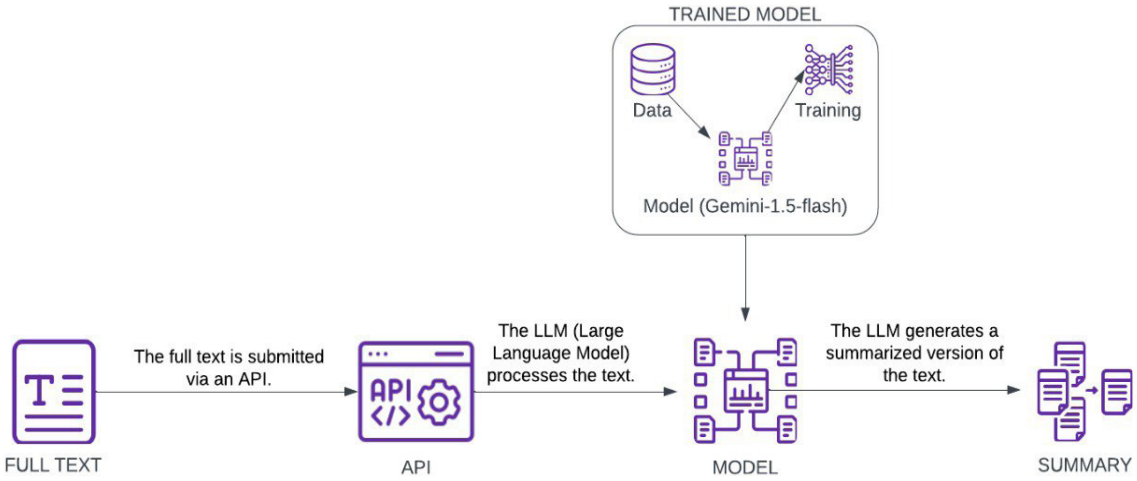


Figure 6: Text Summarization Diagram

The API, which acts as a bridge for the transfer of the output produced in the Voice and Text Summarization models to the user interface, has been written. The API, written using the Flask library, transmits the data received from the user to the artificial intelligence modules and transmits the results obtained to the interface. Audio and text data were processed with the /transcribe and /generate endpoints of the Flask API. The voice and text received from the user were processed in the modules and outputs were generated.

Findings

Within the scope of the study, in the web interface developed for use in judicial units, screens customized according to the roles of the users facilitated the workflow and allowed each user to access only the areas they are authorized to access. In the developed system, there are 6 roles as Citizen, Judge, Prosecutor, Administrator, Lawyer, Police. The dashboards specific to the roles of the users provide statistical information and graphical visualizations about the system. For example, users in the admin role can view data such as the total number of users in the system or active cases through graphs. The graphical

elements and data visualizations used in the design allow users to easily analyze statistics about the system. Figures-7 and 8 list the user interfaces.

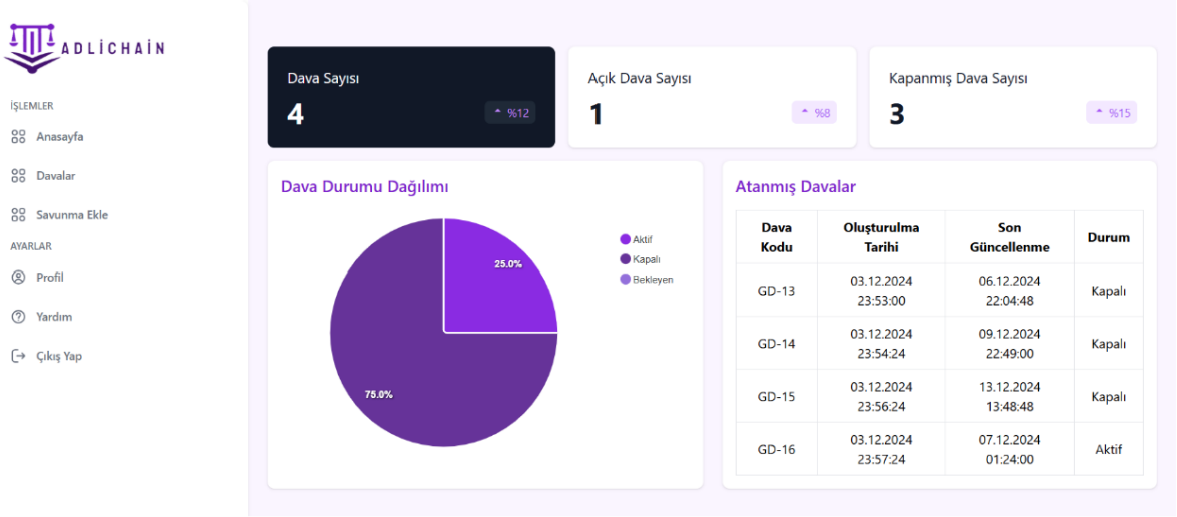


Figure 7: Sample Interface for Lawyer Role

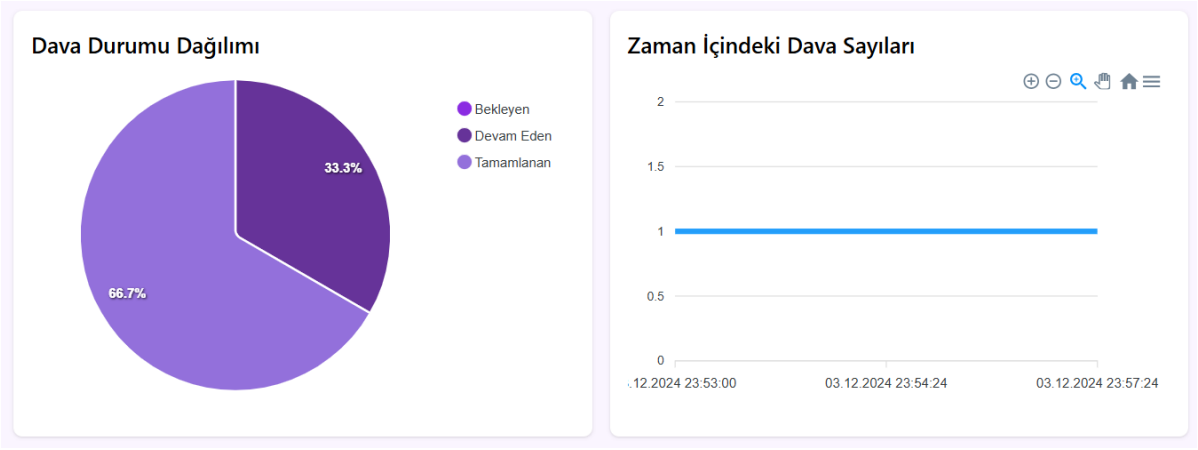


Figure 8: Assigned Cases in Prosecutor Interface

The user interface of the project is integrated with the blockchain network, allowing users to make secure and uninterrupted transactions. This integration is based on the Ethers.js library and the Metamask digital wallet. Ethers.js is a modern JavaScript library used to communicate with the Ethereum blockchain network [41]. In the project, all transactions between the user interface and smart contracts are managed using Ethers.js. Transactions initiated by users through the interface (e.g., filing a lawsuit, uploading evidence) are directed to the relevant smart contracts via Ethers.js. The requests of the users were received from React.js components and calls were made to the contract functions with Ethers.js.

All transactions performed by users in the system were approved through Metamask. When a transaction is initiated, the user can see the transaction details and the gas fee to be paid on Metamask, and after approving the transaction, the transaction is recorded in the blockchain network. Gas fees represent the computational cost required to realise the transaction on the blockchain network. Figure-9 shows the screenshot of Transaction Confirmation and Gas Fee via Metamask.

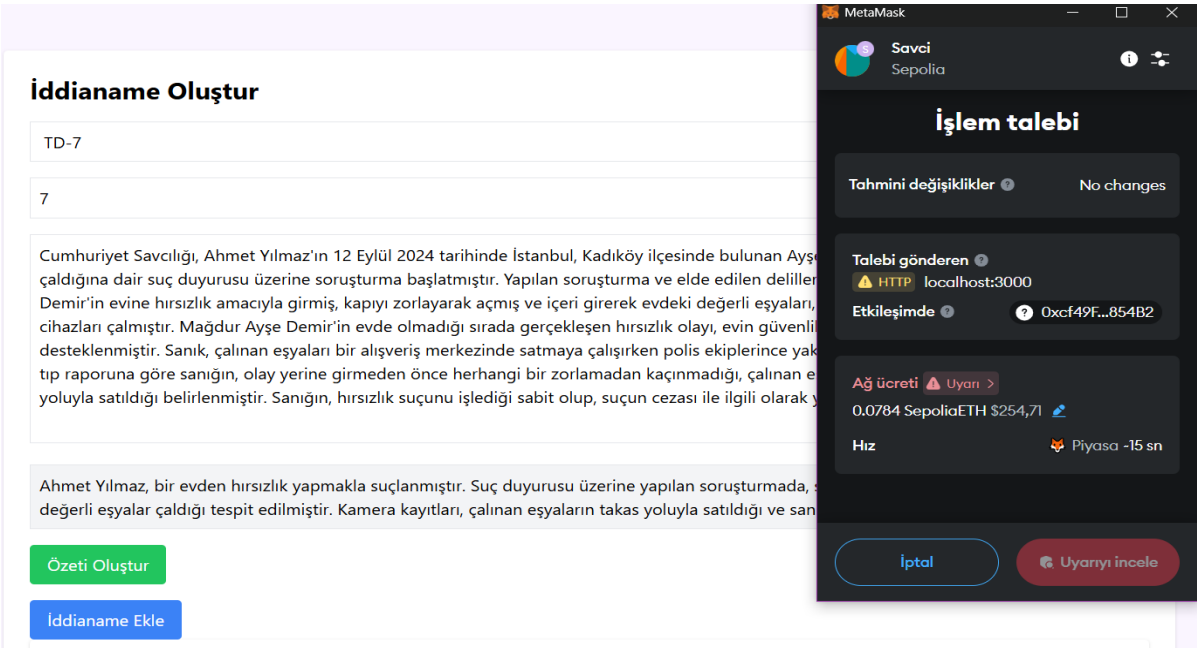


Figure 9: Transaction Confirmation and Gas Fee via Metamask

Blockchain-based System Performance

Within the scope of the project, smart contracts have been developed to enable users to upload evidence, file lawsuits and perform other judicial transactions in a secure and transparent manner. The contracts were deployed on the Ethereum blockchain using the Sepolia Test Network. However, during the development process, several challenges were encountered due to the limitations of smart contracts on the Ethereum blockchain. During the deployment of the smart contracts on the Ethereum blockchain, it was realized that there was a certain bytecode limit. A single comprehensive smart contract that was initially designed could not be deployed because it exceeded this limit. This demonstrated that contract size and complexity were pushing the limits.

The performance of smart contracts is evaluated based on transaction times and gas fees [42]. Tests were conducted with optimized contracts and low gas fees were observed on the Sepolia Test Network. This shows that the optimized code structure is both cost-effective and user-friendly. However, on the real Ethereum network, gas fees can vary significantly depending on factors such as network density and transaction volume.

Smart contracts enable users to conduct transactions in a secure and transparent manner. All transactions are recorded on the blockchain, thus ensuring the verifiability and non-manipulability of evidence. In particular, evidence uploads have become traceable on the blockchain and this process has increased the reliability of forensic evidence.

The bytecode limit problem on the blockchain side was solved by transforming the system into an efficient and modular structure. The use of optimized contracts reduced gas costs and improved performance.

Performance of Voice Recognition and Text Summarization Module

Within the scope of the project, two different solutions, Google Speech-to-Text and Whisper, were used for voice recognition. The performances of these modules were tested with Turkish language support and their processing times were compared based on WER and CER criteria. In the test phase, both modules were run on the same audio recording and the results obtained are analyzed and given in Table 1.

Tablo 1: Whisper ve Google Speech-to-Text Karşılaştırması

Audio File	Method	Duration (s)	WER (%)	CER (%)
Short audio recording	Google Speech-to-Text	2.37	32.14	6.05
Short audio recording	Whisper	5.92	7.14	1.39
Long voice recording	Google Speech-to-Text	9.34	45.28	29.34
Long voice recording	Whisper	19.68	16.98	2.58

The Google Speech-to-Text model provided faster results in terms of processing times. Especially for short audio files, its real-time processing capacity stood out. However, the model made some contextual errors and did not fully adapt to the spelling rules specific to the Turkish language, which slightly reduced its accuracy. However, its speed in processing time is an advantage in scenarios where large data streams need to be processed quickly.

The Whisper model produced more context-appropriate and grammatically meaningful transcriptions. This is a significant advantage, especially in scenarios where context and meaning are critical. However, it required longer processing times compared to Google Speech-to-Text. This delay in processing time can create limitations when working with large datasets. As a result of the comparison, the Whisper model was preferred in a project where contextual relevance is important. It is concluded that both modules can be preferred in different scenarios and can be evaluated depending on the speed and accuracy needs.

Different language models were tested for summarizing long texts and their performance was evaluated. In the testing process, the use of popular models such as Google's "pegasus-xsum" and "mt5-small", Facebook's "bart-large-cnn" were tested on a sample indictment text. The summarization results provided by the models were found to be insufficient in terms of Turkish language structure and contextual appropriateness. Each of these three models obtained the results in Table-2 in text summarization processes where context and meaning are critical, such as indictments.

Table 2: Problems in Open Source Language Models

	Generated Summary	Problems
Pegasus-xsum	"Savc, sann olay srasnda madur Ahmet Kaya'ya ait bir cep telefonu ve nakit para aldn iddia etmektedir."	Türkçe karakterler işlenemedi
mt5-small	"Mehmet Yılmaz Mehmet Yılmaz Mehmet Yılmaz Mehmet Yılmaz"	Anlamsız kelime tekrarları
Bart-large-cnn	"Sanık Mehmet Yılmaz, 20 Ekim 2023 tarihinde Ankara'da gerçekleşen bir hırsızlık olayına."	Devrik ve eksik cümleler

Two different models were obtained by fine-tuning Google's gemini-1.5-flash model with two different datasets using Google AI Studio platform. The model trained with the

"TRT Haber Defense News Dataset" [43] taken from the Kaggle platform produced results that were generally appropriate to the context, but failed to provide consistency in more complex forensic documents. The model trained with the "Custom Dataset Created with the Help of ChatGPT" provided the best results in terms of context compatibility, semantic integrity and accuracy. Table 3 shows the rouge-1, rouge-2, rouge-L, bleu and BertScore of the two models trained with the gemini-1.5-flash model.

Table 3: Test Results of the Model Fine-tuned with Gemini 1.5

Dataset	ROUGE-1	ROUGE-2	ROUGE-L	BLEU	BERTScore
TRT Haber Defense News	0.3473	0.1901	0.2640	9.5117	%62.42
ChatGPT Supported Customized Data Set	0.7718	0.6805	0.7589	56.1260	%88.95

It was determined that the most successful model in terms of context compatibility and semantic integrity in the text summarization task was the model obtained from Google's gemini-1.5-flash model, which was fine-tuned with a special data set. The incomplete and inadequate outputs of the other models showed that these models were not optimized in accordance with the Turkish language structure and failed in context compatibility. Therefore, this model was selected for use in the project and successfully implemented.

File Management System Performance

Instead of storing large files directly on the blockchain, the study uses a decentralized file management system with IPFS and Pinata integration. As the file size increases on the blockchain, there is a significant increase in gas fees and long transaction times. This method played a critical role in preventing the increase in gas costs and optimizing transaction times.

Decentralized storage of files over IPFS increased data security and allowed users to have fast access. Through hash values, file authenticity can be easily checked and data manipulation is prevented. Keeping files in a distributed structure in the IPFS network increased the security of the system and provided a solution that did not require gas fees. Table 4 shows the differences with the use of IPFS.

Table 4: Differences Before and After Using IPFS

Feature	Blockchain (without IPFS)	IPFS & Pinata
Gas Fees	High costs were incurred due to file size.	Free of charge due to the storage of hash information.
Processing Time	Took longer depending on file size.	High speed file upload and access.
Data Security	Decentralized, but experienced difficulties with large files.	Decentralized structure and hash-based accuracy.
User Experience	Limited by high extortion fees and long processing times	User-friendly API and fast access.

By storing only hash information on the blockchain, both transaction time and costs are minimized.

Result and Discussion

This project was developed to speed up judicial processes, increase reliability and create a user-friendly system. In current practices, the processes of taking statements and storing evidence in judicial units are recorded in information systems, but there may be some question marks about the reliability of the data. The fact that the data is distributed and unalterable eliminates this problem. The integrated use of blockchain infrastructure, artificial intelligence and decentralized file management technologies in the system developed in this context has enabled more transparent, secure and efficient management of judicial proceedings.

The blockchain-based system has provided a significant advantage in terms of securely executing and recording transactions. However, due to gas costs and processing time limitations, IPFS and Pinata integration was preferred instead of storing large files directly on the blockchain. This integration reduced costs and made file storage processes more efficient.

Text summarization and Automatic Speech Recognition were also important components of the project. The model, which was trained with a special data set using the fine-tuning method from Google's gemini-1.5-flash model, contributed to the acceleration of forensic processes by producing contextually appropriate and meaningful summaries. Similarly, in Google Speech-to-Text and Whisper voice recognition processes, the processes were optimized by selecting the appropriate module for the project requirements.

Evaluation on Blockchain Based Systems

Tests performed on the Sepolia Testbed for the blockchain-based system showed that the system successfully provided security and verification mechanisms. However, some limitations were encountered in terms of performance and cost. During the development of smart contracts, the contracts had to be optimized due to bytecode limitations. Initially designed as a single contract, the structure was transformed into two separate contracts to overcome these limitations. This structure not only ensured the deployability of the contracts, but also reduced processing times and gas costs. However, since a non-optimized contract could not be deployed, performance comparisons of the different designs could not be made.

In future work, the use of lighter and modular contract designs will enable more efficient management of bytecode boundaries. In conclusion, the blockchain-based system provides a powerful solution for the secure and transparent execution of judicial processes. However, further work in areas such as gas costs, scalability and performance could enable the system to be used more widely and effectively. In the future, with new technologies and design approaches, significant improvements can be made in these areas and the system can serve a wider user base.

Evaluation on Text Summarization and Voice Recognition Systems

In the voice recognition module, Google Speech-to-Text and Whisper models were compared. Google Speech-to-Text stood out for its fast processing times but showed

some shortcomings in terms of contextual relevance and accuracy. Whisper, on the other hand, produced more contextually consistent transcripts, but its long processing times created limitations in big data processing scenarios. In future work, optimization efforts can be made to improve the processing times of models with high contextual accuracy such as Whisper.

In the language model module, pegasus-xsum, mt5-small, bart-large-cnn and 2 models trained on Google Generative AI were tested. The first three models were found to be inadequate in terms of contextualization and accuracy in the Turkish language. Problems such as Pegasus' inability to process Turkish characters, mT5's context-free results, and Bart's generating meaningless contexts with run-on sentences showed that these models are not sufficient for forensic unit data for summarization. Google's Gemini-1.5-flash model was fine-tuned and tested with special Turkish datasets and obtained better results compared to other models. This model made a significant contribution to the project requirements by being able to produce meaningful summaries of long texts in forensic processes. However, it is recommended that this model should be retrained with larger and more diverse Turkish data sets to improve its performance.

In conclusion, although the artificial intelligence modules used in the project largely meet the current needs, it is possible to make the system more powerful with larger data sets, optimized models and development processes based on user feedback. Future studies should aim to develop solutions especially for big data scenarios and to use artificial intelligence more effectively in forensic processes. In this way, a more advanced level can be achieved in the accuracy and acceleration of forensic processes with artificial intelligence technologies.

Suggestions

In blockchain-based systems, more efficient smart contract designs can be evaluated to reduce gas fees and speed up transaction times. In addition, testing different blockchain networks (e.g. Polygon or Avalanche) can increase the cost-effectiveness of the system. It can also be suggested to develop modular and flexible structures for smart contracts to support more complex transactions. Scalability-focused developments will be critical, especially if the system is used by larger user groups.

Alternative model experiments can be conducted to make artificial intelligence modules more efficient. Training powerful models, such as Google's gemini-1.5-flash, which is used especially in text summarization processes, with larger and more diverse data sets can increase performance. In addition, the functionality of the modules can be increased by developing customized summarization capabilities according to user needs. It is aimed to develop more optimized models to increase transaction speeds in artificial intelligence modules.

Reducing file sizes and applying more efficient data compression algorithms can optimize storage and access times for more effective use of decentralized file management systems such as IPFS and Pinata. In addition, network performance can be increased by contributing to the IPFS network to have a larger node structure. Additional encryption

mechanisms and access control systems can be integrated to reduce security risks associated with decentralized solutions. Future work will further improve the reliability, performance, and user experience of the system in line with these proposed improvements.

Declaration of Interest: The authors declare that there is no conflict of interest.

Author Contributions: G.S. developed the research idea and designed the work to be done in the study. T.G. managed the literature review, data collection and analysis processes. G.S. and T.G. performed the statistical analyses together and interpreted the results. T.G. wrote the first draft of the article. G.S. edited the final version of the article and translated it into English. All authors reviewed the article, discussed the content and contributed to the final version.

REFERENCES

[1] F. Ünal and İ. Kiraz, "Türkiye'de e-Devlet uygulamalarının kamu hizmetlerinin sunumunda etkinliği: Adalet Bakanlığı UYAP Bilişim Sistemi örneği," Dumlupınar Üniversitesi Sosyal Bilimler Dergisi, Afro-Avrasya Özel Sayısı, pp. 437–450, 2016.

[2] M. J. R. Hall, "The Relationship between Research and Casework in Forensic Entomology", *Insects*, vol. 12, no. 2, p. 174, Feb. 2021. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.3390/insects12020174.

[3] W. T. Weilbach, "Practical application of distributed ledger technology in support of digital evidence integrity verification processes", thesis, Rhodes Univ., 2018. Accessed: Mar. 18, 2025. [Online]. Available: http://hdl.handle.net/10962/61872

[4] A. J. Akbarfam, M. Heidaripour, H. Maleki, G. Dorai, and G. Agrawal, "ForensiBlock: A Provenance-Driven Blockchain Framework for Data Forensics and Auditability", in *2023 5th IEEE Int. Conf. Trust, Privacy Secur. Intell. Syst. Appl. (TPS-ISA)*, Atlanta, GA, USA, Nov. 1–4, 2023. IEEE, 2023. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.1109/tps-isa58951.2023.00025

[5] A. Sangekar, V. Gumalwad, S. Khillare, K. Warbhog, S. Hadbe, and U. Tupe, "Blockchain-Powered Forensic Evidence Storage: A New Approach to Data Security", in *2024 IEEE Pune Sect. Int. Conf. (PuneCon)*, Pune, India, Dec. 13–15, 2024. IEEE, 2024, pp. 1–6. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.1109/punecon63413.2024.10895214

[6] O. S. Igonor, M. B. Amin, and S. Garg, "The Application of Blockchain Technology in the Field of Digital Forensics: A Literature Review", *Blockchains*, vol. 3, no. 1, p. 5, Feb. 2025. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.3390/blockchains3010005

[7] Ethereum Foundation, "Sepolia Testnet," Ethereum. [Online]. Available: https://ethereum.org/en/developers/docs/networks/#sepolia. [Accessed: Jan. 20, 2025].

[8] C. Bieri, "An Overview into the InterPlanetary File System (IPFS): Use Cases, Advantages, and Drawbacks," in *Communication Systems XIV*, University of Zurich, Zurich, Switzerland, 2021, p. 78.

[9] J. Zhang, Y. Zhao, M. Saleh, and P. J. Liu, "PEGASUS: Pre-training with extracted gap-sentences for ABSTRACTive summarization," arXiv preprint arXiv:1912.08777, 2019. [Online]. Available: https://arxiv.org/abs/1912.08777.

[10] L. Xue et al., "mT5: A massively multilingual pre-trained text-to-text transformer," arXiv preprint arXiv:2010.11934, 2020. [Online]. Available: https://arxiv.org/abs/2010.11934.

[11] M. Lewis *et al.*, "BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension", in *Proc. 58th Annu. Meeting Assoc. Comput. Linguistics*, Online. Stroudsburg, PA, USA: Assoc. Comput. Linguistics, 2020. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.18653/v1/2020.acl-main.703

[12] R. Anil et al., "Gemini: A family of highly capable multimodal models," arXiv preprint arXiv:2312.11805, 2023. [Online]. Available: https://arxiv.org/abs/2312.11805.

[13] K. Mukattash, "The role of forensic phonetics in legal investigation: A case study of two speaker-identified/unidentified recorded samples," [Online]. Available: https://iiste.org/Journals/index.php/JLLL/article/view/34706

[14] S. Dong, K. Abbas, M. Li, and J. Kamruzzaman, "Blockchain technology and application: an overview", *PeerJ Comput. Sci.*, vol. 9, Nov. 2023, Art. no. e1705. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.7717/peerj-cs.1705

[15] G. Habib, S. Sharma, S. Ibrahim, I. Ahmad, S. Qureshi, and M. Ishfaq, "Blockchain Technology: Benefits, Challenges, Applications, and Integration of Blockchain Technology with Cloud Computing", *Future Internet*, vol. 14, no. 11, p. 341, Nov. 2022. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.3390/fi14110341

[16] C. Catalini and J. S. Gans, "Some Simple Economics of the Blockchain", *SSRN Electron. J.*, 2016. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.2139/ssrn.2874598.

[17] M. Janssen, V. Weerakkody, E. Ismagilova, U. Sivarajah, and Z. Irani, "A framework for analysing blockchain technology adoption: Integrating institutional, market and technical factors", *Int. J. Inf. Manage.*, vol. 50, pp. 302–309, Feb. 2020. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.1016/j.ijinfomgt.2019.08.012

[18] M. Andoni *et al.*, "Blockchain technology in the energy sector: A systematic review of challenges and opportunities", *Renewable Sustain. Energy Rev.*, vol. 100, pp. 143–174, Feb. 2019. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.1016/j.rser.2018.10.014

[19] L. Öztürk and E. Sülüngür, "Gümrük Vergilerinde Blockchain Teknolojisinin Kullanılması", *Ekon. Maliye Isletme Derg.*, Jul. 2023. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.46737/emid.1212841

[20] BThaber, "Blockchain'de geleceğin projeleri Deniz Akvaryum ve BAU BlockchainIST Center iş birliği ile tasarlanacak," BThaber, no. 91, Dec. 1, 2024. [Online]. Available: https://www.bthaber.com/blockchainde-gelecegin-projeleri-deniz-akvaryum-ve-bau-blockchainist-center-is-birligi-ile-tasarlanacak/. [Accessed: Dec. 1, 2024].

[21] M. Lusetti, L. Salsi, and A. Dallatana, "A blockchain based solution for the custody of digital files in forensic medicine," *Forensic Science International: Digital Investigation*, vol. 35, p. 301017, 2020. [Online]. Available: https://doi.org/10.1016/j.fsidi.2020.301017.

[22] S. Salisu and V. Filipov, "Blockchain Forensics: A Modern Approach to Investigating Cybercrime in the Age of Decentralisation," *International Conference on Cyber Warfare and Security*, vol. 18, no. 1, pp. 338–347, 2023. [Online]. Available: https://doi.org/10.34190/iccws.18.1.947.

[23] D. Khurana, A. Koli, K. Khatter, and S. Singh, "Natural language processing: state of the art, current trends and challenges", *Multimedia Tools Appl.*, Jul. 2022. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.1007/s11042-022-13428-4

[24] A. Bhatnagar, A. L. Kekatpure, V. R. Velagala, and A. Kekatpure, "A Review on the Use of Artificial Intelligence in Fracture Detection", *Cureus*, Apr. 2024. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.7759/cureus.58364

[25] S. K. Maitlo, I. A. Abbasi, Z. A. Jatoi, and A. Ahmad, "Quantifying Awareness and Attitudes Towards Forensic Linguistics: A Survey of Young Legal Professionals," *Policy Journal of Social Science Review*, vol. 3, no. 2, pp. 12–22, 2025. [Online]. Available: https://journalofsocialsciencereview.com/index.php/PJSSR/article/view/110

[26] J. Wu, L. Cheng, and X. Zhao, "Exploring the efficacy of explicit encoding and fast mapping in second language vocabulary acquisition among Chinese graduate students", *Humanities Social Sci. Commun.*, vol. 12, no. 1, Feb. 2025. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.1057/s41599-025-04431-9

[27] B. Bellaj, A. Ouaddah, E. Bertin, N. Crespi, and A. Mezrioui, "Drawing the Boundaries Between Blockchain and Blockchain-Like Systems: A Comprehensive Survey on Distributed Ledger Technologies", *Proc. IEEE*, vol. 112, no. 3, pp. 247–299, Mar. 2024. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.1109/jproc.2024.3386257

[28] K. Werbach and N. Cornell, "Contracts: Ex Machina", in *Smart Contracts*. Hart Publ., 2021. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.5040/9781509937059.ch-001

[29] P. V. R. P. Raj, S. K. Jauhar, M. Ramkumar, and S. Pratap, "Procurement, traceability and advance cash credit payment transactions in supply chain using blockchain smart contracts", *Comput. & Ind. Eng.*, vol. 167, p. 108038, May 2022. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.1016/j.cie.2022.108038

[30] V. Walunj, V. Rajaraman, J. Dutta, and A. Sharma, "Integrating Crypto-Based Payment Systems for Data Marketplaces: Enhancing Efficiency, Security, and User Autonomy", in *Lecture Notes in Computer Science*. Cham: Springer Nat. Switz., 2024, pp. 443–452. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.1007/978-3-031-80020-7_25

[31] Z. Zhou, T. Sharma, L. Emano, S. Das, and Y. Wang, "Iterative design of an accessible crypto wallet for blind users," in *Proceedings of the 19th USENIX Conference on Usable Privacy and Security (SOUPS '23)*, USENIX Association, USA, 2023, Article 21, pp. 381–398.

[32] J. Anthal, S. Choudhary, and R. Shettiyar, "Decentralizing File Sharing: The Potential of Blockchain and IPFS", in *2023 Int. Conf. Advancement Computation Comput. Technol. (InCACCT)*, Gharuan, India, May 5–6, 2023. IEEE, 2023. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.1109/incacct57535.2023.10141817

[33] V. Bhardwaj, M. T. Ben Othman, V. Kukreja, Y. Belkhier, M. Bajaj, B. S. Goud, A. U. Rehman, M. Shafiq, and H. Hamam**,** "Automatic Speech Recognition (ASR) Systems for Children: A Systematic Literature Review," *Applied Sciences*, vol. 12, no. 9, p. 4419, 2022. [Online]. Available: https://doi.org/10.3390/app12094419

[34] W. Ji, W. Yuan, E. Getzen, K. Cho, M. I. Jordan, S. Mei, et al., "An Overview of Large Language Models for Statisticians," *arXiv preprint arXiv:2502.17814*, 2025. [PDF].

[35] G. Neubig and C. Dyer, "Generalizing and Hybridizing Count-based and Neural Language Models", in *Proc. 2016 Conf. Empirical Methods Natural Lang. Process.*, Austin, Texas. Stroudsburg, PA, USA: Assoc. Comput. Linguistics, 2016. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.18653/v1/d16-1124

[36] Facebook, Inc., "React [Software]," Meta, 2013. [Online]. Available: https://react.dev/.

[37] Tailwind Labs, "Tailwind CSS [Software]," 2017. [Online]. Available: https://tailwindcss.com.

[38] Ethereum Foundation, "Solidity [Programming language]," 2014. [Online]. Available: https://soliditylang.org.

[39] D. Macháček, R. Dabre, and O. Bojar, "Turning Whisper into Real-Time Transcription System", in *Proc. 13th Int. Joint Conf. Natural Lang. Process. 3rd Conf. Asia-Pacific Chapter Assoc. Comput. Linguistics: System Demonstrations*, Bali, Indonesia. Stroudsburg, PA, USA: Assoc. Comput. Linguistics, 2023. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.18653/v1/2023.ijcnlp-demo.3

[40] A. Radford, J. W. Kim, T. Xu, G. Brockman, C. McLeavey, and I. Sutskever, "Robust speech recognition via large-scale weak supervision," in *Proceedings of the 40th International Conference on Machine Learning (ICML'23)*, vol. 202, JMLR.org, 2023, Article 1182, pp. 28492–28518.

[41] R. Koivisto, "Ethers. [Software]," 2016. [Online]. Available: https://ethers.org.

[42] E. Osaba *et al.*, "A Tutorial On the design, experimentation and application of metaheuristic algorithms to real-World optimization problems", *Swarm Evol. Computation*, vol. 64, p. 100888, Jul. 2021. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.1016/j.swevo.2021.100888

[43] D. Şahin, "TRT Haber Savunma News Dataset [Data set]," Kaggle, 2024. [Online]. Available: https://www.kaggle.com/datasets/denizhanahin/trt-haber-savunma-news-dataset.

[44] M. Turki, S. Cheikhrouhou, B. DAMMAK, M. Baklouti, R. Mars, and A. Dhahbi, "NFT-IoT Pharma Chain : IoT Drug traceability system based on Blockchain and Non Fungible Tokens (NFTs)", *J. King Saud Univ. - Comput. Inf. Sci.*, Jan. 2023. Accessed: Mar. 18, 2025. [Online]. Available: https://doi.org/10.1016/j.jksuci.2022.12.016

MAKİNE ÖĞRENMESİ YÖNTEMLERİYLE İZMİR İLİ SU TÜKETİMİNİN İNCELENMESİ VE SU TÜKETİMİNİ TAHMİN ETMEDE ETKİN FAKTÖRLERİN TESPİT EDİLMESİ

Fatih Tek*
İzmir Kâtip Çelebi Üniversitesi, İzmir, Türkiye

Fatma Günseli Yaşar Çıklaçandır
İzmir Kâtip Çelebi Üniversitesi, İzmir, Türkiye
(İletişimden sorumlu yazar) fatihtek35@gmail.com

ÖZET

Çalışmada İzmir şehri için içme suyu tüketimini etkileyen faktörler incelenmiş ve su tüketim miktarını tahmin etmek amacıyla makine öğrenmesi modelleri kullanılmıştır. Çalışma, sürdürülebilir su yönetimi ve altyapı yatırımlarına ilişkin stratejik kararların alınmasında veri tabanlı analizlerle daha etkili bir planlama süreci oluşturulmasına katkı sağlamayı amaçlamaktadır. İzmir Su ve Kanalizasyon İdaresi, Meteoroloji Genel Müdürlüğü, Türkiye İstatistik Kurumu, Çalışma ve Sosyal Güvenlik Bakanlığı ve Merkez Bankası›ndan elde edilen veriler, su tüketimi üzerinde etkili olan faktörleri belirlemek ve en başarılı makine öğrenmesi modelini tespit etmek için kullanılmıştır. Veri seti, Ocak 2003 ile Aralık 2023 arasındaki 252 aylık veriden oluşmaktadır ve bu veriler üzerinde çeşitli özniteliklerin etkililiği değerlendirilmiştir. 5 farklı makine öğrenmesi modeli kullanılarak, özniteliklerin 14 farklı kombinasyonu üzerinde çalışmalar yapılmış ve her modelin performansı, R-kare (R^2) ve Ortalama Mutlak Yüzde Hata (MAPE) metrikleri ile değerlendirilmiştir. Sonuç olarak, AdaBoostRegressor Regresyon Modeli, R^2=0.896 ve MAPE=7.09 değerleri ile en iyi performansı göstermiştir. İkinci en iyi performans ise Rastgele Orman Regresyon modelinden elde edilmiş olup, R^2=0.867 ve MAPE=7.41 değerlerine sahiptir. Bu çalışma, akıllı şehir yönetimi bağlamında su kaynaklarının daha verimli ve sürdürülebilir bir şekilde yönetilmesi adına önemli bir katkı sağlamaktadır.

Anahtar Kelimeler: Makine öğrenmesi, akıllı şehir, içme suyu.

Giriş

Su, hem insanlar hem de gezegenimiz için hayati bir öneme sahiptir ve yaşamın temel kaynaklarından biridir. Meteoroloji Genel Müdürlüğü'nün verilerine göre, Dünya yüzeyinin üçte ikisi su ile kaplıdır. Ancak bu suyun %97,5'i okyanuslardaki tuzlu sudan oluşmaktadır. Geriye kalan %2,5'lik tatlı suyun büyük bir kısmı Antarktika ve Grönland'daki buzullarda ya da derin yer altı katmanlarında depolanmıştır. Doğrudan erişilebilen tatlı su kaynakları ise göller, barajlar, nehirler ve akarsular içinde bulunur ve bu miktar, gezegenimizdeki toplam tatlı su varlığının yalnızca %0,10'unu oluşturmaktadır [1]. Bu nedenle, tatlı su kaynaklarının korunması ve sürdürülebilir kullanımı hayati öneme sahiptir.

Dünya Sağlık Örgütü'nün 2014 yılında yayımladığı Türkiye'nin Su Ayak İzi Raporu'nda, Türkiye'nin su zengini bir ülke olmadığı vurgulanmaktadır. Artan nüfus, sanayileşme ve hızla genişleyen kentleşme nedeniyle, Türkiye'nin 2030 yılı itibarıyla su sıkıntısı yaşayabileceği öngörülmektedir. Bu nedenle, su kaynaklarının sürdürülebilir şekilde kullanılması ve etkin yönetilmesi Türkiye açısından büyük önem taşımaktadır [2].

2024 yılı Birleşmiş Milletler Dünya Su Kalkınma Raporu'nun sonuç bölümünde, iklim değişikliğinin temel insan ihtiyaçları için suyun erişilebilirliğini, kalitesini ve miktarını etkileyeceği belirtilmektedir. Bu durum, milyarlarca insanın su ve sanitasyonla ilgili haklarını etkin bir şekilde kullanmasını tehdit edebilir. Gıda güvenliği, insan sağlığı, kırsal ve kentsel yerleşimler, enerji üretimi, sanayi gelişimi, istihdam, ekonomik büyüme ve ekosistemler suya bağımlıdır ve dolayısıyla iklim değişikliğinin etkilerine karşı savunmasızdır. Bu nedenle, su yönetimi iklim değişikliğine uyum sağlamak ve etkilerini hafifletmek açısından hayati bir rol oynamakta olup, sürdürülebilir kalkınmanın hemen her alanında kritik bir öneme sahiptir [3].

Su yönetimi konusuna geçmeden önce, akıllı şehir kavramını incelemek faydalı olacaktır. Avrupa Parlamentosu, akıllı şehri, çok paydaşlı ve belediye merkezli bir ortaklık temelinde, bilgi ve iletişim teknolojileri (BİT) tabanlı çözümler kullanarak kamu sorunlarını çözmeyi amaçlayan şehir olarak tanımlamaktadır [4]. İngiliz Standartları Enstitüsü'ne göre, akıllı şehir; fiziksel, dijital ve insani sistemlerin yapılandırılmış bir çevreyle etkin entegrasyonu yoluyla ekosistem varlıklarına sürdürülebilir, kapsayıcı ve müreffeh bir gelecek sunmayı amaçlamaktadır [5].Uluslararası Telekomünikasyon Birliği ise akıllı ve sürdürülebilir şehri; mevcut ve gelecek nesillerin ekonomik, sosyal, çevresel ve kültürel gereksinimlerini göz önünde bulundururken, yaşam kalitesini artıran, şehircilik hizmetlerinin verimliliğini yükselten ve rekabet gücünü destekleyen bilgi ve iletişim teknolojilerini ve diğer araçları kullanan yenilikçi bir şehir olarak tanımlamaktadır [6]. Türkiye Cumhuriyeti Çevre ve Şehircilik Bakanlığı'nın 2020-2023 Ulusal Akıllı Şehirler Stratejisi ve Eylem Planı'nda, akıllı şehir; paydaşlar arası iş birliği ile geliştirilen, yenilikçi yaklaşımlar ve yeni teknolojilerden yararlanan, veri ve uzmanlığa dayalı olarak şekillenen, gelecekteki ihtiyaç ve problemleri öngörerek hayata değer katan çözümler sunan, daha yaşanabilir ve sürdürülebilir şehirler olarak ifade edilmektedir [7]. Avrupa Parlamentosu tarafından akıllı şehir yapısında öne çıkarılan altı temel bileşen bulunmaktadır. Bunlar; Akıllı Ekonomi, Akıllı İnsan, Akıllı Yönetişim, Akıllı Yaşam, Akıllı Hareketlilik ve Akıllı Çevre bileşenleridir [8].

Kullanılabilir su miktarının sanılanın aksine oldukça az durumda olmasına ek olarak hızla artan nüfus, küresel sıcaklık ve iklim değişikliği tehlikeleri ile birlikte suyun doğru yönetilmesi hayati önem taşımaktadır. Akıllı şehir misyonu doğrultusunda şehirlerde sürdürülebilir su yönetiminde makine öğrenmesi modellerinden faydalanılması kaçınılmaz bir gereklilik haline gelmiştir.

Bu çalışmada akıllı şehir bileşenlerinden **Akıllı Yönetişim** ve **Akıllı Çevre** bileşenleri ile ilişkili olarak makine öğrenmesi modelleri kullanılarak nüfus bakımından Türkiye'nin 3. büyük şehri olan İzmir şehri için içme suyu tüketimini etkileyen faktörler tespit edilerek su tüketim miktarını tahmin etmede en başarılı makine öğrenmesi modelini belirlemeye çalışacağız.

Akdağ, Diyarbakır il merkezinin 2003-2013 yılları arasındaki verilerini ele alarak Yapay Sinir Ağları, Destek Vektör Makineleri ve Box-Jenkins Yöntemleriyle Kentsel İçme Suyu Tale-

bi Tahmini ve Karşılaştırmalı Analizi çalışmasında en verimli model olarak Yapay Sinir Ağları modelini tespit etmiştir [9].

Önen, Diyarbakır il merkezinin 2005-2014 yılları arasına ait farklı öznitelikleri kullanarak 19 farklı model için Genetik Programlama ve Genetik İfadeli Programlama ile testler gerçekleştirdiği çalışmasında içme suyu tüketiminde su fiyatı ve su kayıplarının etkisinin yüksek, yağış ve nem etkisinin düşük olduğunu tespit etmiştir [10].

Ölçenoğlu, İstanbul ilinin 2010-2022 yılları arasına ait nüfus ve su tüketimi verilerini kullanarak ARIMA ve Toplamsal Winter yöntemleri ile şehrin 2033 yılındaki nüfus ve su tüketimini tahminlemiştir [11].

Oyebode, Güney Afrika'nın Ekurhuleni Şehrinden aylık su tüketimini tahmin etmek için ileri beslemeli yapay sinir ağları (YSA) ve bir destek vektör makinesi (SVM) kullanılmıştır. Çalışmada Ağustos 2010'dan Mart 2018'e kadar bu değişkenlere ilişkin veri setleri kullanılmış olup evrimsel hesaplama ile eğitilen Yapay Sinir Ağı modelinin iyi bir performans gösterdiğini tespit etmiştir [12].

Da Veiga, Brezilya'da 1746 belediye verilerini kullanarak su tüketimini etkileyen faktörlerin tespit edilmesi için 5 farklı çoklu regresyon modeli kullanarak çalışma gerçekleştirmiştir. Sonuçlar, hane geliri, su tarifesi, hane başına düşen kişi sayısı, sıcaklık, çamaşır makinesi olan hanelerin yüzdesi, şehir nüfusu, cinsiyet, şebeke suyu olan hanelerin yüzdesi ve belediyenin GSYİH'sinin kentsel su tüketimini önemli ölçüde etkilediğini göstermektedir [13].

Her şehrin sıcaklık, turizm, sanayi ve tarım gibi faktörlere bağlı olarak kendine özgü dinamikleri bulunmaktadır. Bu nedenle, bir şehirde yüksek başarı gösteren bir modelin, tüm şehirlerde aynı performansı sergileyeceğini varsaymak doğru olmayacaktır.

Çalışma sonucunda elde eden veriler ve analizler şehir yönetiminde söz sahibi olan idareciler için değerli bir rehber niteliği taşıyacaktır. Özellikle sürdürülebilir su yönetimi ve altyapı yatırımlarıyla ilgili stratejik kararların alınmasında, veriye dayalı analizlerle desteklenen öngörüler sunularak daha etkili bir planlama sürecinin oluşturulması amaçlanmaktadır. Akıllı şehir anlayışı çerçevesinde, su kaynaklarının sürdürülebilir, verimli ve adil şekilde yönetilmesi büyük önem taşımaktadır. Bu doğrultuda elde edilecek bulgular, yalnızca İzmir için değil, benzer özelliklerdeki diğer şehirler için de yol gösterici nitelikte olacaktır.

Materyal ve Metot

Veri Kümesi

Çalışmada, Ocak 2003 ile Aralık 2023 tarihleri arasındaki 252 aylık veriler kullanılmıştır. İhtiyaç duyulan veriler, İzmir Su ve Kanalizasyon İdaresi ve Meteoroloji Genel Müdürlüğü ile yapılan resmi yazışmalar yoluyla temin edilmiş, ayrıca Türkiye İstatistik Kurumu, Çalışma ve Sosyal Güvenlik Bakanlığı ile Merkez Bankasının yayınlamış olduğu bültenlerden alınmıştır.

- **İzmir Su ve Kanalizasyon İdaresinden alınan veriler:** Aylık Abone Sayısı, Aylık Su Verilen Hane Sayısı, Aylık Toplam Su Tüketimi Miktarı, 1 tonluk su fiyatı (Konut Abone-1. Kademe)

• **Meteoroloji Genel Müdürlüğünden alınan veriler:** Aylık Ortalama Nispi Nem, Aylık Ortalama Sıcaklık, Aylık Yağışlı Gün Sayısı, Aylık Ortalama Deniz Suyu Sıcaklığı, Aylık Ortalama Aktüel Basınç, Aylık Ortalama 5 cm Toprak Sıcaklığı, Aylık Maksimum Sıcaklığın 30 C Üzeri Gün Sayısı, Aylık Maksimum Yağış Miktarı, Aylık Toplam Yağış Miktarı

• **Türkiye İstatistik Kurumundan alınan veriler:** Yıllık Toplam Nüfus, İzmir Gayri Safi Yurtiçi Hasıla (GSYF)

• **Çalışma ve Sosyal Güvenlik Bakanlığından alınan veriler:** Net Asgari Ücret

• **Merkez Bankası:** Dolar Satış Kuru

Regresyon Analizi Modelleri

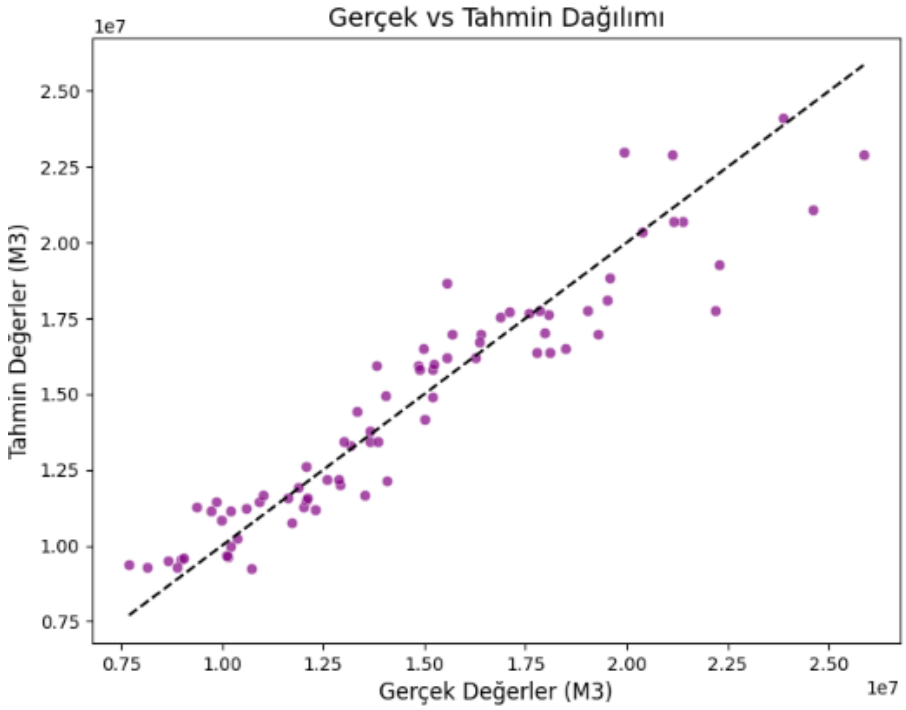
Öznitelikleri 14 farklı kombinasyon ile çalışmamızda belirlediğimiz regresyon modellerine girdi olarak vererek, hangi özniteliğin su tüketimini tahmin etmemizde etkili olduğunu tespit etmey9e çalıştık. Regresyon modeli olarak Rastgele Orman Regresyonu (Random Forest Regressor, RFR, Destek Vektör Regresyonu (Support Vector Regression, SVR), K-En Yakın Komşuluk (K-Nearest Neighbor, KNN), AdaBoostRegresor modelleri kullanılmış olup modellerin performansı, R-kare (R-squared) ve Ortalama Mutlak Yüzde Hata (Mean Absolute Percentage Error, MAPE) metrikleri ile değerlendirilmiştir.

3. Deneysel Sonuçlar

Çalışmada elde edilen deneysel sonuçlar Tablo 1'de listelenmiş olup AdaBoostRegressor Regresyon modeli R-kare (R^2) için 0.896, Ortalama Mutlak Yüzde Hata (MAPE) için %7.09 değerleri ile en iyi performans gösteren model olmuştur. İkinci en iyi performans gösteren model R^2=0.928, MAPE=%7.41 değerleri ile Rastgele Orman Regresyon modeli olmuştur. En başarılı model olan AdaBoostRegressor'ın dağılım grafiği Şekil 1'de verilmiştir.

Tablo 1: Makine Öğrenmesi Modelleri Performansları

	Model İsmi	R-squared (R^2)	Mean Absolute Percentage Error (MAPE)
1	Rastgele Orman Regresyonu (Random Forest Regressor)	0.867	7.41%
2	Destek Vektör Regresyonu (SVR - Support Vector Regression)	-0.05	23.86%
3	Keras (TensorFlow) ANN	0.652	14.09%
4	KNN Regresyon Modeli (Scikit-Learn)	0.791	10.68%
5	AdaBoostRegressor ile Regresyon Modeli (Scikit-Learn)	0.896	7,09%



Şekil 2: AdaBoostRegressor Regresyon Modeli Performansının Dağılım Grafiği

Çalışmalar sonucunda İzmir'in su tüketimini tahmin etmede en etkili özniteliklerin sırasıyla Nüfus Sayısı, Hane Sayısı, Aylık Ortalama Sıcaklık, Aylık Ortalama Aktüel Basınç, Aylık Ortalama Nispi Nem, Su Fiyat Endeksi olduğu tespit edilmiştir.

Sonuç

İzmir şehrine ait 2003-2023 yılları arasındaki aylık veriler, resmi yazışmalar ve kurumlar tarafından yayınlanan bültenlerden temin edilmiştir. Veriler, ön işleme aşamalarından geçirilerek regresyon analizi için uygun hale getirilmiş ve 14 farklı kombinasyonla öznitelikler makine öğrenmesi modellerine girdi olarak sunulmuştur.

İzmir'in su tüketimini tahmin etmede en başarılı model AdaBoostRegressor olmuştur. Bu model, R^2=0.896 ile en yüksek determinasyon katsayısına sahip olup, su tüketim değerlerini en iyi şekilde açıklayabilmektedir. Ayrıca, %7.09 MAPE ile hata oranı en düşük model olarak öne çıkmaktadır. Rastgele Orman Regresyonu modeli de R^2 = 0.867 ve %7.41 MAPE ile oldukça başarılı bir performans sergilemiştir. Buna karşılık, Destek Vektör Regresyonu (SVR) modeli negatif R^2 değeri (-0.05) ve %23.86 MAPE ile su tüketimini tahmin etmede yetersiz kalmıştır.

İzmir ili su tüketimini tahmin etmede en etkili öznitelikler incelendiğinde Nüfus Sayısı ve Hane Sayısı demografi faktörlerinin ön planda olduğu görülmektedir. Nüfus sayısı ve hane sayısının artması beklendiği gibi toplam su talebini arttırmaktadır. Bu iki öznitelik su tüketimini etkileyen en önemli faktörler olmuştur. Aylık Ortalama Sıcaklık değerinin yüksek olması, özellikle yaz aylarında tarımsal sulama, içme suyu tüketimi, turizm faaliyetleri gibi etkenler su tüketimini arttırmaktadır. Yaz aylarında artan turist sayısı, otel ve restoran gibi işletmelerde su kullanımı ile bireysel su tüketimini önemli ölçüde artmaktadır. Aylık Ortalama Aktüel Basınç ve Aylık Ortalama Nispi Nem gibi meteorolojik değişkenler buharlaşma oranını ve

su kaynaklarının verimliliğini etkileyerek su tüketimini arttırmaktadır. Su Fiyat Endeksi (1 ton su ücreti/asgari ücret) ise ekonomik bir faktör olarak su kullanımında etkisi bulunmakta olup su fiyatının artması bireysel ve endüstriyel kullanıcıların tüketim alışkanlığını değiştirmelerini sağlayabilir, su tasarrufuna teşvik edebilir.

Elde edilen sonuçlar doğrultusunda, şehrin su tüketimini tahmin etmek için gerekli teknik altyapı oluşturulmuş olup, sürdürülebilir su yönetimiyle ilgili önemli kararların veriye dayalı bilgiler ışığında alınmasına katkı sağlanmıştır.

Çıkar Beyanı**:** Yazarlar, bu çalışmada herhangi bir çıkar çatışması bulunmadığını beyan ederler. Çalışma, bağımsız bir şekilde gerçekleştirilmiş olup, finansal veya ticari çıkarlarla ilişkilendirilmiş herhangi bir etki altında değildir. Yazarlar, tüm bulguların bilimsel ve objektif bir şekilde sunulduğunu ve araştırma sürecinin etik kurallara uygun olarak yürütüldüğünü onaylamaktadırlar.

Teşekkürler: Yazarlar, bu çalışmanın hazırlanmasında destek sağlayan İzmir Su ve Kanalizasyon İdaresi, Meteoroloji Genel Müdürlüğü'ne teşekkür eder.

Yazar Katkıları: Çalışmada Fatih Tek'in fikrin oluşturulma sürecinde, literatür araştırmasında, verilerin işlenip analiz edilmesinde ve bildirinin yazımı esnasında katkıları bulunmaktadır. Fatma Günseli Yaşar Çıklaçandır'ın fikrin oluşturulma sürecinde, literatür araştırmasında, bildirinin yazımının eleştirel incelenmesinde ve bildirinin yazımında katkıları bulunmaktadır.

REFERANSLAR

[1] Meteoroloji Genel Müdürlüğü, Dünyada Su, https://www.mgm.gov.tr/genel/hidrometeoroloji.aspx?s=3

[2] Dünya Sağlık Örgütü, Türkiye'nin Su Ayak izi Raporu,16, 2014.

[3] Birleşmiş Milletler Dünya Su Kalkınma Raporu,149, 2024.

[4] Avrupa Parlementosu S.9,2014.

[5] İngiliz Standartları Enstitüsü PAS 180, 2014.

[6] Uluslararası Telekomünikasyon Birliği (ITU), 2016.

[7] Çevre, Şehircilik ve İklim Değişikliği Bakanlığı, "2020-2023 Ulusal Akıllı Şehirler Stratejisi ve Eylem Planı", 2019.

[8] Avrupa Parlamentosu, Mapping Smart Cities in EU, http://www.europarl.europa.eu/RegData/etudes/etudes/join/2014/507480/IPOLITRE_ET(2014)507480_EN.pdf

[9] R. Akdağ, "Yapay Sinir Ağları, Destek Vektör Makineleri ve Box-Jenkins Yöntemleriyle Kentsel İçmesuyu Talebi Tahmini ve Karşılaştırmalı Analizi", Business and Economics Research Journal, 123-138, 7, 1 2016.

[10] F. Önen, B. Aslan, ve N. Hamidi, "Diyarbakır Kenti İçmesuyu İhtiyacının Genetik İfadeli Programlama İle Modellenmesi", DÜMF MD, 9, 2, 859–870, 2018.

[11] A. Ölçenoğlu, ve O. Borat, "Holt-Winters ve Box-Jenkins Modellerini Kullanarak Su Tüketimi Tahmini: İstanbul Örneği", İstanbul Ticaret Üniversitesi Teknoloji ve Uygulamalı Bilimler Dergisi, 6, 2, 81-96, 2024. https://doi.org/10.56809/icujtas.1330019.

[12] O. Oyebode, ve D.E. Ighravwe, "Urban Water Demand Forecasting: A Comparative Evaluation of Conventional and Soft Computing Techniques", Resources, 8(3), 156, 2019. https://doi.org/10.3390/resources8030156.

[13] F. daVeiga, A. Kalbusch, ve E. Henning, "Drivers of Urban Water Consumption in Brazil: A Countrywide, Cross-sectional Study", Urban Water Journal, 20, 10, 1462–1470, 2022. https://doi.org/10.1080/1573062X.2022.2041049.

MANIPULATING IMAGES WITH SL(2,Z): A GROUP-THEORETIC APPROACH

Bilal Demir*
Necatibey Faculty of Education, Department of Mathematics, Balıkesir University, Balıkesir, Türkiye

Özden Koruoğlu
Necatibey Faculty of Education, Department of Mathematics, Balıkesir University, Balıkesir, Türkiye
(Corresponding author) bdemir@balikesir.edu.tr

ABSTRACT

We investigate the action of the uni-modular group on the integer lattice Z^2 in the context of discrete image manipulation. Beginning with a geometric interpretation of image transformations as linear maps, we highlight how many common manipulations—such as shearing, rotation, and reflection—correspond to elements of preserving both area and lattice structure. We then explore the algebraic structure of as a non-abelian group generated by a small set of matrices, and present a constructive algorithm to express arbitrary elements as words in these generators. This decomposition has both theoretical and computational significance, offering insight into the symmetries of discrete images and applications in lattice-based graphics, texture tiling, and symbolic image encoding. Illustrative examples and visualizations will be provided to demonstrate the dynamic interplay between algebraic operations and geometric effects on digital images.

Keywords: Image Processing, matrix transformations, word problem.

Introduction

Digital image processing is the study of algorithms and techniques used to analyze, enhance, and transform visual information represented as discrete grids of pixel values. At the core of many image operations—ranging from geometric distortions and texture mapping to pattern recognition and computer vision—lies the manipulation of the spatial structure of an image. These manipulations are often governed by mathematical transformations that reshape or reposition image content while preserving or modifying certain properties[1,2].

Linear transformations, in particular, play a central role in image manipulation. Represented by 2×2 matrices acting on R^2 or Z^2, they allow for operations such as rotation, shearing, reflection, and scaling. When an image is regarded as a function on a grid (e.g., $\mathbb{Z}^2$), applying a matrix transformation corresponds to re-mapping pixel positions according to linear rules. Understanding the algebraic structure underlying these transformations—es-

pecially when constrained to integer matrices—opens the door to elegant mathematical interpretations and efficient algorithms[2]. In this article, we focus on the special linear group $SL(2,\mathbb{Z})$, which encapsulates a class of transformations particularly well-suited to discrete and lattice-based image manipulation.

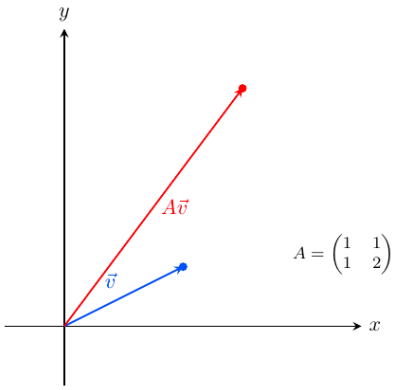


Figure 1: Transformation of the vector $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ under $A=\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$

The group $SL(2,\mathbb{Z})$, consisting of all 2x2 integer matrices with determinant 1, holds a special place in both pure and applied mathematics[3]. From a theoretical perspective, it is a classical example of a non-abelian discrete group with rich algebraic structure. Its elements act naturally on the integer lattice $\mathbb{Z}^2$, preserving area and orientation—properties that are highly desirable in image manipulation tasks where reversibility and structure-preserving transformations are important.

$$SL(2,\mathbb{Z}) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a,b,c,d \in \mathbb{Z},\ \ ad - bc = 1 \right\}$$

What makes $SL(2,\mathbb{Z})$, particularly appealing for image processing is that its action on discrete grids preserves the lattice structure and volume (since $detA=1$). This ensures that transformed images remain within the integer grid, avoiding interpolation or rounding errors that arise in more general transformations[2]. Moreover, every element of $SL(2,\mathbb{Z})$ can be built from a finite set of generators, allowing for a compact, efficient representation and manipulation of transformations. These features make it not only mathematically elegant but also computationally practical.

This paper explores the interplay between discrete image manipulation and the group $SL(2,\mathbb{Z})$ the group of 2x2 integer matrices with determinant one. We examine how transformations from this group act naturally on pixel grids, preserving area and grid alignment, making them highly suitable for structure-preserving image transformations. Additionally, we discuss the algebraic underpinnings of the group, and outline an algorithmic approach to express any element of $SL(2,\mathbb{Z})$ as a word in terms of its generators, enabling efficient transformation representation and implementation.

Linear Transformations in Image Processing

In digital image processing, an image is typically represented as a discrete collection of pixels arranged in a rectangular grid. Mathematically, this grid can be modeled as a finite subset of the integer lattice $\mathbb{Z}^2$. Each point $(x, y) \in \mathbb{Z}^2$ corresponds to a pixel location, and is associated with a value (or a vector of values) that encodes color or intensity[1].

This representation naturally lends itself to mathematical treatment using tools from linear algebra. Since pixels are located at integer-coordinate positions, any transformation of the image can be thought of as a function acting on these $\mathbb{Z}^2$ points. Linear transformations, in particular, play a central role because they preserve lines, grid structure (in specific cases), and allow for operations such as rotation, reflection, shearing, and scaling[2].

A linear transformation T from a vector space V to a vector space W is a function that satisfies two properties:

$T(u+v)=T(u)+T(v)$ for all vectors $u, v \in V$.

$T(cu) = cT(u)$ for all vectors $u \in V$ and all scalars c.

In the context of image processing, we are often concerned with linear transformations from R^2 to (or their discrete counterpart, transformations on $\mathbb{Z}^2$. These transformations can be represented by 2×2 matrices. Given a vector $v = (x,y)$, the linear transformation is defined by:

$$T(v) = Av = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$$

where $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is a 2x2 matrix with real-valued entries. When an image is regarded as a function on a grid (e.g., $\mathbb{Z}^2$), applying a matrix transformation corresponds to re-mapping pixel positions according to these linear rules.

Fundamental linear transformations in image manipulation are rotation, shearing and reflection. A rotation θ around the origin can be represented as;

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Shearing changes the shape of an image by moving points parallel to a line.

$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$ creates k times shearing in the horizontal and vertical lines, respectively.

A reflection flips the image across an axis.

$\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ are reflections of a point across the and axis respectively.

These transformations, when applied to the pixel coordinates of an image, alter its spatial arrangement, enabling a wide range of image manipulation effects.

The Group $SL(2, \mathbb{Z})$

The uni-modular group, denoted as $SL(2,\mathbb{Z})$, is defined as the set of all 2x2 matrices with integer entries and a determinant of 1, a group with significant importance across various mathematical fields. The action of this group on the upper half-plane $H=\{z \in C: im(z) > 0\}$ is described through fractional linear (Möbius) transformations. These transformations map points in H to other points in H, and this action induces a group action of $SL(2,\mathbb{Z})$ on H [3,4].

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} z \to \frac{az + b}{cz + d}$$

Elements of $SL(2,\mathbb{Z})$ forms an algebraic group under the matrix multiplication. This group is generated by the matrices;

$$T = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad and \quad S = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Hence one can obtain the presentation;

$$SL(2,\mathbb{Z}) = < T, S : T^4 = (TS)^6, (TS)^3 = T^2 >$$

This means in mathematically, every element in $SL(2,\mathbb{Z})$ can be written as a word in terms of generators. It is easy to see that the matrix T makes 90 degrees counter clock-wise about the origin. And 5 makes 1 times shearing. By the presentation the manipulation of any element in $SL(2,\mathbb{Z})$ can be described consecutive process of rotations and shearing.

As we mentioned before the action of $SL(2,\mathbb{Z})$ on H is discrete. Therefore $SL(2,\mathbb{Z})$ pos-sesses a fundamental domain. Under the action of $SL(2,\mathbb{Z})$ on the upper half-plane H, a fundamental domain is a region $F \subset H$ such that every orbit of the group intersects F at least once, and no two distinct points in the interior of F are equivalent under the group action.

A standard choice of fundamental domain for $SL(2,\mathbb{Z})$ is:

$$F = \left\{ z \in \mathbb{C} : Im(z) \geq 1 \, , -\frac{1}{2} \leq Re(z) \leq \frac{1}{2} \right\}$$

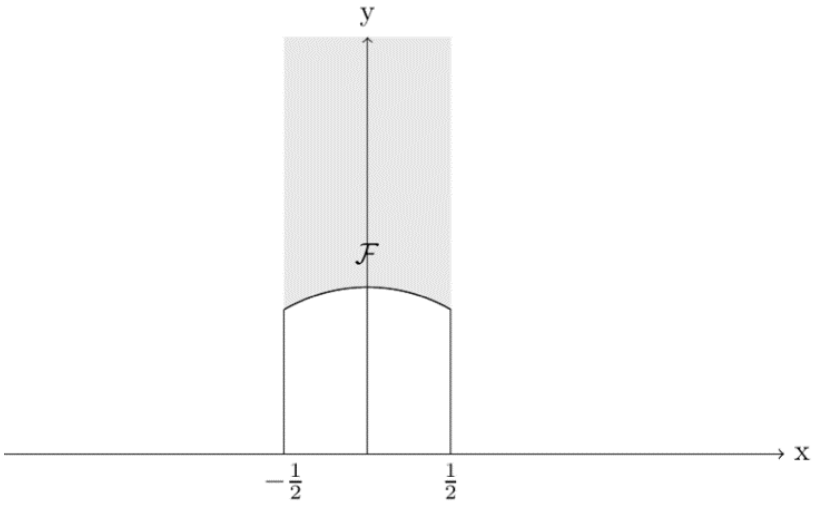


Figure 2: A Fundamental Domain of

In the view of this study, this group acts on the integer lattice Z^2 by matrix multiplica-tion. Geometrically, each element of $SL(2,\mathbb{Z})$ defines a linear transformation of the plane that preserves orientation and area. That is, these matrices distort the lattice without stretching or compressing volumes, and without flipping the plane.

Visually, applying a matrix from $SL(2,\mathbb{Z})$ rearranges the grid structure of an image in a way that keeps the integer grid aligned and consistent, making it ideal for exact, reversible image manipulations. Since the transformations are invertible and preserve the structure of the lattice, they are especially well-suited for applications in discrete geometry and digital image transformations.

Applications of in Image Manipulation

The group $SL(2,\mathbb{Z})$ provides a natural framework for manipulating digital images while preserving key geometric properties such as area and lattice structure. Since digital images are modeled as subsets of the integer grid Z^2, applying a matrix from $SL(2,\mathbb{Z})$ results in a new configuration of pixels that remains within the integer lattice, eliminating the need for interpolation or resampling.

Typical transformations from $SL(2,\mathbb{Z})$ include Shears and skews, which slide one axis relative to another, integer-based rotations around the origin (in special cases), combinations of simple transformations that remain invertible due to the determinant condition. These transformations can be used to perform reversible distortions or scramblings of images for encryption or artistic purposes, construct symmetry-preserving filters or tiling effects, Explore pattern invariants under linear actions in image analysis. Since every element of $SL(2,\mathbb{Z})$ is invertible and volume-preserving, transformations can be undone precisely, making them valuable in both compression algorithms and reversible image processing tasks.

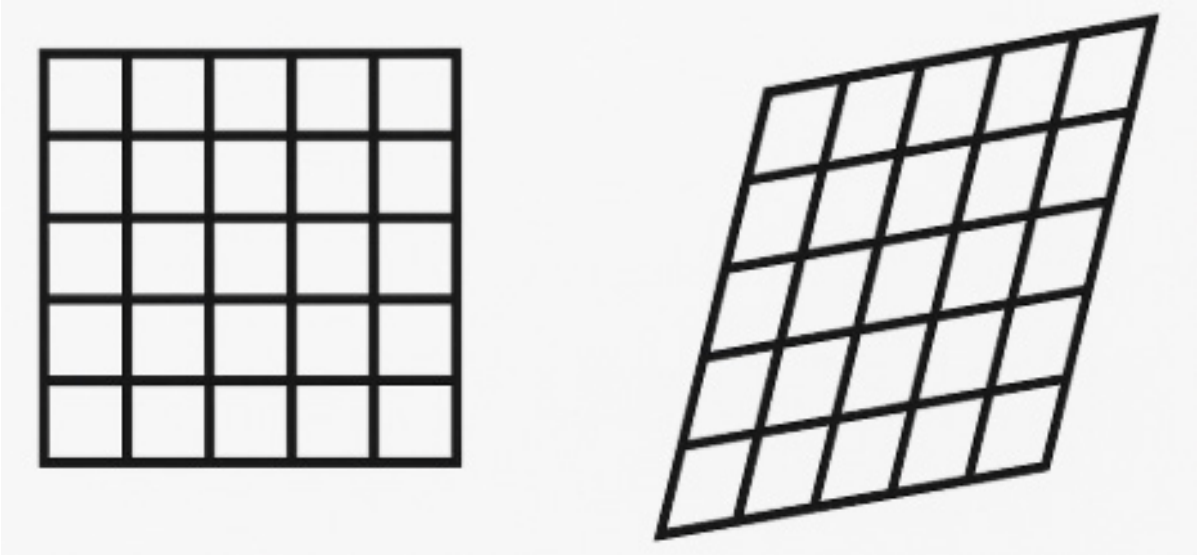


Figure 3: Shearing the grid by the matrix S

Expressing Elements of as Words

One of the remarkable features of the group $SL(2,\mathbb{Z})$ is that it is finitely generated. Every element of $SL(2,\mathbb{Z})$ can be expressed as a word S in T and and their inverses. This section outlines an algorithm to express a given element $M \in SL(2,\mathbb{Z})$ as a product (word) in S and T. The procedure is based on interpreting the action of $SL(2,\mathbb{Z})$ on the upper half-plane H and repeatedly applying generators to bring a point back into a fundamental domain.

Let $Z \in H$ be a point in the upper half-plane. Our goal is to find a matrix $M \in SL(2,\mathbb{Z})$ such that Mz=y, in the standard fundamental domain. Simultaneously, we obtain a word expression for M in terms of S and T[4,5].

Algorithm Steps:
1. Initialize: ○ Let $M \leftarrow IM$ ○ Let $y \leftarrow Mz$ 2. (repeat until $\|y\| \geq 1$ and $Re(y) \in [-1/2, 1/2]$: ○ Step A: If $Re(y) \notin [-1/2, 1/2]$, set n:=$\lfloor Re(y)+0.5 \rfloor$ ▪ Update: $M \leftarrow S^{-n} M$ ▪ $y \leftarrow S^{-n} y$ ○ Step B: If $\|y\|<1$ apply inversion: ▪ Update: $M \leftarrow TM$ ▪ $y \leftarrow Ty$ ○ Repeat 3. Return: Final matrix M and the corresponding word in S and T.

This process produces a decomposition of any $M \in SL(2,Z)$ as a word:

$$M=T^{i0} S^{j0} T^{i1} S^{j1} \ldots T^{in} S^{jn}$$

with $1 \leq i_k \leq 4, 1 \leq j_k \leq 6$. This representation is not necessarily unique, but it provides a constructive method for encoding transformations in terms of a minimal generating set.

Now we give an example of the whole process. Suppose the manipulation is applied to the image via the matrix $A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$, see in figure 4.

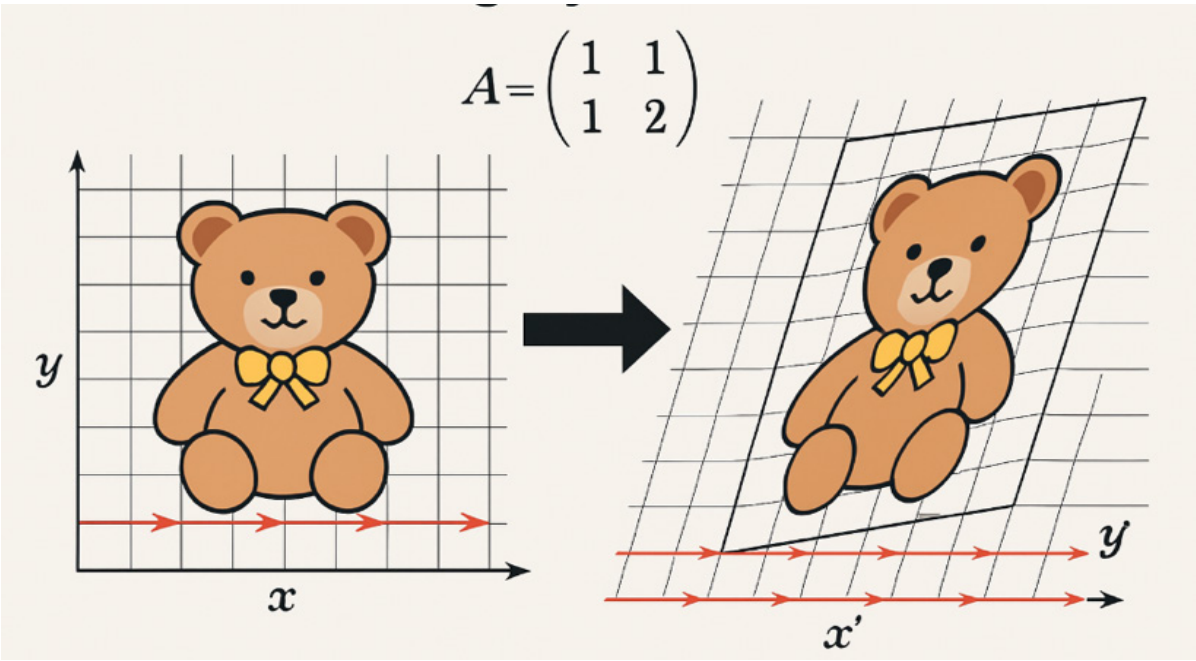


Figure 4: Manipulation by the matrix A.

In order to obtain the matrix A in terms of generators we need to choose a point in H, say z = i. Now apply $Az = \frac{i+1}{i+2} = \frac{3+i}{5} = y$. In step 2, we need to bring back this point to the fundamental domain. Observe that $Re(Az) = \frac{3}{5} > \frac{1}{2}$. Hence new M is $S^{-1}A = \begin{bmatrix} 0 & -1 \\ 1 & 2 \end{bmatrix}$.

Then one must apply S^{-1} to the point $y=\frac{3+i}{5}$.

$$S^{-1}y = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \frac{3+i}{5} = \frac{-2+i}{5}$$

Since $\left|\frac{-2+i}{5}\right| < 1$, we need the transformation T.

$$T.S^{-1}.A = \begin{bmatrix} -1 & -2 \\ 0 & -1 \end{bmatrix}$$

$$Ty = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \frac{-2+i}{5} = 2 + i$$

It is easy to see that this is a 2 unit translation. The final element that we need is S^{-2}. As a result;

$$A=STS^2$$

Declaration of Interest: The authors declare that there is no conflict of interest.

Acknowledgements: The author acknowledges the use of OpenAI's ChatGPT for assistance in generating several of the illustrative images presented in this paper.

Author Contributions: All authors contributed equally to the conception, writing, and development of this work.

REFERENCES

[1] R.C. Gonzalez and R.E. Woods, Digital Image Processing, 4th ed. Upper Saddle River, NJ: Pearson, 2018.

[2] R. Szeliski, Computer Vision: Algorithms and Applications, 1st ed. London: Springer, 2010.

[3] D.L. Johnson, Presentations of Groups, 2nd ed. Cambridge: Cambridge University Press, 1997.

[4] Coxeter, Harold SM, and William OJ Moser. Generators and relations for discrete groups. Vol. 14. Springer Science & Business Media, 2013.

[5] Yamamura, Akihiro. "Public-key cryptosystems using the modular group." International Workshop on Public Key Cryptography. Berlin, Heidelberg: Springer Berlin Heidelberg, 1998.

MESLEK YÜKSEKOKULU ÖĞRENCİLERİNİN YAPAY ZEKÂYA YÖNELİK GÖRÜŞLERİNİN BELİRLENMESİ: ORDU ÜNİVERSİTESİ ÖRNEĞİ

Servet Kılıç*
Ordu Üniversitesi, Teknik Bilimler Meslek Yüksekokulu, Ordu, Türkiye
**(İletişimden sorumlu yazar) biltek5266@gmail.com*

ÖZET

Bu araştırma, meslek yüksekokulu öğrencilerinin sınıf düzeylerine, bölümlerine ve yapay zeka (YZ) destekli uygulamaları kullanım durumlarına göre (deneyim) YZ'ye yönelik görüşlerini ortaya koymaktadır. Araştırma nicel araştırmalar arasında yer alan tarama yöntemiyle yürütülmüştür. Araştırmada belirli bir zamana yayılmadan anlık olarak bir grubun görüşlerini ortaya konulduğu için kesitsel tarama yoluyla veriler toplanmıştır. Katılımcılar, Ordu Üniversitesi'nin Meslek Yüksekokulları'nın farklı bölümlerinde öğrenim gören 207 öğrenciden oluşmaktadır. Araştırma sonuçlarına göre; öğrencilerin sınıf düzeyine göre yapay zeka ile ilgili görüşlerinde bir farklılık görülmemiştir. Daha önce yapay zeka destekli uygulamalar kullanan öğrencilerin görüşleri diğerlerine göre daha pozitiftir. Öğrenciler daha çok; ChatGPT, Gemini, Copilot ve FireFly gibi yapay zeka destekli uygulamaları tercih etmektedir. Bilgisayar Teknolojileri ile diğer teknik bölümlerde (Elektrik ve Enerji, Makine ve Metal Teknolojileri) öğrenim gören öğrencilerin görüşleri arasında anlamlı bir fark ortaya çıkmazken, İşletme Yönetimi Bölümü öğrencileri arasında anlamlı farklılıklar oluşmuştur.

Anahtar Kelimeler: Yapay zeka, meslek yüksek okulu, teknik ve sosyal bilimler

Giriş

Günümüzde yapay zeka (YZ), eğitimden sağlığa, ticaretten sanayiye ve turizmden iletişime kadar pek çok alanda köklü değişimlere yol açan öncü bir teknoloji olarak öne çıkmaktadır. Otomasyon, veri analizi, karar destek sistemleri, yorumlama ve değerlendirme gibi özellikleri sayesinde insan yaşamını dönüştürmekte, kolaylaştırmakta ve geliştirmektedir [1]. McCarthy (2004) YZ'yi, insana ait zihinsel özelliklerin mühendislik ve bilgisayar programları sayesinde makinelere aktarılması olarak tanımlamıştır. Nabiyev (2012) ise YZ'yi; makinelerin akıl yürütme, problem çözme, analiz, yorumlama ve genelleme gibi insana ait bilişsel becerileri sergilemesi olarak ifade etmiştir. En genel tanımıyla YZ, insan zekâsıyla ilişkilendirilen problem çözme, mantıksal sorgulama, analiz etme, yorumlama ve değerlendirme gibi bazı üst düzey zihinsel becerilerin bilgisayar programları aracılığıyla makinelere aktarılması çabasıdır [4].

Makinelerin düşünebilmesi ve karar verebilme özellikleri sayesinde günlük yaşamın birçok alanında YZ destekli uygulamalardan faydalanılabilmektedir. Eğitimciler; bireysel öğrenme ortamları oluşturabilir, uyarlanabilir sistemleri sayesinde öğrencilerin bireysel becerilerini ve ihtiyaçlarını belirleyebilir, öğrencileri izleyebilir, geri bildirim alabilir, verileri analiz

edebilir ve değerlendirmelerde bulunabilir [5]. Şirketler, kişilerin izlediği videolardan, arama motorlarına yazdıkları ifadelerden ve alış veriş yaptıkları sitelerden yaralanarak müşterilerin isteklerini ve yönelimlerini belirleyebilirler. Hastaların sağlık kayıtlarının analizi yoluyla hastaların geçmiş sağlık verileri taranarak risk analizi yapılabilir, ilaç etkileşimleri incelenebilir ve hastalara tedavi önerileri sunulabilir [6]. Böylece doktorlar, hastaların klinik durumlarını daha kapsamlı bir şekilde değerlendirme imkânı bulabilirler. Görüldüğü gibi farklı meslek grupları ve iş alanlarında YZ destekli uygulamalar kullanılmakta ve insan yaşamı daha kolay hale getirilebilmektedir.

YZ'nin sunduğu dönüşüm, sadece insanların yaşadığı ortamlarla ilgili olmayıp aynı zamanda bireylerin mesleki yeterliliklerini, bilgi birikimlerini ve dijital okuryazarlıklarını da etkilemektedir. Üniversite öğrencileri süreç içerisinde meslek hayatına dair teorik ve uygulamalı eğitimler almaktadır. Öğrencilerin çağın gereklerine uygun bir vizyona sahip olmaları, mezun olduklarında iş gücü piyasasında rekabetçi bireyler olmaları açısından önemlidir. Bu dönüşümün merkezinde yer alan meslek yüksekokulları (MYO) iş gücüne doğrudan katılım sağlayan bireyler yetiştirmesi bakımından kritik bir konumda bulunmaktadır. MYO'lar ağırlıklı olarak uygulamalı eğitime odaklanırken öğrencilere doğrudan mesleki beceriler kazandırmayı amaçlamaktadır. Özellikle dijitalleşme ve otomasyonun yoğun olduğu sektörlerde çalışabilecek öğrenciler, mezun olduklarında YZ ile doğrudan etkileşim içinde olan iş alanlarında çalışabilirler. Öğrencilerin YZ'ye yönelik farkındalık düzeylerinden yola çıkılarak görüşlerinin belirlenmesi onların meslek hayatına atılmadan önceki hazır bulunuşluk düzeylerini de ortaya koyabilir. Çalışmadan elde edilecek bulgular, hem öğretim programlarının YZ ile ilgili öğrencilerin bilgi, beceri ve farkındalıklarını artırma noktasındaki durumunu ortaya koyabilecektir. Bu sayede eğitim politikacılarına, yöneticilere ve eğitimcilere öğretim programlarının işlevselliği noktasında bilgiler sunulabilecektir. Bu kapsamda bu araştırmada, MYO öğrencilerinin YZ'ye ilişkin görüşlerinin belirlenmesi amaçlanmıştır.

Araştırmanın Problemleri

1. Öğrencilerin YZ hakkındaki görüşleri sınıf seviyelerine göre nasıl değişmektedir?

2.Öğrencilerin YZ hakkındaki görüşleri daha önce uygulamaları kullanma durumlarına göre (Deneyim) nasıl değişmektedir?

3.Öğrencilerin YZ hakkındaki görüşleri bölümlerine göre nasıl değişmektedir?

Yöntem

Araştırma nicel araştırmalar arasında yer alan tarama yöntemiyle yürütülmüştür. Tarama araştırmaları ile bir grubun bazı konular hakkındaki düşüncelerini ortaya koymak amaçlanmaktadır. Sonuçların genellenebilir olması amacıyla katılımcı sayısı ve çeşitliliği ön plana çıkmaktadır [7]. Veriler belirli bir zamana yayılmadan anlık olarak kesitsel tarama yoluyla toplanmıştır.

Katılımcılar

Katılımcılar amaçlı örnekleme yöntemleri arasında yer alan kolay ulaşılabilir uygun örnekleme tekniği ile belirlenmiştir [8]. Bu çerçevede Ordu üniversitesi Teknik ve Sosyal Bilimler Meslek Yüksekokulu'nda öğrenim gören toplam 207 öğrenci araştırmaya katılmıştır.

Demografik özelliklerine göre öğrencilerin dağılımları Tablo 1'de verilmiştir.

Tablo 1: Demografik bilgilere bağlı katılımcı sayıları

Kategori	Değişkenler	N
Sınıf Düzeyi	1. Sınıf	83
	2. Sınıf	124
Deneyim Durumu	Deneyimli	162
	Deneyimsiz	45
Bölüm	Bilgisayar Teknolojileri	109
	Elektrik ve Enerji	25
	İşletme Yönetimi	45
	Makine ve Metal Teknolojileri	28

Veri Toplama Araçları

Öğrencilerin yapay zekaya yönelik görüşlerini belirlemek amacıyla Süleymanoğulları ve diğerleri (2024) tarafından geliştirilen ölçek kullanılmıştır. Ölçek, etkileşim, yaygın etki ve kaygı olmak üzere 3 faktörlü olarak toplam 25 maddeden oluşmaktadır. 5'li likert tipinde hazırlanan ölçekte; en düşük 25 puan ve en yüksek 125 puan alınmaktadır. Ölçeğin genel Cronbach Alpha katsayısı ise 0.869 olarak tespit edilmiştir. Veriler 2024-2025 akademik yılı bahar döneminde yüz yüze ve çevrimiçi olarak toplanmıştır.

Verilerin Analizi

İlk olarak öğrencilerin yapay zekaya yönelik puanlarının tanımlayıcı istatistiksel analizleri yapılmıştır. Verilerin normal dağılım gösterip göstermediğini belirlemek amacıyla Kolmogorov-Smirnov testi uygulanmıştır. Sınıf düzeyine göre bakıldığında; 1. sınıfların puanları (0.167, p>0.05) ve 2.sınıfların puanları (0.200, p>0.05) normal dağılım göstermiştir. Deneyim durumlarına göre; YZ uygulamalarını kullananların puanları (0.162, p>0.05) ve kullanmayanların puanları (0.129, p>0.05) normal dağılım göstermiştir. Bilgisayar Teknolojileri (0.218, p>0.05),Elektrik ve Enerji (0.756, p>0.05), İşletme Yönetimi (0.458, p>0.05) ve Makine ve Metal Teknolojileri bölümü (0.289, p>0.05) öğrencilerinin puanları normal dağılım göstermiştir.

Sınıf düzeyi ve deneyim durumlarına göre öğrencilerin görüşlerinde anlamlı bir değişiklik olup olmadığını belirlemek için bağımsız örneklem T-testi yapılmıştır. Bölüm farklılıklarına göre anlamlı bir değişiklik olup olmadığını kontrol etmek için ise tek yönlü ANOVA testi yapılmıştır. Test sonucu anlamlı olan farklılıkların hangi gruplar arasında olduğunu belirlemek için Post-hoc testi arasında yer alan Tukey testi yapılmıştır. Öğrencilerin genel farkındalık düzeyini belirleyebilmek için genel bir değerlendirme ölçütü ortaya konulmuştur (Tablo 2). Ölçeğin her bir maddesine verilen puan aralıkları ve tüm maddelerden elde edilen toplam puan birlikte değerlendirilerek bir ölçüt ortaya konulmuştur.

Tablo 2: MYO öğrencilerinin düzeylerini ortaya koyan değerlendirme ölçütleri

Ölçek birimi aralığı	Ölçek toplam puanı aralığı	Değerlendirme ölçüt düzeyi
1.00 - 1.79	25 - 45	Çok düşük
1.80 - 2.59	46 - 65	Düşük
2.60 - 3.39	66 - 85	Orta
3.40 - 4.19	86 - 105	Yüksek
4.20 - 5.00	106 - 125	Çok Yüksek

Bulgular, Tartışma ve Sonuçlar

Değişkenlere bağlı olarak elde edilen verilerin tanımlayıcı istatistikleri Tablo 3'te verilmiştir.

Tablo 3: Bağımsız değişkenlere göre öğrencilerin tanımlayıcı istatistikleri

Değişkenler	N	M	SD
1. Sınıf	83	90.20	13.8
2. Sınıf	124	87.05	11.7
Deneyimli	162	90.43	9
Deneyimsiz	45	80.68	8.12
Bilgisayar Teknolojileri	109	91.2	13.8
Elektrik ve Enerji	25	87.9	11.7
İşletme Yönetimi	45	81.3	9
Makine ve Metal Teknolojileri	28	88.5	8.12

Tablo 3 incelendiğinde her iki sınıf düzeyinde de öğrencilerin YZ'ye yönelik görüşleri olumlu ve yüksek düzeydedir. Deneyimli olan öğrencilerin görüşleri olumlu ve yüksek düzeyde iken, deneyimsiz olan öğrenciler orta düzeydedir. Teknik ve sayısal ağırlıklı bölümdeki öğrencilerin görüşleri olumlu ve yüksek düzeyde iken, işletme bölümü öğrencilerin görüşleri orta düzeydedir.

Sınıf düzeylerine göre öğrencilerin yapay zekaya yönelik görüşleri

Levene testi sonuçlarına bakıldığında sınıf düzeylerine göre öğrencilerin puanlarının varyanslarının eşit olduğu belirlenmiştir ($F=4.368$, $p>0.05$). Tablo 4'te sınıf düzeyine göre öğrencilerin puan ortalamaları arasındaki farklılıkları ile ilgili bağımsız örneklem T-testi sonuçları verilmiştir.

Tablo 4: Sınıf düzeyine göre öğrencilerin puan ortalaması farklılıkları

Sınıf Düzeyi	N	M	Ss	sd	t	p
1. sınıf	83	90.20	10.47	201.36	1.870	.063
2.sınıf	124	87.05	13.70			

Tablo 4 incelendiğinde, 1. sınıfların YZ'ye yönelik görüşlerine ait puan ortalaması ile 2.sınıfların puan ortalaması arasında anlamlı bir fark ortaya çıkmadığı görülmektedir ($t=1.870$, $p>0.05$).

Ordu üniversitesi MYO öğretim programları incelendiğinde henüz YZ ile ilgili derslerin ve içeriklerin yer almadığı görülmektedir. Yapay zekâya dair bilgi veya farkındalık artırıcı etkinlikler her iki sınıf düzeyinde de yer almaması, öğrencilerin görüşlerinin benzer düzeyde olmasında etkili olmuş olabilir. Ayrıca, MYO'ların iki yıllık bir eğitim vermesi, geçen bir yıllık sürecin öğrencilerin gelişen teknolojiye yönelik görüşlerinin değişmesinde etkili olmayabilir.

Deneyim durumlarına göre öğrencilerin yapay zekaya yönelik görüşleri

Levene testi sonuçlarına bakıldığında deneyim durumuna göre öğrencilerin puanlarının varyanslarının eşit olduğu belirlenmiştir ($F=0.646$, $p>0.05$). Tablo 5'te deneyim durumlarına göre öğrencilerin ortalamaları arasındaki puan farklılıkları ile ilgili bağımsız örneklem T-testi sonuçları verilmiştir.

Tablo 5: Deneyim durumlarına göre öğrencilerin puan ortalaması farklılıkları

Deneyim Durumu	N	M	Ss	sd	t	p
Deneyimli	162	90.43	11.35	205	4.845*	.001
Deneyimsiz	45	80.68	13.89			

Tablo 5 incelendiğinde, deneyimli olan öğrencilerin yapay zeka yönelik görüşlerine ait puan ortalaması ile kullanmayan öğrencilerin puan ortalaması arasında anlamlı bir fark olduğu görülmektedir (t=4.845, p<0.05). Deneyimli olan öğrencilerin görüşleri anlamlı olarak farklılaşmıştır.

YZ destekli uygulamaları kullanan öğrenciler, bu teknolojilerin ne işe yaradıklarını, nasıl çalıştıklarını, faydalarını ve zararlarını daha iyi deneyimleyebilirler. Bu deneyim, öğrencilerin daha bilinçli ve olumlu görüşler geliştirmelerine katkı sağlayabilir [10]. Benzer şekilde, Keuning ve diğerleri (2024), öğrencilerin YZ araçlarını kullanma deneyimlerinin, onların bu teknolojilere yönelik tutumlarını şekillendirdiğini ifade etmişlerdir. Ayrıca, YZ uygulamalarını deneyimleyen öğrenciler, bu teknolojilere karşı daha fazla güven duyabilirler. Bu güven, öğrencilerin YZ araçlarını eğitim süreçlerinde daha etkin kullanmalarına ve bu teknolojilere yönelik daha olumlu görüşler geliştirmelerine yol açabilir. Choung ve diğerleri (2023) öğrencilerin yapay zekâ sistemlerine duyduğu güvenin, bu sistemleri kullanma niyetlerini artırdığını ifade etmiştir.

Bölümlere bağlı olarak öğrencilerin yapay zekaya yönelik görüşleri

Levene testi sonuçlarına bakıldığında bölüm farklılıklarına göre öğrencilerin puanlarının varyanslarının eşit olduğu belirlenmiştir (F=3.633, p>0.05). Tablo 6'da bölümlere bağlı olarak öğrencilerin ortalamaları arasındaki puan farklılıkları ile ilgili Tek Yönlü Anova Testi sonuçları sunulmuştur.

Tablo 6: Bölümlere bağlı öğrencilerin puan ortalaması farklılıkları

	Kareler Toplamı	Sd	Kareler Ortalaması	F	p
Gruplar arası	3096.273	3	1032.091	7.105	.000
Gruplar içi	29488.684	203	145.264		
Toplam	32584.957	206			

Tablo 6 incelendiğinde, farklı bölümlerde öğrenim gören öğrencilerin YZ'ye yönelik görüşleri arasında anlamlı bir farklılık ortaya çıkmıştır [F(3,203)=7.105, p=0.000]. Bu farklılığın hangi bölümler arasında olduğunu belirlemek için Post-hoc testleri arasında yer alan Tukey testi yapılmıştır (Tablo 7)

Tablo 7: Bölümlerarası puan ortalaması farklılıkları

Gruplar	N	X	Ss	Bilgisayar Teknolojileri	Elektrik ve Enerji	İşletme Yönetimi
Bilgisayar Teknolojileri	109	91.2	13.8			
Elektrik ve Enerji	25	87.9	11.7	p>.05		
İşletme Yönetimi	45	81.3	9	P<.05	p>.05	
Makine ve Metal Teknolojileri	28	88.5	8.12	p>.05	p>.05	p>.05

Tablo 7 incelendiğinde, Bilgisayar Teknolojileri bölümünde öğrenim gören öğrenciler ile İşletme Yönetimi bölümünde öğrenim gören öğrencilerin YZ ile ilgili görüşlerine yönelik puan ortalamaları arasında anlamlı bir fark bulunmuştur ($p<0.05$). Teknik ve sayısal ağırlıklı bölümlerde öğrenim gören öğrencilerin görüşleri arasında bir farklılık ortaya çıkmamıştır.

Teknik ve sayısal ağırlıklı eğitim veren MYO'ların öğretim programları incelendiğinde, bilgisayar kullanımı, teknik çizim, otomasyon, algoritma ve temel programlama gibi dersler birçok bölümde yer almaktadır. Bu nedenle öğrenciler derslerinde disiplinler arası olarak makine öğrenimi, veri analitiği ve otomasyon gibi ortak konulara ortak olarak odaklanabilmektedirler [13]. Bilgisayar teknolojisi öğrencileri; algoritma, veri yapıları, programlama, ağ sistemleri, siber güvenlik ve temel veri bilimi gibi konulara odaklanırken, YZ'nin potansiyelini ve sınırlarını daha gerçekçi bir şekilde değerlendirebilirler [10]. İşletme yönetimi öğrencilerinin öğretim programlarında bilgisayar destekli derslere çok fazla rastlanmamaktadır. Öğrenciler daha çok mezun olduktan sonra işletmelerde YZ destekli uygulamaları deneyimleyebilmektedirler. Ayrıca, bu öğrenciler YZ'nin iş gücü üzerindeki etkileri ve etik konular nedeniyle daha temkinli bir yaklaşım sergilemiş olabilirler [14].

Çıkar Beyanı: Her hangi bir çıkar çatışması yoktur.

REFERANSLAR

[1] S. J. Russell, and P. Norvig, "Artificial Intelligence: A Modern Approach", Upper Saddle River: Pearson Education, Inc., 2020.

[2] J. Mccarthy, "What Is Artificial Intelligence?", http://www-formal.stanford.edu/jmc/, 2004.

[3] V. V. Nabiyev, "Yapay Zeka: İnsan-bilgisayar Etkileşimi", Ankara: Seçkin Yayıncılık, 2012.

[4] A. T. Mathew, and P. Mani, "Strength of Deep Learning-based Solutions to Secure Healthcare IoT: A Critical Review", The Open Biomedical Engineering Journal, 17, 1, 2023.

[5] F. Aşık, A. Yıldız, S. Kılınç, N. Aytekin, R. Adalı ve K. Kurnaz, "Yapay Zekânın Eğitime Etkileri", Uluslararası Sosyal ve Beşerî Bilimler Araştırma Dergisi, 10, 98, 2100-2107, 2023.

[6] F. Jiang, Y. Jiang, H. Zhi, Y. Dong, H. Li, S. Ma,... and Y. Wang, "Artificial Intelligence in Healthcare: Past, Present and Future", Stroke and Vascular Neurology, 2, 4, 2017.

[7] J. R. Fraenkel, N. E. Wallen, and H. H. Hyun, "How to Design and Evaluate Research in Education", McGraw-Hill Humanities/Social Sciences/Languages, 2011.

[8] C. Teddlie, and F. Yu, "Mixed Methods Sampling: A Typology with Examples", Journal of Mixed Methods Research, 1, 1, 77-100, 2007.

[9] M. Süleymanoğulları, A. Özdemir ve A. Tekin, "Yapay Zekâ Ölçeği: Geçerlik ve Güvenirlik Çalışması", Education Science and Sports, 6, 1, 13-27, 2024.

[10] A. Turkaya ve E. B. Özdemir, "Yapay Zekâ Teknolojileri Kullanımının Ön Lisans Öğrencilerinin Dijital Okuryazarlık Düzeylerine Etkisi", Journal of University Research, 7, 4, 400-465, 2024.

[11] H. Keuning, I. Alpizar-Chacon, I. Lykourentzou, L. Beehler, C. Köppe, I. de Jong, and S. Sosnovsky, "Students' Perceptions and Use of Generative AI Tools for Programming Across Different Computing Courses", In Proceedings of the 24th Koli Calling International Conference on Computing Education Research (pp. 1-12), 2024.

[12] H. Choung, P. David, and A. Ross, "Trust in AI and its role in the acceptance of AI Technologies", International Journal of Human–Computer Interaction, 39, 9, 1727-1739, 2023.

[13] H. S. Eti, "Yükseköğretimde Yapay Zekâ: Öğrenci Tutumları ve Akademisyen Görüşleri", *İktisadi İdari ve Siyasal Araştırmalar Dergisi*, 10, 26, 132-153, 2024.

[14] K. Turgut, "Yapay Zeka'nın Yüksek Öğretimde Sosyal Bilim Öğretimine Entegrasyonu", Ankara Uluslararası Sosyal Bilimler Dergisi,1, 7, 2024.

MORPHOLOGICAL AND STATISTICAL ANALYSIS OF IRIS FLOWER SPECIES: A STUDY ON CLASSIFICATION AND DIFFERENTIATION

Kemal Polat
Bolu Abant Izzet Baysal University, Faculty of Engineering, Department of Electrical and Electronics Engineering, Bolu, Türkiye

Büşra Kaya Olgaç*
Bolu Abant Izzet Baysal University, Faculty of Engineering, Department of Electrical and Electronics Engineering, Bolu, Türkiye
(Corresponding author) kpolat@ibu.edu.tr

ABSTRACT

Machine learning is a sub-branch of artificial intelligence that enables a computer system to improve its performance by learning from data and improving that data. Machine learning is also one of the widely used methods in classification problems. It consists of data preparation, model training, estimation and evaluation steps. This article aims to classify Iris Flower and evaluate the results. The iris flower dataset contains three flower classes: Versicolor, Setosa, Virginica. Performance criteria evaluations were made with k-NN (k-Nearest Neighbors) classifier, SVM (Support Vector Machines), Decision Tree algorithms for these data. Classifiers were trained with 50-50% train-test and 75-25% train-test for each of the normal, min-max, z-score normalization methods. The k-NN algorithm gave the best performance with an AUC value of 1.00 for each of the Versicolor, Setosa, and Virginica classes at 50-50% train-test separation with normal normalization. The highest performance for each of the Versicolor, Setosa, Virginica classes was the SVM (Linear) algorithm with an AUC value of 1.00 at the 75-25% train-test separation. The highest performance for the Setosa class was all algorithms with an AUC value of 1.00 at the Min-Max normalization 50-50% train-test separation. For the Versicolor class, the SVM algorithm was the one with the AUC value of 0.99, while for the Virginica class, the SVM algorithm was again the one with the highest accuracy with the AUC value of 0.99. For the Setosa class, all algorithms had the highest performance with an AUC value of 1.00 in the Min-Max normalization 75-25% train-test separation. For the Versicolor class, the SVM algorithm was the one with the AUC value of 0.99, while for the Virginica class, the SVM algorithm was again the one with the highest accuracy with the AUC value of 0.99. In the z-Score normalization 50-50% train-test separation, the SVM (Linear) algorithm was the one with the AUC value of 1.00 for each of the Versicolor, Setosa, and Virginica classes. The SVM (Linear) algorithm was again the one that gave the highest performance with an AUC value of 1.00 for each of the Versicolor, Setosa, and Virginica classes at the 75-25% train-test separation for z-Score normalization.

This article emphasizes the high accuracy rates of machine learning in classification in the study. Machine learning studies in the fields of medicine, engineering and statistics will play an important role in improving the conditions.

Keywords: Machine learning; iris flower, k-NN, SVM (Linear), decision tree.

Introduction

Artificial intelligence and its methods have entered our lives with the developing technology to be used in various fields. Artificial intelligence is an advanced technology method that models human characteristics such as thinking and making decisions by comparison[1]. Artificial intelligence can also make decisions and produce ideas on desired subjects with its ability to think[2]. Machine learning (ML) is one of the subclasses of artificial intelligence. ML is a system that fulfills desired responsibilities such as classification, information extraction, etc.[3]. ML and Artificial Intelligence methods have made it possible to analyze large volumes of information and identify hidden patterns in data [4]. The use of machine learning is useful for many purposes, which highlights its revolutionary potential. Due to its capacity to manage complex data-driven tasks and its adaptability, machine learning is an indispensable tool in modern technology [5].

This project demonstrates a machine learning workflow using the Iris flower dataset. Classification models are created using the Iris dataset of iris flower species. The dataset is divided into groups for testing and training by percentage. Then, using the training data, datasets are trained for k-NN, SVM, Decision Tree Classifiers. Predictions are generated during the test set and the accuracy of the model is evaluated. Thus, the evaluation, prediction, model training and data preparation stages for machine learning are completed. The aim is to predict the flower species in the Iris flower dataset with high accuracy. Thus, the most appropriate classification method can be selected.

Many methods have been studied in the classification of species of Iris flowers. Each study has been based on a different method. Among the studies that have been conducted, those that have achieved significant success have been examined and are as follows.

In this study by Diptam et al. [6], Iris flower classification was performed using Neural Network. Iris flower species were tried to be identified. Classification of Iris dataset, by examining the dimensions of the petal and sepal of the Iris flower, the estimation process is demonstrated by analyzing the patterns. Artificial neural networks are effectively used in this classification problem. In this study, multilayer feedforward networks trained with backpropagation learning algorithm are used.

In their study, Poojitha et al. [7] transferred the iris flower dataset to MATLAB and used for three different types of clustering. They clustered the dataset they will use in the study using the k-means algorithm and neural network in the MATLAB environment. These neural networks are effectively used in clustering operations of large datasets. In addition, they are widely applied in areas such as pattern recognition, feature extraction, image segmentation and data mining. The results obtained include the iris dataset divided into three types in an unsupervised manner.

This method proposed by Vaishali et al. [8] is applied to Iris datasets and divides the dataset into four classes. Zero misclassifications are obtained in the test sets for the first class dataset and the results obtained for all other datasets are compared with the results found in the literature.

Materials and Methods

In the prepared study, a software was developed with Python programming language using iris dataset. In this software, k-NN, SVM and Decision Tree algorithms were applied and the dataset was modeled in Colab environment.

Materials

Table 1: Materials of Iris Flower

Type of Iris Flower	Number	Classifier	Performance Criteria
Setosa	50	k-NN SVM Decision Tree	Sensitivity Specificity Accuracy ROC curves confusion matrix, AUC value
Versicolor	50		
Virginia	50		

In this study, the classification of Iris flower was made. First of all, angle access iris flower data set was taken[9]. The received data sets were divided into two as test and training sets to be trained in k-NN, SVM and Decision Tree algorithms. The normalization of the data set was done separately with Normal, Min-Max, z-Score normalization methods. The classifiers were trained in such a way that the data sets would have 50-50% train-test, 75-25% train-test ratios. First of all, performance criteria were calculated for each of the k-NN, SVM and Decision Tree algorithms by using the Normal normalization method so that the data set would have 50-50% train-test and 75-25% train-test ratios. With the Min-Max normalization method, the data set will have 50-50% train-test and 75-25% train-test ratios, and using k-NN, SVM and Decision Tree algorithms, the performance criteria were calculated for each. With the z-Score normalization method, the data set will have 50-50% train-test and 75-25% train-test ratios, and using k-NN, SVM and Decision Tree algorithms, the performance criteria were calculated for each.

Methods

In the method section of the study, iris data set, k-NN, SVM and Decision Tree algorithms and performance evaluation criteria will be given under subheadings.

k-NN (k-Nearest Neighbors)

The k-NN algorithm is widely used in different areas such as fault cause detection, power system protection and medical detection. This algorithm is efficient and has low misclassification rate compared to other classifiers such as search trees, support vector machines and neural networks [10]. It is known as one of the simple algorithms used in ML. Although it is suitable for both classification and regression, it is mostly preferred in classifying objects. It is also a useful algorithm for filling missing data and resampling [11].

SVM (Support Vector Machines)

SVM algorithm, one of the common learning algorithms, is designed for regression and classification problems. SVM algorithm is used in areas such as face detection, image classification, text categorization, etc. [12]. There are 2 types of SVM algorithms. Linear type SVM algorithm is useful in cases where data needs to be separated in a linear manner, which means that the dataset can be divided into two classes separated by a single straight line. The nonlinear type SVM algorithm is useful in cases where the data needs to be separated nonlinearly. This means that the dataset cannot be separated into classes in a linear manner [13].

Decision Tree

The decision tree algorithm, which falls into the category of supervised learning algorithms, is generally used to solve classification problems. It can also be applied in both classification and regression analysis. The algorithm consists of nodes represented and the dataset, where the branches of the result of each node are represented. There are two types of nodes: the first is the decision node, which is used for decision making and contains more than one branch; the second is the leaf node, which is the output of the decision nodes and has no other branches. It is named after the tree shape [14]. Decision trees are one of the powerful approaches used in many different fields such as pattern recognition, ML and image processing. They are usually preferred for grouping operations. Decision trees consist of nodes and branches. Nodes represent cases belonging to a certain classification category, while each subset represents a value that can be used by the node [15].

Performance Criteria

In the classification processes, various evaluation criteria such as Receiver Operating Characteristic Curve (ROC), Area Under the ROC Curve (AUC), sensitivity, specificity, accuracy and F1 value are widely used. The ROC curve shows the probability distribution of different classes. The AUC value expresses the area under the ROC curve [16].

Experiment and Discussion

In the study, the classification process was performed using k-NN, SVM and Decision Tree machine learning algorithms using the Python programming language in the iris dataset consisting of 50 iris flower information for the classes Versicolor, Setosa and Virginica.

For each of the classes Versicolor, Setosa, Virginica, k-NN, SVM and Decision Tree ML algorithms were used in normal normalization 50-50% train-test separation. Accuracy, sensitivity, specificity, confusion matrix values are shown in the table 2, while ROC curves and AUC values are given in the graphs.

Table 2: Normal Normalization 50-50% Train-Test Table

	Accuracy	Sensitivity	Specificity	Confusion Matrix
k-NN (k=1)	0,96	1,00	0,94	[25 0 0] [0 25 0] [0 3 22]
k-NN (k=2)	0,91	1,00	0,86	[25 0 0] [0 25 0] [0 7 18]
k-NN (k=3)	0,92	1,00	0,88	[25 0 0] [0 25 0] [0 6 19]
k-NN (k=4)	0,91	1,00	0,86	[25 0 0] [0 25 0] [0 7 18]
k-NN (k=5)	0,95	1,00	0,92	[25 0 0] [0 25 0] [0 4 21]
SVM(Linear)	1,00	1,00	1,00	[25 0 0] [0 25 0] [0 0 25]
SVM(Poly)	0,95	1,00	0,92	[25 0 0] [0 25 0] [0 4 21]
SVM(Rbf)	0,97	0,96	0,98	[25 0 0] [0 24 1] [0 1 24]
SVM(Sigmoid)	0,04	0,00	0,03	[1 0 24] [14 0 11] [23 0 2]
Decision Tree(Gini)	0,91	0,96	0,88	[25 0 0] [0 24 1] [0 6 19]
Decision Tree (Entropy)	0,89	0,96	0,86	[25 0 0] [0 24 1] [0 7 18]

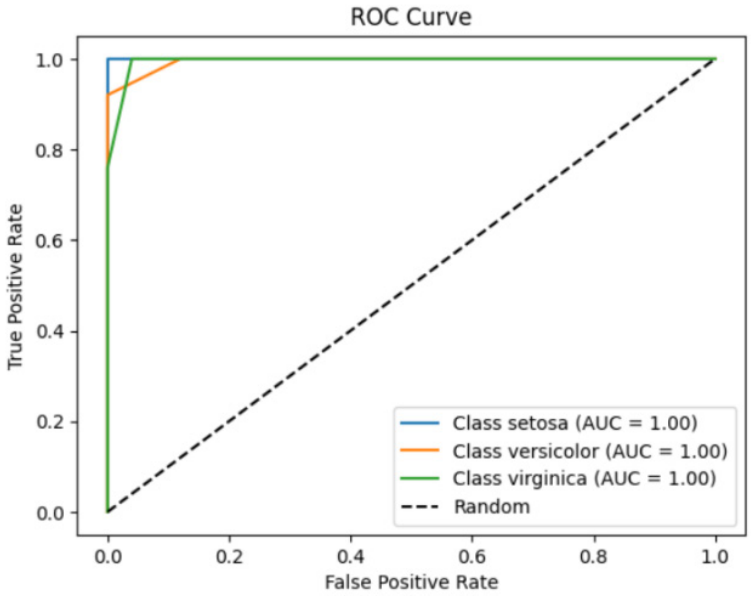


Figure 1: ROC Curve of k-NN

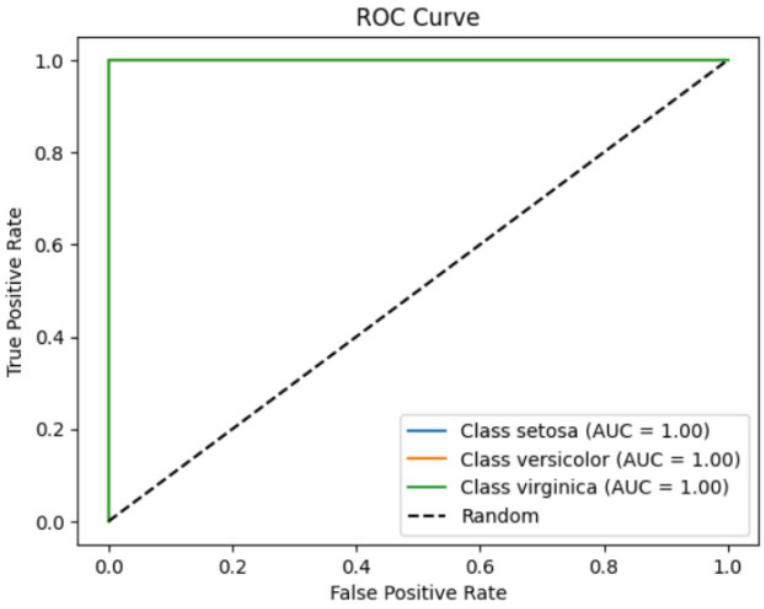


Figure 2: ROC Curve of SVM

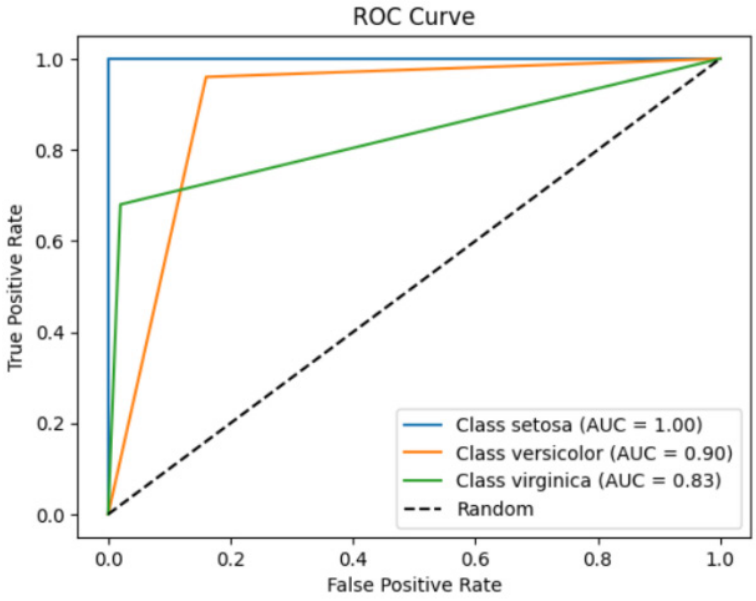


Figure 3: ROC Curve of Decision Tree

When the graphs are examined, it is observed that the AUC value of all three algorithms for the Setosa class is 1.00, reaching a high prediction rate. For the Versicolor class, the highest rate is seen with an AUC value of 1.00 for k-NN, SVM algorithms, while the rate is determined as 0.90 in the Decision Tree algorithm. For the Virginica class, the highest rate is seen with an AUC value of 1.00 for k-NN, SVM algorithms, while the rate is determined as 0.83 in the Decision Tree algorithm.

Normal normalization 75-25% train-test separation for each class of Versicolor, Setosa, Virginica k-NN, SVM and Decision Tree ML algorithms were used. Accuracy, sensitivity, specificity, confusion matrix values are shown in the table 3, while ROC curves and AUC values are given in the graphs.

Table 3: Normal Normalization 75-25% Train-Test Table

	Accuracy	Sensitivity	Specificity	Confusion Matrix
k-NN (k=1)	0,95	1,00	0,92	[12 0 0] [0 13 0] [0 2 11]
k-NN (k=2)	0,92	1,00	0,88	[12 0 0] [0 13 0] [0 3 10]
k-NN (k=3)	0,97	1,00	0,96	[12 0 0] [0 13 0] [0 1 12]
k-NN (k=4)	0,97	1,00	0,96	[12 0 0] [0 13 0] [0 1 12]
k-NN (k=5)	0,97	1,00	0,96	[12 0 0] [0 13 0] [0 1 12]
SVM(Linear)	1,00	1,00	1,00	[12 0 0] [0 13 0] [0 0 13]
SVM(Poly)	0,95	0,92	0,96	[12 0 0] [0 12 1] [0 1 12]
SVM(Rbf)	0,92	0,92	0,92	[12 0 0] [0 12 1] [0 2 11]
SVM(Sigmoid)	0,32	0,00	1,00	[12 0 0] [13 0 0] [13 0 0]
Decision Tree(Gini)	0,95	0,92	0,96	[12 0 0] [0 12 1] [0 1 12]
Decision Tree(Entropy)	0,95	0,92	0,96	[12 0 0] [0 12 1] [0 1 12]

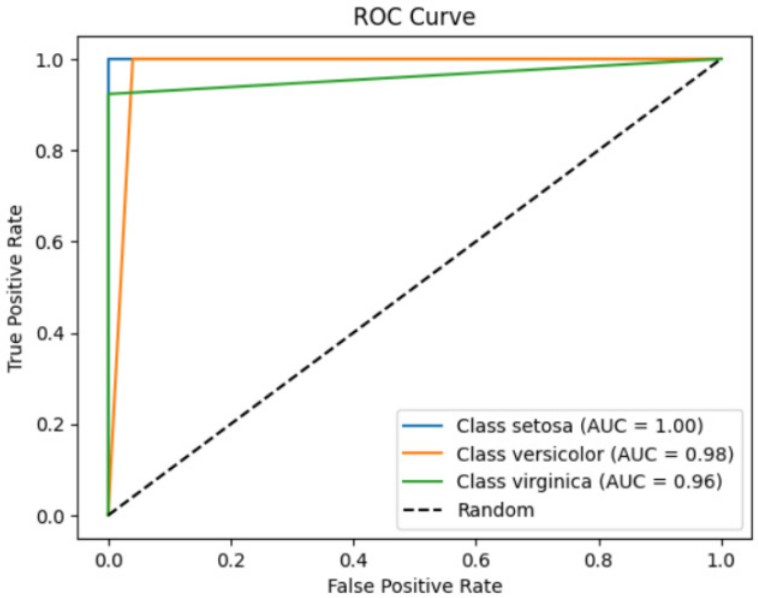


Figure 4: ROC Curve of k-NN

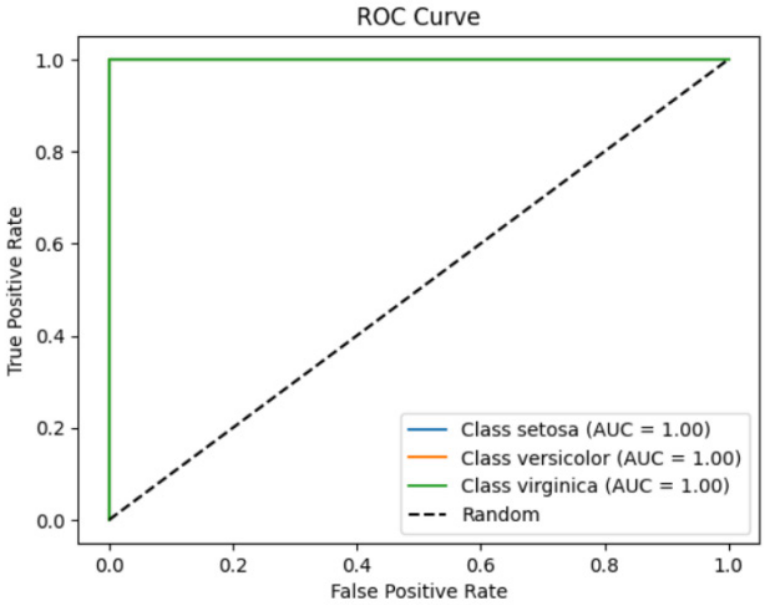


Figure 5: ROC Curve of SVM

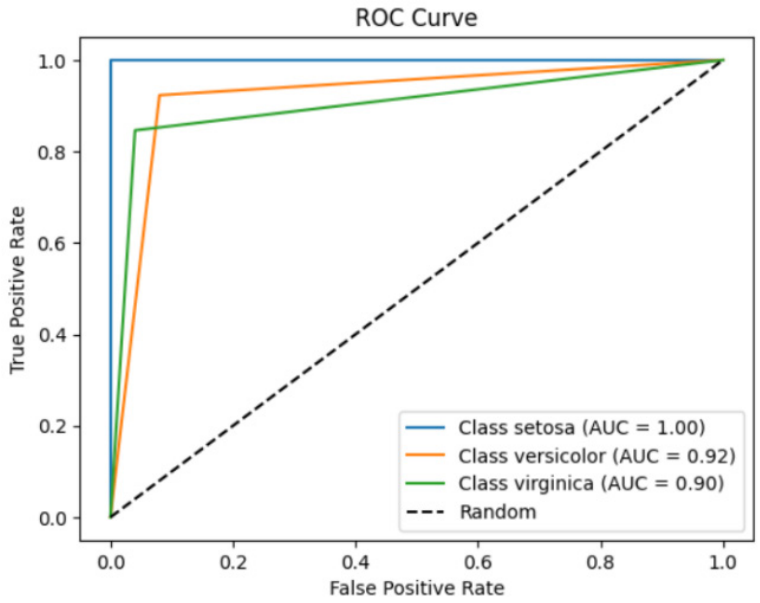


Figure 6: ROC Curve of Decision Tree

When the graphs are examined, it is observed that the AUC value of all three algorithms for the Setosa class is 1.00, reaching a high prediction rate. For the Versicolor class, the SVM algorithm appears to have the highest rate with an AUC value of 1.00, while the rate was found to be 0.92 in the Decision Tree algorithm and 0.98 in the k-NN algorithm. For the Virginica class, the SVM algorithms appear to have the highest rate with an AUC value of 1.00, while the rate was found to be 0.90 in the Decision Tree algorithm and 0.96 in the k-NN algorithm.

Min-Max normalization 50-50% train-test separation for each class of Versicolor, Setosa, Virginica k-NN, SVM and Decision Tree ML algorithms were used. Accuracy, sensitivity, specificity, confusion matrix values are shown in the table 4, while ROC curves and AUC values are given in the graphs.

Table 4: Min-Max Normalization 50-50% Train-Test Table

	Accuracy	Sensitivity	Specificity	Confusion Matrix
k-NN (k=1)	0,95	1,00	0,92	[25 0 0] [0 25 0] [0 4 21]
k-NN (k=2)	0,91	1,00	0,86	[25 0 0] [0 25 0] [0 7 18]
k-NN (k=3)	0,96	1,00	0,94	[25 0 0] [0 25 0] [0 3 22]
k-NN (k=4)	0,95	1,00	0,92	[25 0 0] [0 25 0] [0 4 21]
k-NN (k=5)	0,97	1,00	0,96	[25 0 0] [0 25 0] [0 2 23]
SVM(Linear)	0,95	0,96	0,94	[25 0 0] [0 24 1] [0 3 22]
SVM(Poly)	0,92	1,00	0,88	[25 0 0] [0 25 0] [0 6 19]
SVM(Rbf)	0,96	0,96	0,96	[25 0 0] [0 24 1] [0 2 23]
SVM(Sigmoid)	0,25	0,00	0,08	[18 0 7] [6 0 19] [14 10 1]
Decision Tree(Gini)	0,91	0,96	0,88	[25 0 0] [0 24 1] [0 6 19]
Decision Tree(Entropy)	0,88	0,96	0,84	[25 0 0] [0 24 1] [0 8 17]

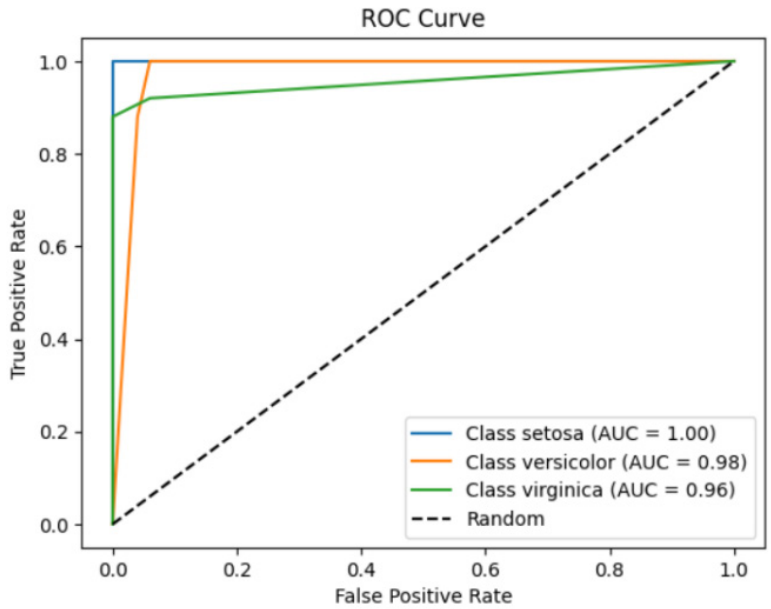


Figure 7: ROC Curve of k-NN

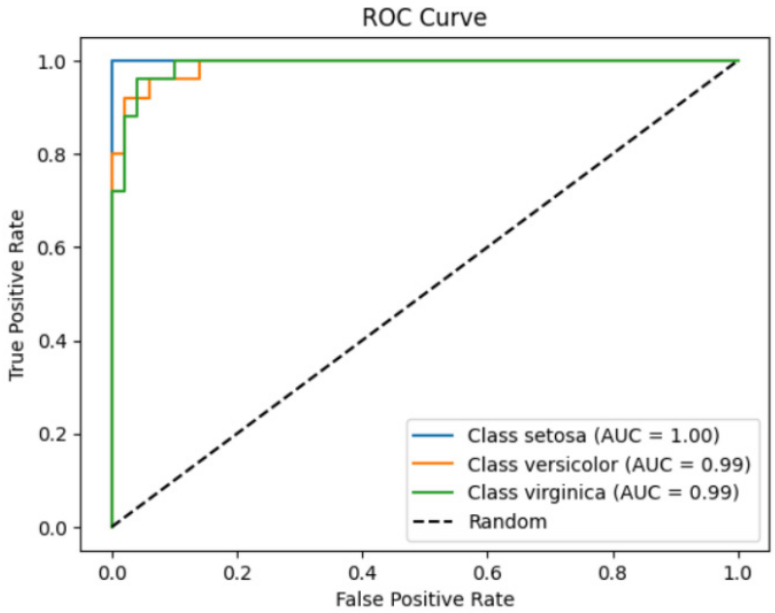


Figure 8: ROC Curve of SVM

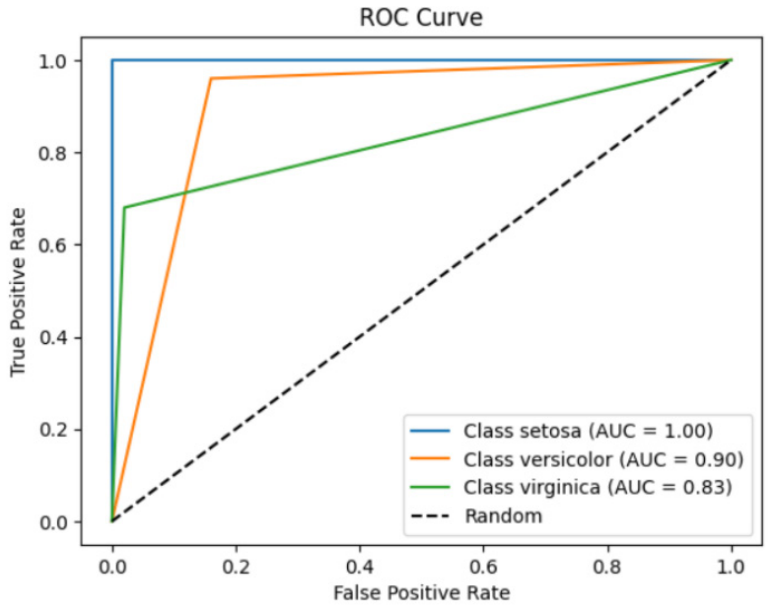


Figure 9: ROC Curve of Decision Tree

When the graphs are examined, it is observed that the AUC value of all three algorithms for the Setosa class is 1.00, reaching a high prediction rate. For the Versicolor class, the SVM algorithm appears to have the highest rate with an AUC value of 0.99, while the rate in the Decision Tree algorithm was found to be 0.90 and the k-NN algorithm was found to have 0.98. For the Virginica class, the SVM algorithms appear to have the highest rate with an AUC value of 0.99, while the rate in the Decision Tree algorithm was found to be 0.83 and the rate in the k-NN algorithm was found to be 0.96.

Min-Max normalization 75-25% train-test separation for each class of Versicolor, Setosa, Virginica k-NN, SVM and Decision Tree ML algorithms were used. Accuracy, sensitivity, specificity, confusion matrix values are shown in the table 5, while ROC curves and AUC values are given in the graphs.

Table 5: Min-Max Normalization 75-25% Train-Test Table

	Accuracy	Sensitivity	Specificity	Confusion Matrix
k-NN (k=1)	0,95	1,00	0,92	[12 0 0] [0 13 0] [0 2 11]
k-NN (k=2)	0,92	1,00	0,88	[12 0 0] [0 13 0] [0 3 10]
k-NN (k=3)	0,95	1,00	0,92	[12 0 0] [0 13 0] [0 2 11]
k-NN (k=4)	0,92	1,00	0,88	[12 0 0] [0 13 0] [0 3 10]
k-NN (k=5)	0,95	1,00	0,92	[12 0 0] [0 13 0] [0 2 11]
SVM(Linear)	0,92	0,92	0,92	[12 0 0] [0 12 1] [0 2 11]
SVM(Poly)	0,92	0,92	0,92	[12 0 0] [0 12 1] [0 2 11]
SVM(Rbf)	0,95	0,92	0,96	[12 0 0] [0 12 1] [0 1 12]
SVM(Sigmoid)	0,26	0,00	0,84	[9 0 3] [4 0 9] [8 4 1]
Decision Tree (Gini)	0,97	0,92	1,00	[12 0 0] [0 12 1] [0 0 13]
Decision Tree (Entropy)	0,92	0,92	0,92	[12 0 0] [0 12 1] [0 2 11]

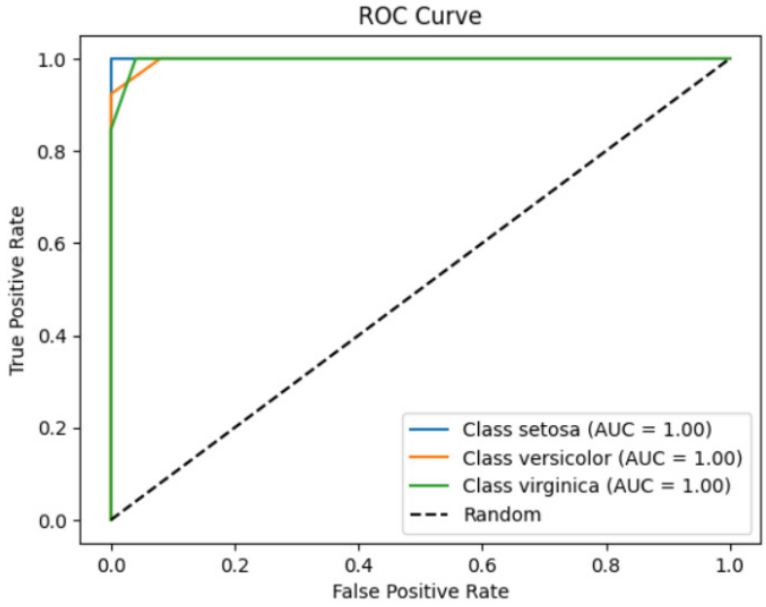


Figure 10: ROC Curve of k-NN

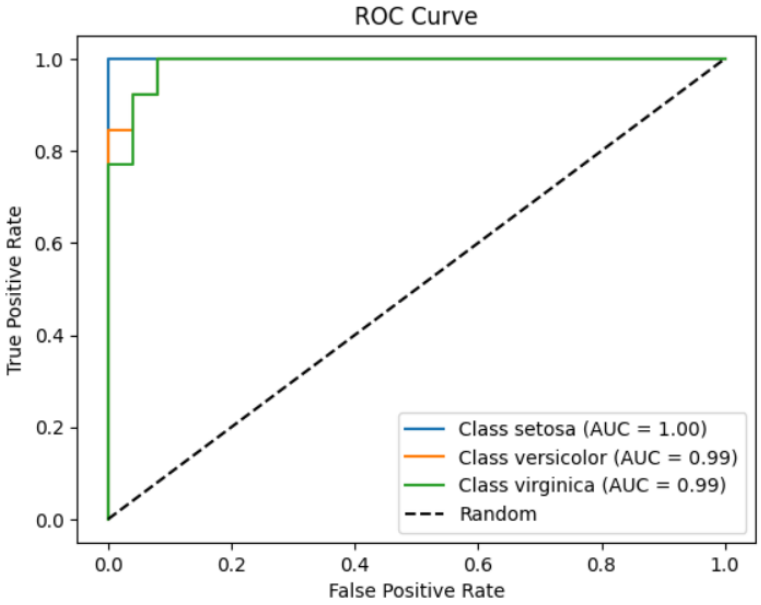


Figure 11: ROC Curve of SVM

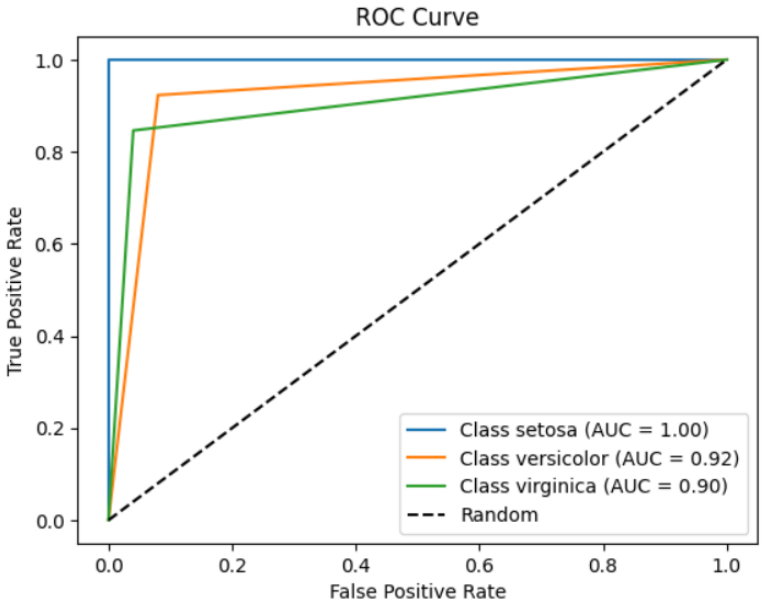


Figure 12: ROC Curve of Decision Tree

When the graphs are examined, it is observed that the AUC value of all three algorithms for the Setosa class is 1.00, reaching a high prediction rate. For the Versicolor and Virginica classes, the k-NN algorithm has the highest rates with an AUC value of 1.00.

Z-Score normalization 50-50% train-test separation for each of the Versicolor, Setosa, Virginica classes k-NN, SVM and Decision Tree ML algorithms were used. Accuracy, sensitivity, specificity, confusion matrix values are shown in the table 6, while ROC curves and AUC values are given in the graphs.

Table 6: Z-Score Normalization 50-50% Train-Test Table

	Accuracy	Sensitivity	Specificity	Confusion Matrix
k-NN (k=1)	0,93	1,00	0,90	[25 0 0] [0 25 0] [0 5 20]
k-NN (k=2)	0,91	1,00	0,86	[25 0 0] [0 25 0] [0 7 18]
k-NN (k=3)	0,92	0,96	0,90	[25 0 0] [0 24 1] [0 5 20]
k-NN (k=4)	0,91	1,00	0,86	[25 0 0] [0 25 0] [0 7 18]
k-NN (k=5)	0,95	1,00	0,96	[25 0 0] [0 25 0] [0 4 21]
SVM(Linear)	0,97	0,96	0,98	[25 0 0] [0 24 1] [0 1 24]
SVM(Poly)	0,87	1,00	0,80	[25 0 0] [0 25 0] [0 10 15]
SVM(Rbf)	0,96	0,96	0,96	[25 0 0] [0 24 1] [0 2 23]
SVM(Sigmoid)	0,84	0,56	0,98	[25 0 0] [0 14 11] [0 1 24]
Decision Tree(Gini)	0,91	0,96	0,88	[25 0 0] [0 24 1] [0 6 19]
Decision Tree (Entropy)	0.88	0,96	0,84	[25 0 0] [0 24 1] [0 8 17]

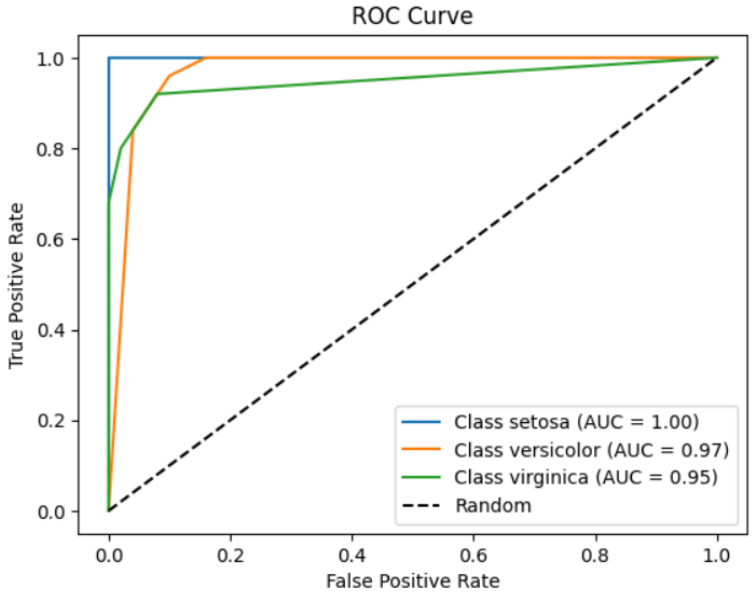


Figure 13: ROC Curve of k-NN

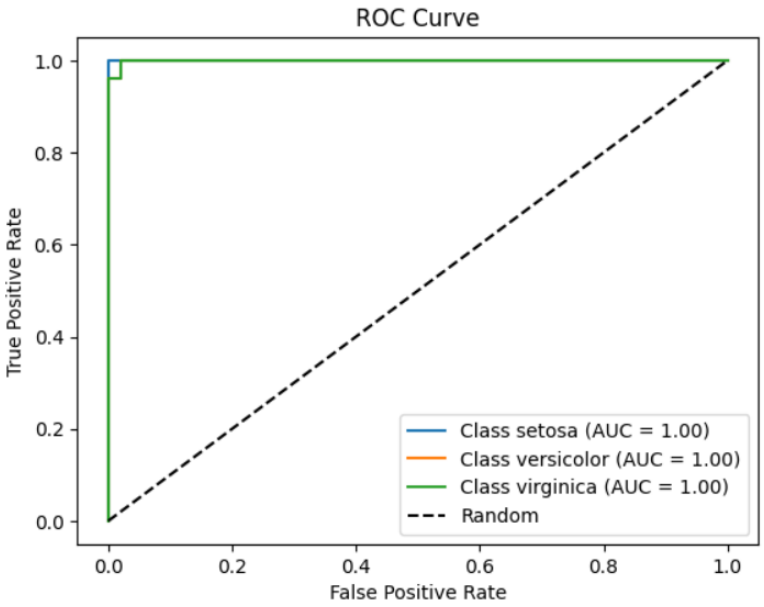


Figure 14: ROC Curve of SVM

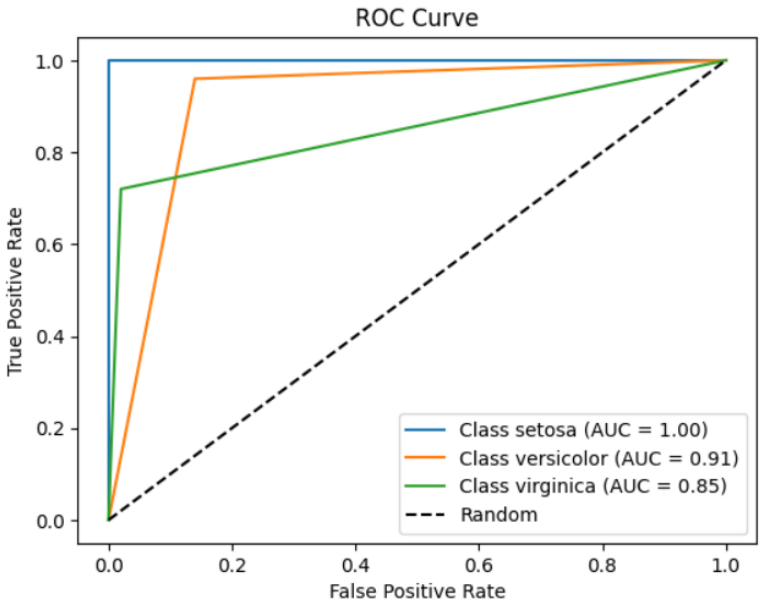


Figure 15: ROC Curve of Decision Tree

When the graphs are examined, it is observed that the AUC value of all three algorithms for the Setosa class is 1.00, reaching a high prediction rate. For the Versicolor class, the SVM algorithm appears to have the highest rate with an AUC value of 1.00, while the rate was found to be 0.91 in the Decision Tree algorithm and 0.97 in the k-NN algorithm. For the Virginica class, the SVM (Support Vector Machines) algorithms appear to have the highest rate with an AUC value of 1.00, while the rate was found to be 0.85 in the Decision Tree algorithm and 0.95 in the k NN algorithm.

Z-Score normalization 75-25% train-test separation for each of the Versicolor, Setosa, Virginica classes k-NN, SVM and Decision Tree ML algorithms were used. Accuracy, sensitivity, specificity, confusion matrix values are shown in the table 7, while ROC curves and AUC values are given in the graphs.

Table 7: Z-Score Normalization 75-25% Train-Test Table

	Accuracy	Sensitivity	Specificity	Confusion Matrix
k-NN (k=1)	0,95	1,00	0,92	[12 0 0] [0 13 0] [0 2 11]
k-NN (k=2)	0,92	1,00	0,88	[12 0 0] [0 13 0] [0 3 10]
k-NN (k=3)	0,92	1,00	0,88	[12 0 0] [0 13 0] [0 3 10]
k-NN (k=4)	0,92	1,00	0,88	[12 0 0] [0 13 0] [0 3 10]
k-NN (k=5)	0,92	1,00	0,88	[12 0 0] [0 13 0] [0 3 10]
SVM(Linear)	0,97	0,92	1,00	[12 0 0] [0 12 1] [0 0 13]
SVM(Poly)	0,89	1,00	0,84	[12 0 0] [0 13 0] [0 4 9]
SVM(Rbf)	0,95	0,92	0,96	[12 0 0] [0 12 1] [0 1 12]
SVM(Sigmoid)	0,89	0,76	0,96	[12 0 0] [0 10 3] [0 1 12]
Decision Tree(Gini)	0,95	0,92	0,96	[12 0 0] [0 12 1] [0 1 12]
Decision Tree(Entropy)	0,92	0,92	0,92	[12 0 0] [0 12 1] [0 2 11]

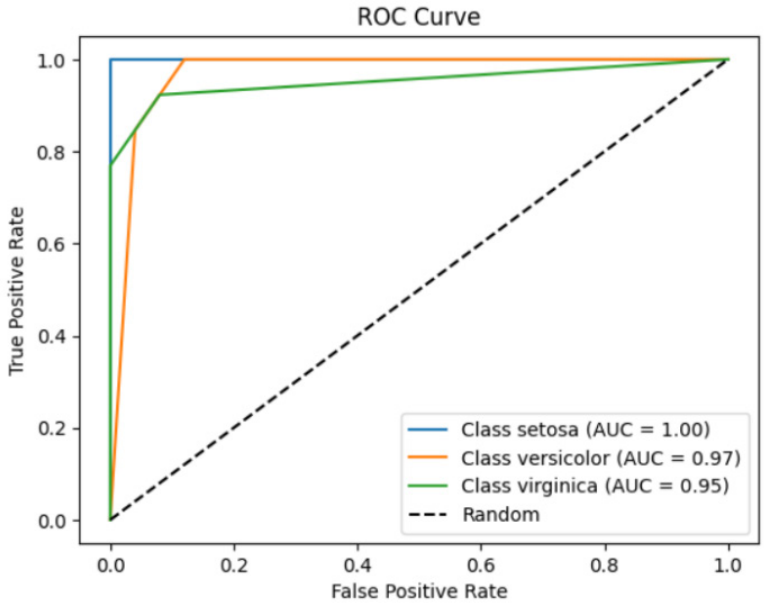


Figure 16: ROC Curve of k-NN

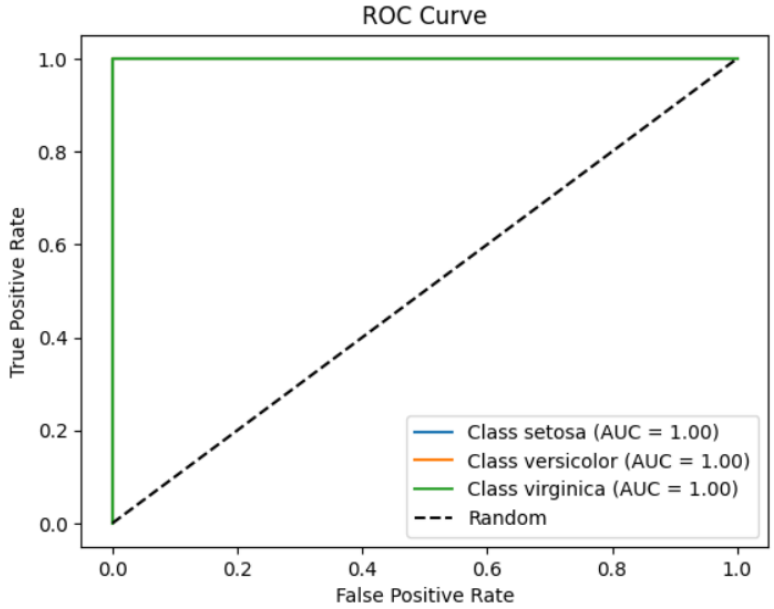


Figure 17: ROC Curve of SVM

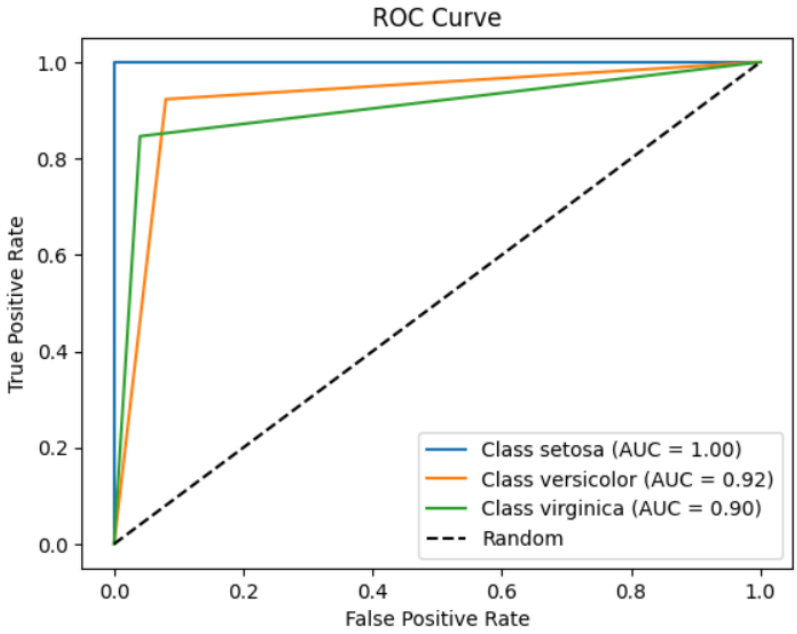


Figure 18: ROC Curve of Decision Tree

When the graphs are examined, it is observed that the AUC value of all three algorithms for the Setosa class is 1.00, reaching a high prediction rate. For the Versicolor class, the SVM algorithm appears to have the highest rate with an AUC value of 1.00, while the rate was found to be 0.92 in the Decision Tree algorithm and 0.97 in the k-NN algorithm. For the Virginica class, the SVM algorithms appear to have the highest rate with an AUC value of 1.00, while the rate was found to be 0.90 in the Decision Tree algorithm and 0.95 in the k-NN algorithm.

Conclusion

The type of iris flower was determined using SVM, k-NN, Decision Tree algorithms and the performance of classification accuracy was evaluated as a result of this study. When looking at the results of each of the performances obtained, it was seen that the SVM algorithm had the highest accuracy rate in predicting flower types.

An artificial intelligence model was proposed for the classification of iris flowers with the results obtained from the evaluation of the obtained data. As a result of the studies, the proposed model aims to support the literature in the application areas of artificial intelligence. In the new studies to be conducted, it is planned to improve the accuracy rates by working on unused artificial intelligence models.

Declaration of Interest: The authors declares that there is no financial or personal conflict of interest regarding this study. No interest group or commercial organization has supported the conduct of the study, analysis of the data, and writing of the article. The authors confirms that there is no conflict of interest in the study.

REFERENCES

[1] İ. Özkan, and E. Ülker, "Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri", Gaziosmanpaşa Bilimsel Araştırma Dergisi, 6, 3, 85-104, 2017.

[2] B. Aksoy, H.D. Halis, O.K.M. Salman, "Elma Bitkisindeki Hastalıkların Yapay Zekâ Yöntemleri ile Tespiti ve Yapay Zekâ Yöntemlerinin Performanslarının Karşılaştırılması", International Journal of Engineering and Innovative Research, 2, 3, 194-210, 2020.

[3] T. Mitchell, "Machine Learning", McGraw Hill, 1997.

[4] K.B. Johnson, W. Wei, D. Weeraratne, M.E. Frisse, K. Misulis, K. Rhee, et al., "Precision Medicine, AI, and The Future of Personalized Health Care", Clinical Translational Science, 14, 1, 86-93, 2021.

[5] R. Veena, K. Preethi, "Iris Flower Classification Using Machine Learning", International Research Journal of Modernization in Engineering Technology and Science, 6, 7, 2024.

[6] D. Diptam, R. Argha, C. Kaustav, "Training Aritificial Neural Network Using Particle Swarm Optimization Algorithm", International Journal on Computer Science And Engineering(IJCSE), 3, 3, 2013.

[7] V. Poojitha, J. Shilpi, "A Collecation of IRIS Flower Using Neural Network CLusterimg tool in MATLAB", International Journal on Computer Science And Engineering(IJCSE)

[8] A. Vaishali, R.K. Rathy, "An Efficient Neura-Fuzzy Approach For Classification of Dataset", Optimization and Information Technology, 2014. DOI:10.24432/C56C76

[9] T. Yaochi, C. Yunchi, L. Kuohao,"Applications of K-nearest neighbor algorithm in intelligent diagnosis of wind turbine blades damage", Renewable Energy, 212, 855-864, 2023.

[10] N. Zamri, M. Pairan, W. Amira, W. Azman, S.S. Abas, L. Abdullah, S. Naim, Z. Tarmudi, M. Gao, "River Quality Classification Using Different Distances in K-Nearest Neighbors Algorithm ", Procedia Computer Science, 204, 180-186, 2022.

[11] A. Lambora, K. Gupta, K. Chopra, "Genetic algorithm - A literature review", Cloud and Parallel Computing (COMITCon), 380–384, 2019.

[12] B. Malti ,G. Apoorva , C. Apoorva, "A comparative analysis of K-Nearest Neighbor, Genetic, Support Vector Machine, Decision Tree, and Long Short Term Memory algorithms in machine learning", Decision Analytics Journal, 3,100071, 2022.

[13] M. Bansal, R. Yadav, P.K. Ujjwal, "Palmistry using machine learning and opencv", Fourth International Conference on Inventive Systems and Control (ICISC), 536–539, 2020.

[14] I. Rakibul, S. Azrin, T. Nuruzzaman, "A comparative analysis of machine learning algorithms with tree-structured parzen estimator for liver disease prediction", Healthcare Analytics, 6, 100358, 2024.

[15] H.C. Bayrakçı, A.B. Keşkekçi, R. Arslan, "Classification of Iris Flower by Random Forest Algorithm", Dergipark, 2, 7-14, 2022.

NUMERICAL SOLUTION OF THE PANTOGRAPH TYPE DIFFERENTIAL EQUATION USING JACOBI NEURAL NETWORK

M. Mustafa Bahşı*
Manisa Celal Bayar University, Kırkağaç Vocational School, Manisa, Türkiye

Ayşe Kurt Bahşı
Manisa Celal Bayar University, Faculty of Engineering and Natural Science, Manisa, Türkiye
**(Corresponding author) mustafa.bahsi@cbu.edu.tr*

ABSTRACT

This study presented a new method for solving the pantograph type differential equations with various initial and boundary conditions using the Jacobi Neural Network (JNN). Jacobi polynomials are used as activation functions in the middle layer of the neural network, so the architecture has been called Jacobi neural network. The trial solution of the pantograph type differential equation is considered as the output of the feed-forward neural network. The neural network consists of adjustable weights with Newtons' like method. The architecture of the JNN consists of three layers named input layer, hidden layer and output layer. In the input layer, an independent variable is given. In the hidden layer, Jacobi polynomials have been used to activate the input from the input layers. The output layer consists of a linear combination of the inputs with weights of the hidden layer. Various problems have been solved by the presented method for the effectiveness of the method. The JNN solutions are compared with different methods. In the application of the method, the convergence of the weights and its stability were examined.

Keywords: Jacobi polynomials, neural network, pantograph equation.

Introduction

Pantograph-type delay differential equations (PTDDEs) represent a special class of functional differential equations characterized by the presence of proportional delays in the independent variable. These equations arise naturally across a wide range of disciplines, including electromagnetism, control theory, number theory, probability, and quantum mechanics, underscoring their significance in both theoretical and applied mathematics. The presence of delayed terms makes solving PTDDEs analytically and numerically challenging. In recent years, various methods have been developed to address these difficulties. For instance, Albidah et al. [1] utilized the homotopy perturbation method to obtain both analytical and numerical solutions for PTDDEs. Similarly, Alrebdi and Al-Jeaid [2] employed the Laplace transform technique to derive accurate solutions. Furthermore, Sriwastav and Barnwal [3] proposed a Bernstein polynomial-based approach to obtain numerical solutions for Lane-Emden-type PTDDEs. These studies [4–6] collectively demonstrate the effectiveness of diverse analytical and numerical techniques for tackling PTDDEs.

Beyond conventional numerical methods, artificial neural networks (ANNs) have increasingly been adopted for solving differential equations. Notably, neural network architectures incorporating orthogonal polynomials have shown great potential for enhancing solution accuracy and computational efficiency. For example, Yang et al. [7] applied the Jacobi spectral collocation method to solve fractional pantograph delay differential equations. Such studies [8–10] illustrate the promising role of orthogonal polynomials within neural network frameworks.

In this work, a novel approach, termed the Jacobi Neural Network (JNN) method, is proposed for the solution of PTDDEs. In the JNN approach, Jacobi polynomials are employed as activation functions within the hidden layer of the neural network, and the network's output is modeled as the solution function. The network weights are optimized using a Newton-like optimization algorithm, thereby enhancing solution accuracy. This methodology is designed to effectively address PTDDEs under both initial and boundary conditions. In summary, the objective of this study is to demonstrate the effectiveness of the proposed JNN approach in solving PTDDEs and to offer a viable alternative to existing methods in literature. To this end, the proposed method has been applied to a variety of test problems, and the results have been compared and evaluated against other established techniques.

Background

In this section, we will review the basic concepts on which our study is based. We introduce the notations used, Jacobi polynomials, briefly explain the present the pantograph equations.

Jacobi Polynomials

The orthogonal Jacobi polynomials $P_n^{(\alpha,\beta)}(x)$ to develop a new approach to obtain numerical solutions of the PTDD equation. The orthogonal Jacobi polynomials are defined with respect to $\omega^{\alpha,\beta}(x) = (1-x)^{\alpha}(1+x)^{\beta}$ $(\alpha > -1, \beta > -1)$(weight function) on (-1,1) and it is proved that the Jacobi polynomials satisfy the following relation.

$$P_n^{(\alpha,\beta)}(x) = \sum_{k=0}^{n} 2^{-k} \binom{n+\alpha+\beta+k}{k} \binom{n+\alpha}{n-k} (x-1)^k;\ k = 0, 1, 2, \dots, n$$

Note that α and β are special parameters of the Jacobi polynomials. The Jacobi polynomials transform into some known orthogonal polynomials with respect to α and β. Some of the most important are Legendre ($\alpha= \beta=0$), Chebyshev ($\alpha= \beta=-1/2$) and Gegenbauer ($\alpha= \beta$) polynomials [11, 12].

Pantograph Equations

The closed form of the kth order pantograph type delayed differential equation, which includes the term $y(\lambda x)$ and its derivatives, where $0<\lambda<1$, is given below in closed form.

$$F\left(y^{(k)}(x), .., y'(x), y(x), y^{(k-1)}(\lambda x), \dots, y'(\lambda x), y(\lambda x), x\right) = 0, a \le x \le b \qquad (1)$$

The mixed conditions of the given equation are as follows.

$$c_i^1 y^{(i)}(a) + c_i^2 y^{(i)}(b) + c_i^3 y^{(i)}(c_i) = \gamma_i, \quad i = 0, 1, 2, \ldots, k-1 \tag{2}$$

where for $\forall i\ c_i^1, c_i^2, c_i^3 \in \mathbb{R}$ and $a, b, c_i \in \mathbb{R}$ such that $a < c_i < b$.

Jacobi Neural Networks Model

In this section, we will introduce the model based on artificial neural network architecture that is used to find numerical solutions of pantograph type delayed differential equations. Jacobi polynomials have been used as activation functions in the developed artificial neural network architecture-based model. So, the model is called Jacobi Neural Network Model.

Architecture of the Jacobi Neural Network Model

Jacobi Neural Network Model are composed of an input layer, a hidden layer and an output layer. The architecture of JNN model can be seen in Figure 1. The model architecture is designed to have an artificial neuron in the input layer, an adjustable number of artificial neurons in the hidden layer and an artificial neuron in the output layer. The independent variable of the problem is given in the input inlayer. In the hidden layer, the independent variables are activated with activation functions in the artificial neurons. Jacobi Polynomials with different orders are used as activation functions. The number of artificial neurons in the hidden layer will depend on the problem. The number of neurons must be greater than the order of the problem.

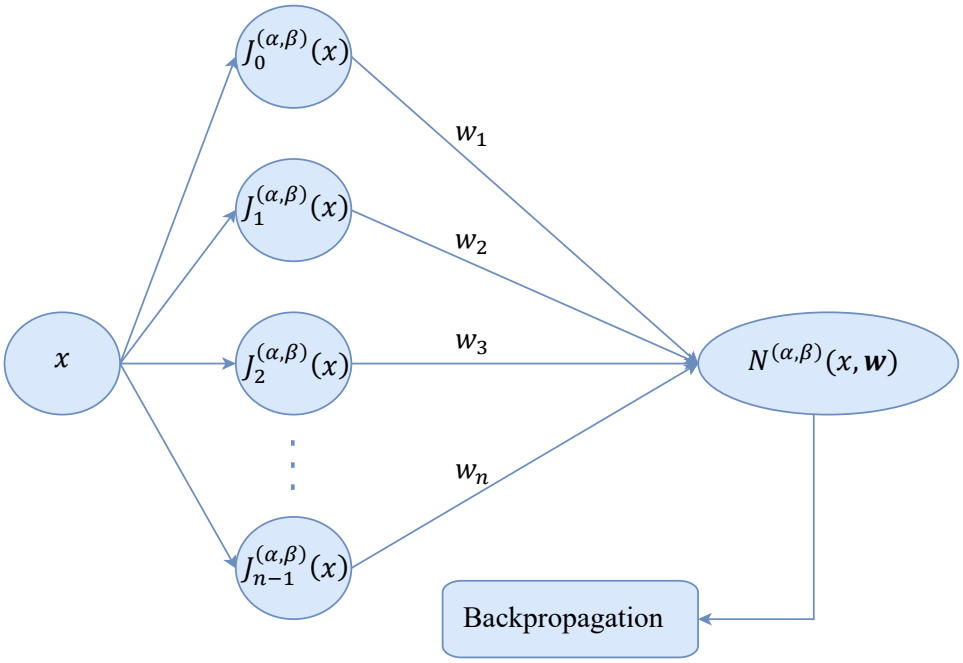


Figure 1: The architecture of JNN model.

Jacobi Neural Network for Solving Pantograph Equations

A trial solution suitable for the JNN architecture defined in the previous section will be proposed. Let the number of artificial neurons in the hidden layer of the model be n. The

activation function of each artificial neuron will consist of Jacobi polynomials with different degrees. Let the activation function of the ith neuron be $J_{i-1}^{(\alpha,\beta)}(x)$ and the connection weight of this neuron with the output layer be w^i. α and β are the hyper parameters for the model coming from the definition of the Jacobi polynomial of the JNN model.

Let's define the trial solution of the model as $N^{(\alpha,\beta)}(x, \boldsymbol{w})$. x is the independent variable (input data) w= [w_1, w_2, ..., w^n] are the weights.

$$N^{(\alpha,\beta)}(x, \boldsymbol{w}) = \sum w_i J_{i-1}^{(\alpha,\beta)}(x) \tag{3}$$

The trial solution $N^{(\alpha,\beta)}(x, \boldsymbol{w})$ will be shown as $N\,(x, w)$ for ease of notation. If the trial solution is applied to problem (1) and condition (2), it becomes a

$$\underset{\boldsymbol{w}}{Min} \sum_x \left(F\left(N^{(k)}(x, \boldsymbol{w}), \dots, N'(x, \boldsymbol{w}), N(x, \boldsymbol{w}), N^{(k)}(\lambda x, \boldsymbol{w}), \dots, N'(\lambda x, \boldsymbol{w}), N(\lambda x, \boldsymbol{w}), x\right)\right)^2 \tag{4}$$

minimization problem, with constraints

$$G_i(\boldsymbol{w}) = c_i^1 N^{(i)}(a, \boldsymbol{w}) + c_i^2 N^{(i)}(b, \boldsymbol{w}) + c_i^3 N^{(i)}(c_i, \boldsymbol{w}) - \gamma_i, \;\; i = 0, 1, 2, \dots, k-1 \tag{5}$$

Learning Algorithm of the Jacobi Neural Network Model

JNN will be trained on the points $x_i = a + \frac{b-a}{h-1}(i-1); i = 1,2, \dots, h,$ which divide the interval $[a, b]$ where the problem is defined into h equal parts. The cost function at points of the minimization problem presented in (4) is defined below.

$$E(\boldsymbol{w}) = \frac{1}{2h}\sum_{i=1}^{h}\left(F\left(N^{(k)}(x_i, \boldsymbol{w}), \dots, N'(x, \boldsymbol{w})N(x_i, \boldsymbol{w}), N^{(k)}(\lambda x_i, \boldsymbol{w}), \dots, N(\lambda x_i, \boldsymbol{w}), x\right)\right)^2 \tag{6}$$

The Lagrange function defined as the sum of the linear sum of each of the k constraints presented in (5) with the Lagrange multipliers $\lambda_0, \lambda_1, \dots, \lambda_{k-1}$, and the cost function is presented below.

$$L(\boldsymbol{w}, \boldsymbol{\lambda}) = E(\boldsymbol{w}) + \sum_{i=0}^{k-1} \lambda_i G_i(\boldsymbol{w}) \tag{7}$$

where $\boldsymbol{\lambda} = [\lambda_0, \lambda_1, \dots, \lambda_{k-1}]$. .

Backpropagation, a widely used method in machine learning for training neural networks and updating their parameters, will be employed to train the JNN. This algorithm aims to iteratively adjust each weight so that the network output approximates the desired result more closely. To facilitate this process, the Newton optimization method will be utilized [13, 14].

A column vector v, consisting of n weights and k Lagrange multipliers, is defined accordingly.

$$\boldsymbol{v} = \begin{bmatrix} \boldsymbol{w}_{n\times 1}^T \\ \boldsymbol{\lambda}_{k\times 1}^T \end{bmatrix}_{(n+k)\times 1}$$

The elements of the vector $\boldsymbol{v}$ consist of the weights and Lagrange multipliers. Let $\boldsymbol{v}^0$ denote the initial values, which are assigned randomly, and $\boldsymbol{v}^i$ represent the updated weights and λ_i Lagrange multipliers obtained at the end of the -th iteration. The update rule for vector $\boldsymbol{v}$ is given by the following equation.

$$\boldsymbol{v}^{i+1} = \boldsymbol{v}^i + \Delta\boldsymbol{v}^i; i = 0, 1, 2, \ldots \quad (8)$$

where

$$\Delta\boldsymbol{v}^i = \boldsymbol{W}^{-1}(\boldsymbol{w}^{i-1}, \boldsymbol{\lambda}^{i-1})\left(-\boldsymbol{M}(\boldsymbol{w}^{i-1}, \boldsymbol{\lambda}^{i-1})\right)$$

such that

$$\boldsymbol{W}(\boldsymbol{w}, \boldsymbol{\lambda}) = \begin{bmatrix} \boldsymbol{H}_{n\times n} & : & \boldsymbol{N}_{n\times k} \\ \boldsymbol{N}^T{}_{k\times n} & : & \boldsymbol{0}_{k\times k} \end{bmatrix}_{(n+k)\times(n+k)} \quad \boldsymbol{M}(\boldsymbol{w}, \boldsymbol{\lambda}) = \begin{bmatrix} \boldsymbol{L}_{n\times 1} \\ \boldsymbol{G}_{k\times 1} \end{bmatrix}_{(n+k)\times 1}$$

$$\boldsymbol{H}(\boldsymbol{w}, \boldsymbol{\lambda}) = \begin{bmatrix} \frac{\partial^2 L(\boldsymbol{w},\boldsymbol{\lambda})}{\partial w_1 \partial w_1} & \frac{\partial^2 L(\boldsymbol{w},\boldsymbol{\lambda})}{\partial w_1 \partial w_2} & \cdots & \frac{\partial^2 L(\boldsymbol{w},\boldsymbol{\lambda})}{\partial w_1 \partial w_n} \\ \frac{\partial^2 L(\boldsymbol{w},\boldsymbol{\lambda})}{\partial w_2 \partial w_1} & \frac{\partial^2 L(\boldsymbol{w},\boldsymbol{\lambda})}{\partial w_2 \partial w_2} & \cdots & \frac{\partial^2 L(\boldsymbol{w},\boldsymbol{\lambda})}{\partial w_2 \partial w_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial^2 L(\boldsymbol{w},\boldsymbol{\lambda})}{\partial w_n \partial w_1} & \frac{\partial^2 L(\boldsymbol{w},\boldsymbol{\lambda})}{\partial w_n \partial w_2} & \cdots & \frac{\partial^2 L(\boldsymbol{w},\boldsymbol{\lambda})}{\partial w_n \partial w_n} \end{bmatrix}_{n\times n}$$

$$\boldsymbol{N} = \begin{bmatrix} \frac{\partial G_0(\boldsymbol{w})}{\partial w_1} & \frac{\partial G_1(\boldsymbol{w})}{\partial w_1} & \cdots & \frac{\partial G_{k-1}(\boldsymbol{w})}{\partial w_1} \\ \frac{\partial G_0(\boldsymbol{w})}{\partial w_2} & \frac{\partial G_1(\boldsymbol{w})}{\partial w_2} & \cdots & \frac{\partial G_{k-1}(\boldsymbol{w})}{\partial w_2} \\ \vdots & \vdots & & \vdots \\ \frac{\partial G_0(\boldsymbol{w})}{\partial w_n} & \frac{\partial G_1(\boldsymbol{w})}{\partial w_n} & \cdots & \frac{\partial G_{k-1}(\boldsymbol{w})}{\partial w_n} \end{bmatrix}_{n\times k}$$

$$\boldsymbol{L}(\boldsymbol{w}, \boldsymbol{\lambda}) = \begin{bmatrix} \frac{\partial L(\boldsymbol{w},\boldsymbol{\lambda})}{\partial w_1} \\ \frac{\partial L(\boldsymbol{w},\boldsymbol{\lambda})}{\partial w_2} \\ \vdots \\ \frac{\partial L(\boldsymbol{w},\boldsymbol{\lambda})}{\partial w_n} \end{bmatrix}_{n\times 1} \quad \boldsymbol{G}(\boldsymbol{w}) = \begin{bmatrix} G_0 \\ G_1 \\ \vdots \\ G_{k-1} \end{bmatrix}_{k\times 1}$$

Numerical Examples

In this section, the developed JNN algorithm is applied to two examples. The proposed method was implemented using the Python programming language, with the aid of the SymPy and NumPy libraries. During the application of the method, the following considerations were considered:

- In the solutions, the default values of the alpha and beta hyperparameters were set to zero.
- The maximum number of iterations was set to 20.
- An early stopping mechanism was triggered when the weights were obtained identically for three consecutive iterations with a precision of 20 decimal digits.

•Let the exact solution of the problem be denoted by $y(x)$, and the numerical solution obtained at the n-th iteration by $N(x, \boldsymbol{w}) = N_n(x, \boldsymbol{w})$, over the problem domain $[a, b]$. The error function and the maximum absolute error are defined respectively as follows:

$$E_n = \| y(x) - N_n(x, \boldsymbol{w}) \|_\infty = max\{| N_n(x, \boldsymbol{w}) - y(x)|, a \leq x \leq b\} \quad (9)$$

Example 1: Nonlinear pantograph type delayed differential equation

$$y'(x) + 2y^2\left(\frac{x}{2}\right) = 1, \qquad x \geq 0 \quad (10)$$

with initial condition $y(0)= 0$ [15].

As shown in Table 1, the JNN algorithm demonstrates high accuracy in solving the nonlinear pantograph-type delayed differential equation presented in (10). The error values E_n decrease significantly as the number of neurons n increases, indicating enhanced approximation capability with higher network capacity.

Moreover, the early stopping criterion was satisfied at progressively earlier iterations with increasing c, showing that convergence was achieved more rapidly when more neurons were utilized. Specifically, for n=12, the method reached the stopping condition at the 6th iteration, achieving a remarkably low error of 1.85 e - 15, which highlights the method's efficiency and precision.

Table 1: Numerical Results and Early Stopping Iterations for Example 1

n	8	10	12
E_n	1.85 e - 9	2.27 e - 12	1.85 e - 15
Early stopping	$15th$	$8th$	$6th$

Figure 2 illustrates the absolute error $e(x)=|y(x)-N(x,\boldsymbol{w})|$ between the exact solution and the numerical approximation obtained using the JNN algorithm with $n=8$ neurons. The error function remains on the order of e-9 throughout the domain [0,1], demonstrating the method's high accuracy even with a relatively low number of neurons.

Although small oscillations are observed, the error exhibits a generally smooth and bounded behaviour, with the maximum error occurring near the left boundary. This behaviour reflects the model's robustness in approximating delayed nonlinear dynamics with acceptable numerical stability.

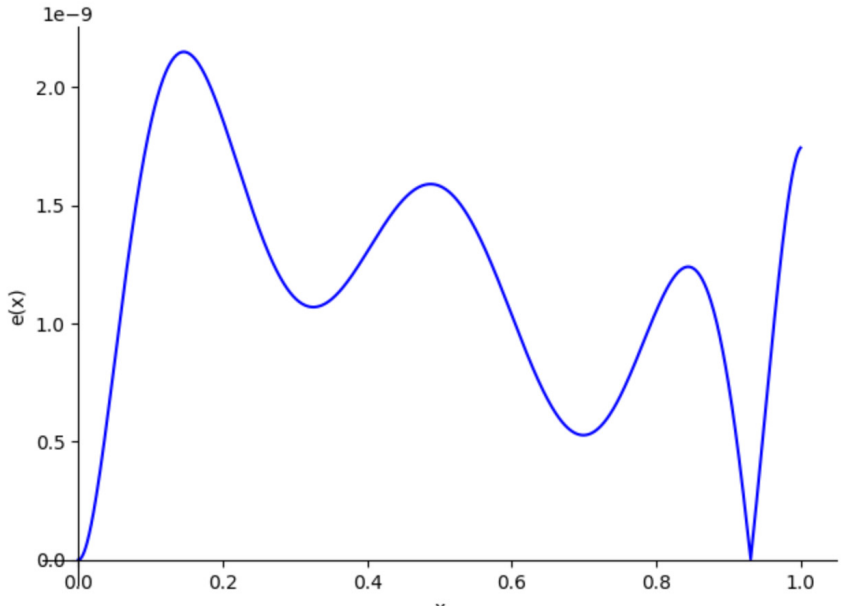


Figure 2: Absolute error function $e(x)$ for Example 1 with $n=8$ over the domain [0,1].

For $n=8$, the JNN method yields a 7th-degree polynomial solution. The maximum absolute error obtained using this method is $1.80e$-9. This result is significantly more accurate than those reported in the literature for the same polynomial degree. In comparison, another study employing the PIA(1,1) algorithm of the same order reported a maximum absolute error of $1.23e$-7 [6]. Furthermore, it was observed in a related study that the absolute maximum error obtained via the Homotopy Analysis Method (HAM) was $1.2e$-6, while the Optimal Homotopy Analysis Method (OHAM) reduced this to $4e$-8[15]. These comparisons clearly demonstrate that the proposed JNN method achieves the highest accuracy among the referenced approaches.

Example 2: Second order nonlinear pantograph type delayed differential equation

$$y''(x) - (|y(x)|^2 + |y(x)|^3)y\left(\frac{x}{2}\right) = 0, \qquad x \geq 0 \tag{11}$$

with boundary condition $y(0)=1$ and $y(1) = 1/2$ [16].

Table 2 presents the numerical results obtained using the JNN method for solving the second-order nonlinear pantograph-type delayed differential equation given in Example 2. The table shows the error values E_n corresponding to different neuron counts n=8, 10, 12, as well as the iteration number at which early stopping occurred. It is observed that as the number of neurons increases, the error E_n decreases significantly. Specifically, the error reduces from 7.13e-7 for n=8 to 1.10e-8 for n=12. Notably, early stopping was consistently achieved at the 5-th iteration across all tested values of n, indicating both the stability and efficiency of the JNN method in handling delayed nonlinear problems with high accuracy.

Table 2: Numerical Results and Early Stopping Iterations for Example 2

n	8	10	12
E_n	7.13 e - 7	3.08 e - 8	1.10 e - 8
Early stopping	$5th$	$5th$	$5th$

Conclusions

In this study, a Jacobi Neural Network (JNN) framework was introduced for solving pantograph-type functional differential equations characterized by proportional delay. The proposed method integrates Jacobi polynomials as activation functions in the hidden layer of a feed-forward neural network, enabling effective approximation of solutions under different initial and boundary conditions. The numerical results of two illustrative examples confirm the accuracy and efficiency of the JNN approach. In the first example, the method yielded a significantly lower maximum absolute error compared to established techniques such as PIA(1,1), HAM, and OHAM. In the second example, the robustness and convergence of the method were demonstrated by a consistent reduction in error as the number of neurons increased, along with stable early stopping behaviour.

Overall, the JNN method offers a highly accurate and computationally stable alternative for solving both linear and nonlinear pantograph-type delay differential equations. Its adaptability and precision suggest significant potential for applications across a wide range of scientific and engineering problems involving functional differential equations.

Declaration of Interest: The authors declare that there is no conflict of interest.

Acknowledgements: The authors acknowledge the Organizing and Scientific Committees of ICADA 2025 for their valuable efforts in organizing the conference.

Author Contributions: M. M. Bahşı conceived the idea of the study, implemented the JNN method, carried out the coding process, and wrote the manuscript. A. Kurt Bahşı defined the mathematical steps of the method, adapted the optimization technique to the JNN method, conducted the comparative analysis of the results, and contributed to the editing of the manuscript.

REFERENCES

[1] A.B. Albidah, N.E. Kaanan, A. Ebaid and H.K. Al-Jeadid, "Exact and numerical analysis of the pantograph delay differential equation via the homotopy perturbation method", Mathematics, 11.4, 944, 2023.

[2] R. Alrebdi, and H. K. Al-Jeaid, "Accurate solution for the Pantograph delay differential equation via Laplace transform", Mathematics 11.9, 2031, 2023.

[3] N. Sriwastav, and A.K. Barnwal, "Numerical solution of Lane-Emden pantograph delay differential equation: stability and convergence analysis", International J. of Mathematical Modelling and Numerical Optimisation, 13.1, 64-83, 2023.

[4] M.M. Bahşı, M. Çevik, and M. Sezer, "Orthoexponential polynomial solutions of delay pantograph differential equations with residual error estimation", Applied Mathematics and Computation, 271, 11-21, 2015.

[5] D. Methi, A. Kumar, and R. Aggarwal, "Analysis of nonlinear pantograph delay differential equations using differential transform technique", AIP Conference Proceedings, 3191, 1, 2025.

[6] M.M. Bahşi, and M. Çevik. "Numerical solution of pantograph-type delay differential equations using perturbation-iteration algorithms", J. of Applied Mathematics, 2015.1, 139821, 2015.

[7] C. Yang, J. Hou, and X. Lv, "Jacobi spectral collocation method for solving fractional pantograph delay differential equations", Engineering with Computers, 38.3, 1985-1994, 2022.

[8] B. Thakur, and S. Gupta, "An Iterative Algorithm for Numerical Solution of Nonlinear Fractional Differential Equation using Legendre Wavelet Method", IAENG International J. of Applied Mathematics, 54.3, 380-389, 2024.

[9] R. Alrebdi, and H.K. Al-Jeaid. "Two Different Analytical Approaches for Solving the Pantograph Delay Equation with Variable Coefficient of Exponential Order", Axioms, 13.4, 229, 2024.

[10] W.M. Abd-Elhameed, and M.M. Alsuyuti, "New spectral algorithm for fractional delay pantograph equation using certain orthogonal generalized Chebyshev polynomials", Communications in Nonlinear Science and Numerical Simulation, 141, 108479, 2025.

[11] B.Y. Guo, J. Shen, and LL Wang, "Generalized Jacobi polynomials/functions and their applications", Applied Numerical Mathematics, 59.5, 1011-1028, 2009.

[12] M.M. Bahşı, and M. Çevik, "Improved Jacobi Matrix Method for Solving Multi-Functional Integro-Differential Equations with Mixed Delays", Celal Bayar University Journal of Science, 16.4, 393-401, 2020.

[13] D.P. Bertsekas, Constrained optimization and Lagrange multiplier methods, Academic Press, 2014.

[14] K.D. Dwivedi, and J.F. Gómez-Aguilar, "An efficient numerical method to solve ordinary differential equations using Fibonacci neural networks", Computational and Applied Mathematics, 42.1, 54, 2023.

[15] N. Ratib Anakira, A.K. Alomari, and I Hashim, "Optimal homotopy asymptotic method for solving delay differential equations", Mathematical Problems in Engineering 2013.1, 498902, 2013.

[16] M.A.Z, Raja, "Numerical treatment for boundary value problems of pantograph functional differential equation using computational intelligence algorithms", Applied Soft Computing, 24, 806-821, 2014.

OPTIMIZING URBAN INDOOR FARMING: UV LIGHT AND AUTOMATED IRRIGATION FOR BOK CHOY CULTIVATION

Carlo R. Hadjirul*
Department of Computer Science, Western Mindanao State University, Zamboanga City, Philippines

Jopel E. Enriquez
Department of Computer Science, Western Mindanao State University, Zamboanga City, Philippines

Ghaleeb Kangiluhan
Department of Computer Science, Western Mindanao State University, Zamboanga City, Philippines

Aradzna M. Kamman
Department of Computer Science, Western Mindanao State University, Zamboanga City, Philippines
(Corresponding author) qb202101178@wmsu.edu.ph

ABSTRACT

This study investigates the efficiency and effectiveness of an automated urban indoor farming system for Bok Choy cultivation, incorporating UV light and automated irrigation, and compares its performance with traditional farming methods. The study involved 9 Bok Choy plants per setup, conducted in a controlled environment where soil quality, plant spacing, and environmental conditions were maintained consistently. The automated system optimized light exposure through UV lighting and regulated irrigation using soil moisture sensors, while the traditional system relied on manual watering and monitoring. Results indicated that after 2 weeks, the automated system led to an average of 5 leaves per plant, whereas the traditional system produced an average of 3 leaves. Additionally, the automated system maintained stable soil moisture levels, reducing water waste by approximately 20%, and minimized environmental stress, leading to more uniform growth. A t-test comparing the growth rates of both setups showed no significant difference in leaf growth ($p = 0.157$), indicating that leaf development between indoor and traditional farming was similar. However, a t-test conducted on the survival rates revealed a statistically significant difference ($p = 0.008$), with the indoor system achieving a 100% survival rate compared to the 77.77% survival rate in the traditional setup by Week 6. These findings suggest that while both systems produce similar growth rates, the automated indoor system offers superior pest management, resource efficiency, and plant survival. The results highlight the potential of automated farming systems to enhance plant growth, reduce labor, and support sustainable agricultural practices in urban settings. Further research is needed to assess the long-term scalability, cost-efficiency, and broader applicability of automated farming technologies.

Keywords: Automated urban farming, UV light optimization, automated irrigation, bok choy cultivation, sustainable agriculture.

Introduction

Urbanization has driven the demand for sustainable agricultural solutions capable of optimizing efficiency within confined urban spaces. Indoor farming has emerged as a promising avenue, providing controlled environments suitable for year-round crop cultivation [1]. Bok Choy, known for its rapid growth, nutritional richness, and culinary versatility, has become a staple in many cuisines [2]. However, successful indoor cultivation of Bok Choy requires precise management of environmental variables such as light exposure and soil moisture levels.

Traditional farming methods often struggle to maintain stable growing conditions indoors, leading to suboptimal yields and resource inefficiencies [3]. Furthermore, the risk of pest outbreaks in an uncontrolled environment presents a constant threat to Bok Choy crops. To address these challenges, integrating automated technologies has become crucial. This study aims to develop an automated urban indoor farming system specifically tailored for Bok Choy cultivation. The system consists of two core components: automated UV light control and soil moisture detection.

Bok Choy faces several critical challenges that impact its productivity and quality. Among these challenges is its vulnerability to various pests and diseases. In the absence of modern pest management strategies and disease-resistant varieties, Bok Choy is often susceptible to outbreaks that can devastate crops and reduce yields. Traditional farming methods may not offer sufficient control over these threats, resulting in economic losses for farmers who are dependent on steady harvests. Addressing these challenges is essential for sustaining Bok Choy production in conventional agricultural settings.

Automated UV light control ensures that Bok Choy receives optimal light exposure throughout its growth cycle, enhancing photosynthesis and promoting robust plant development. By programming the system to adjust UV light levels according to predetermined schedules, energy consumption can be optimized while mimicking natural day/night cycles. Additionally, the system incorporates soil moisture detection mechanisms, which trigger an irrigation system when the soil becomes dry. This real-time monitoring and response mechanism ensures consistent hydration levels, promoting healthy plant growth and minimizing water waste.

Related Studies

Urban agriculture, empowered by technological advancements, offers a compelling solution to address food security concerns in densely populated areas. Automated urban indoor farming systems are gaining increasing recognition for their potential to optimize crop production within controlled environments. This contrasts sharply with the challenges faced by traditional outdoor farming methods, which often struggle with unpredictability, climate change, and resource inefficiencies.

Climate change can impact agricultural practices, particularly vegetable production

in the Philippines. Changes in temperature and precipitation patterns have a detrimental effect on crop growth, especially in outdoor farming systems. Extreme weather events such as typhoons and droughts pose a severe risk to crops like Bok Choy, which can be adversely affected by water stress and fluctuations in temperature [4]. This research aligns with [5], which emphasized the vulnerability of vegetables like Bok Choy to climate-induced stresses, underlining the importance of shifting to controlled environments where conditions can be optimized.

In line with promoting more sustainable agricultural practices, the Department of Agriculture - Bureau of Plant Industry (DA-BPI) has been instrumental in advancing research on indoor farming methods, including the use of UV light for vegetable cultivation. Their support of agricultural innovation in the Philippines has the potential to reduce reliance on traditional farming, which is often constrained by adverse weather events, while also improving crop yields in urban settings. Budavári et al.'s work on controlled environment agriculture (CEA) further highlights the role of artificial lighting in indoor farming systems, showing how optimized light exposure can boost plant health and productivity [6]. The use of UV light in these systems, explored by Dr. Murat Kacira, has proven to enhance both resource efficiency and crop yield. Kacira's contributions to the development of vertical farming systems and lighting technologies underscore the key benefits of CEA, including enhanced food security and reduced environmental impact [7].

The Philippine Council for Agriculture, Aquatic, and Natural Resources Research and Development (PCAARRD) also plays a pivotal role by funding research on indoor farming technologies. Studies that focus on optimizing the use of UV lights in these systems can further drive progress toward more sustainable and resilient agricultural systems in urban environments. Research by Dr. Toyoki Kozai on plant physiology and growth under artificial lighting, particularly LED systems, has helped clarify the mechanisms of plant growth in controlled environments. His work has enabled farmers to fine-tune lighting strategies that maximize crop productivity, particularly in vertical farms [8].

The significant findings from [9], who studied the impact of extended light exposure on leafy greens like Bok Choy, align with these advancements. Their study demonstrated that providing 14-16 hours of light per day substantially increased growth and biomass in leafy greens, offering insights into the critical role that light exposure plays in optimizing crop yields in controlled settings.

Moreover, Muniappan and Heinrichs [10] contributed to understanding the pest and disease challenges faced by vegetable growers, including Bok Choy farmers in Asia. They identified key pests such as aphids and caterpillars that threaten Bok Choy crops, particularly in outdoor farming systems. The risk of pests and diseases can be minimized in controlled environments, where automated monitoring systems can prevent outbreaks and protect crop health.

Overall, the contributions of these researchers have significantly advanced our understanding of indoor farming technologies, particularly the use of UV lights for crop production. Their findings provide valuable insights into optimizing lighting strategies, thereby enhancing crop yield, quality, and resource efficiency in indoor farming systems.

Methodology

Design

The diagram in Figure 1 depicts the components of an automated indoor farming system. The central unit is the Arduino Uno, which controls the system. The LED UV lights are powered through a breadboard and connected to the Arduino for regulated lighting. A moisture sensor monitors the soil's moisture levels, sending data to the Arduino. If the soil is dry, the system activates the mini sprinkler connected to a water pump, which draws water from a water container to irrigate the plants. Additionally, an LCD screen displays the system's data, such as the moisture level and other relevant parameters. This setup automates both lighting and watering, optimizing plant care in urban farming environments.

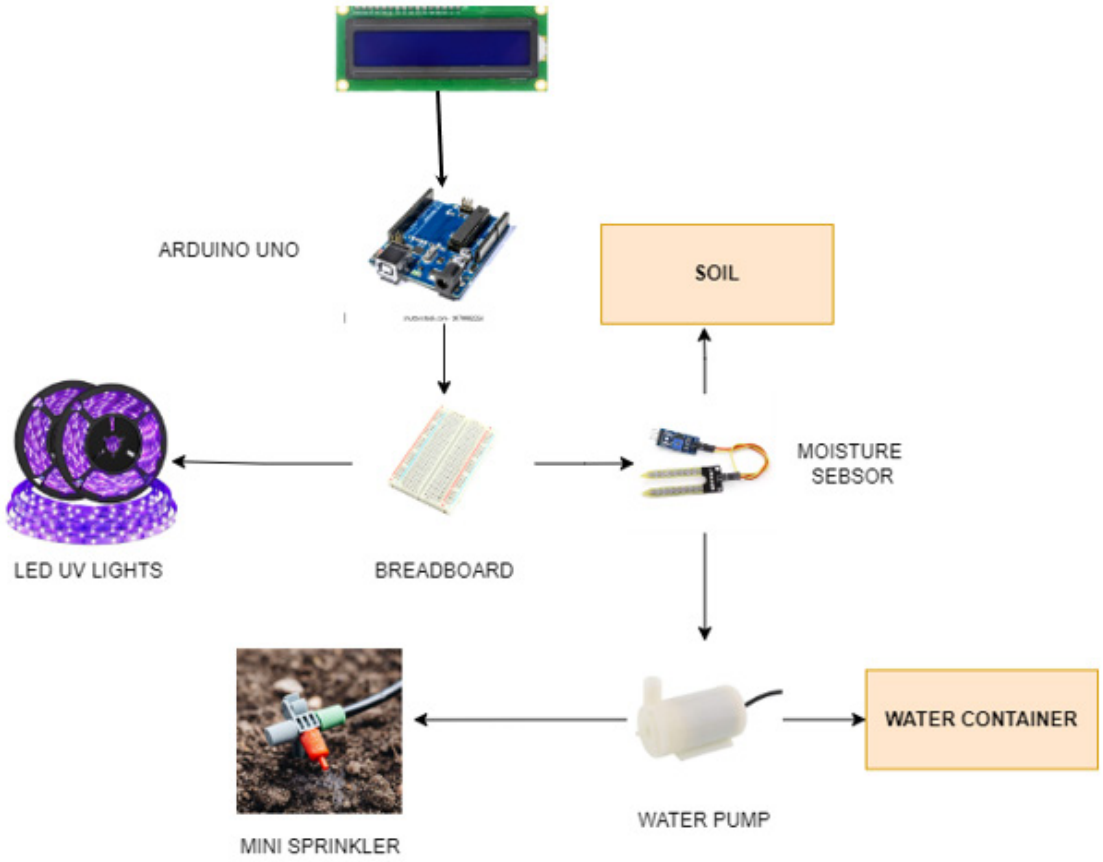


Figure 1: System Component

The flowchart in Figure 4 outlines the operation of an automated system that controls UV light and watering based on plant needs. Upon being plugged into the electrical outlet, the system is managed by a microcontroller or control unit. It checks if the UV light has been on for 15 hours—if so, it turns off for 9 hours, then reactivates. Simultaneously, the system checks soil moisture; if the soil is dry, the automated watering system is activated and continues until the soil reaches the desired moisture level. If the conditions are not met, no actions are taken, and the system waits for the necessary changes to occur. This process ensures optimized light exposure and hydration for plant growth.

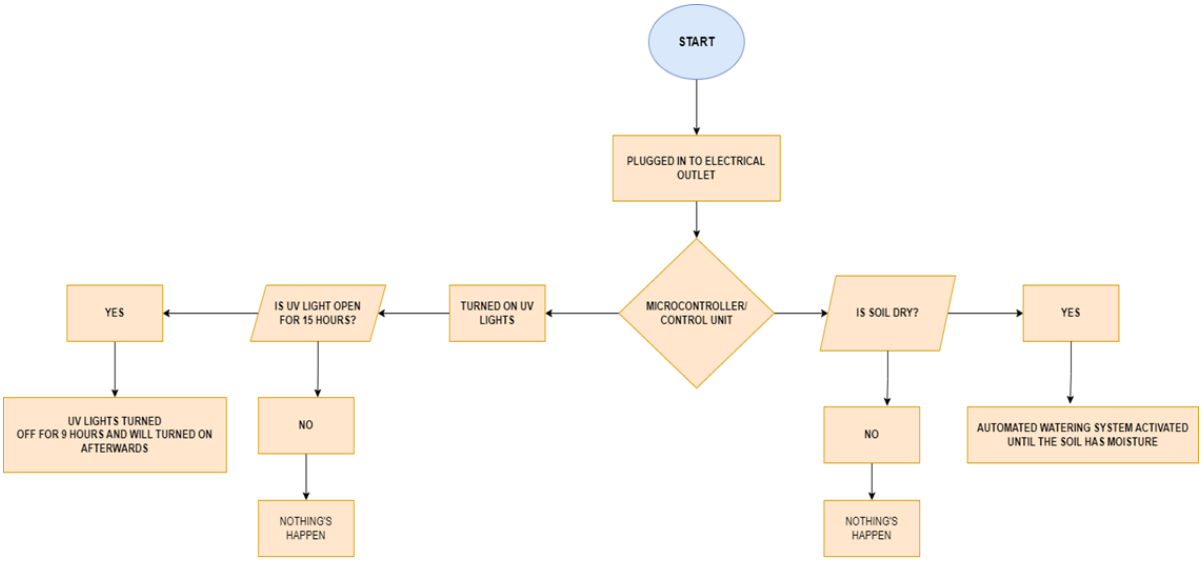


Figure 4: Process for controlling the UV light and automated watering system

The proposed frame in Figure 3 measures 36 inches in width and 28 inches in length. The materials used to construct the frame are 100% recycled, sourced from an old wooden cabinet.

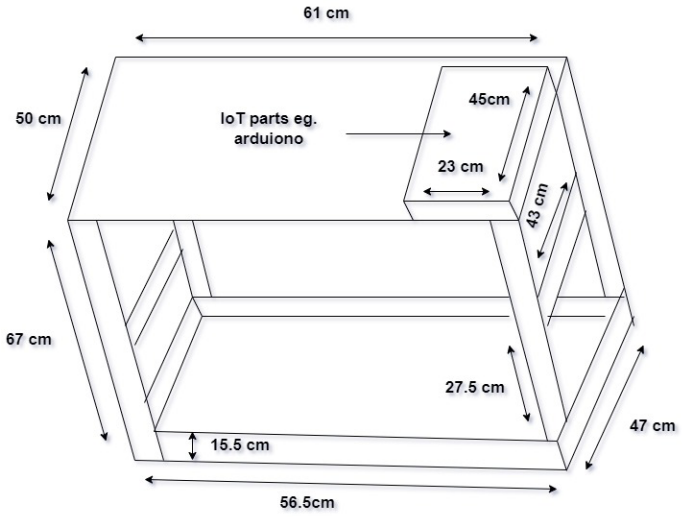


Figure 3: Bok Choy frame

Experimental Setup and Procedure

This study follows an experimental research design aimed at evaluating the impact of an automated indoor farming system in comparison to traditional farming methods for Bok Choy cultivation. The experiment is conducted in a controlled environment, where Bok Choy plants are grown under two conditions: one using an automated setup that incorporates IoT-based UV lighting and soil moisture monitoring, and the other using a traditional setup based on manual irrigation and observation. Both setups maintain consistent soil composition, plant spacing, and environmental conditions to ensure accurate comparisons. Key growth parameters, such as leaf count, plant height, survival rate, soil moisture levels, and pest prevalence, are systematically recorded and analyzed. The study aims to assess

the extent to which automation enhances crop growth and resource efficiency, contributing to the development of sustainable urban farming practices. This study planted 9 Bok Choy plants per setup to ensure optimal spacing and resource distribution.

Both the automated indoor farming and traditional setups in Figure 4 were provided with approximately 10-12 kg of soil per container, ensuring uniformity in soil type and quality. The plants were spaced evenly, with a 3 to 4-inch gap between each one. In the automated indoor farming setup, the 9 Bok Choy seedlings were planted and connected to an IoT-controlled system for automated irrigation. After planting, the system was activated, and the automated watering system was triggered due to initial soil dryness. The moisture detector continuously monitored soil moisture levels, activating the watering system only when necessary to maintain optimal moisture. This process was repeated, ensuring consistent hydration without excess water usage.

Figure 4: (a) Automated indoor farming (b) traditional farming bed frame

Additionally, the indoor environment was maintained at a moderate temperature, as Bok Choy is sensitive to cold conditions, which can slow growth or, in severe cases, cause plant death. Studies by Zhang et al. [15] indicate that temperatures below 15°C (59°F) can significantly slow Bok Choy's physiological processes, reducing photosynthesis rates and hindering nutrient absorption. In contrast, keeping the temperature within the optimal range of 18°C to 24°C (65°F to 75°F) promotes faster cell division and healthy leaf expansion, both essential for robust growth and high yields. Ensuring a warm and steady indoor environment is therefore crucial to supporting Bok Choy's rapid growth cycle and increasing productivity.

For the traditional setup, the 9 Bok Choy seedlings were planted with the same spacing and soil conditions. Water was applied manually, and the soil was carefully monitored

to maintain proper moisture levels. Watering was carried out periodically, and additional monitoring ensured no environmental disruptions, such as animal disturbances, that could affect plant growth. This setup adjustment facilitated a controlled and reliable comparison between the two methods, focusing on resource effectiveness, growth measurements, and plant health outcomes.

Bok Choy will follow a development cycle that can range from 80 to 100 days, depending on soil conditions. High-quality soil with sufficient organic matter, proper drainage, and a balanced pH of 6.0 to 7.5 is essential for maximizing yield within this timeframe. According to Ng et al. [16], nutrient-rich soil promotes faster root development and robust leaf growth, enabling Bok Choy to reach maturity on the shorter end of the growth spectrum. Conversely, suboptimal soil conditions may delay growth and extend the harvest period. This underscores the importance of maintaining consistent soil quality across both setups to ensure valid comparisons.

Statistical Tools

To assess the differences between the indoor and traditional farming setups, a t-test was conducted using (1). The t-test was used to determine if there were statistically significant differences in leaf growth and survival rates between the two farming methods. The t-test results for leaf growth indicated no significant difference, while the survival rate data showed a statistically significant difference, with the indoor setup demonstrating better survival rates.

$$t = \frac{\overline{x_1} - \bar{x}_2}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}} \quad (1)$$

where

- t = Student's t-test
- x_1 = mean of first group
- $\bar{x}_2$ = mean of second group
- s_1^2 = standard deviation of group 1
- s_2^2 = standard deviation of group 2
- n_1= number of observations in group 1
- n_2 = number of observations in group 2

Results and Discussion

Comparative Analysis

By Day 34 to 50, the Bok Choy, as shown in Figure 5 in the indoor setup, showed positive results, with all nine plants surviving and producing an average of 6 to 11 leaves per plant. There were no signs of pest or insect damage, thanks to the controlled environment. However, two of the Bok Choy plants remained smaller in size compared to the others, likely due to minor discrepancies in light exposure or initial planting conditions.

In the traditional setup, growth was more irregular. Four of the nine Bok Choy plants did not match the growth of the other five, which developed comparably well. A negligible trace of pesticides was observed, and the average leaf count ranged from 5 to 9 leaves

per plant. However, the growth of the four plants was smaller, likely caused by environmental stressors such as inconsistent soil quality, weather fluctuations, or pest pressure. This comparison underscores the benefits of indoor farming in providing a more stable and supportive environment for uniform plant growth.

Figure 5: Day 35 of comparison between indoor and traditional

Over the course of the cultivation period as tabulated in Table 1 and as shown in Figure 6, significant differences were observed between the indoor and traditional farming setups for Bok Choy. In the indoor setup, all plants survived during Week 1 with 2 leaves each, showing no signs of pesticide exposure and a 100% survival rate. By Week 2, the survival rate remained high at 100%, with plants developing between 2 and 5 leaves. By Weeks 3 to 6, the plants showed steady growth, ranging from 4 to 11 leaves, and maintained 100% survival throughout. There were no signs of pests or disease, and the plants consistently exhibited a healthy green color.

Table 1: Bok Choy Growth Analysis between Indoor and Traditional

Week	Sign of Pesticides (%)		Avg. Leaf Number per Bok Choy		Survival Rate (%)	
	Indoor	Traditional	Indoor	Traditional	Indoor	Traditional
1	0%	0%	2	2	100%	100%
2	0%	0%	4	2	100%	88.90%
3	0%	0%	4.75	3	100%	77.77%
4	0%	0%	6.5	4	100%	77.77%
5-6	0%	4%	8.5	5	100%	77.77%

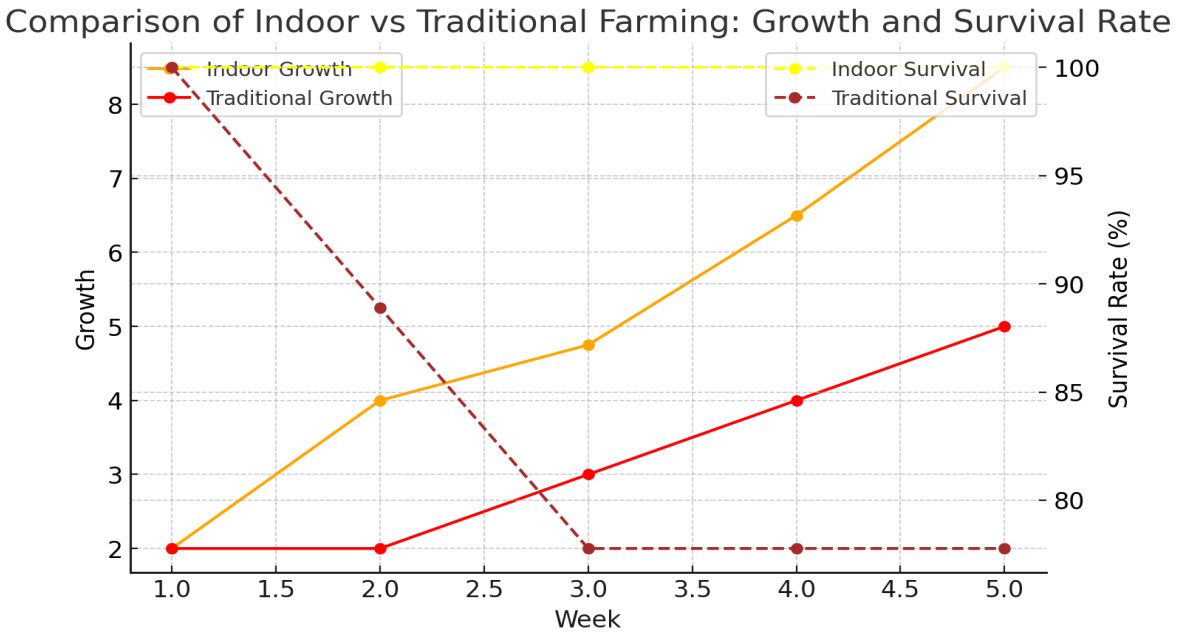


Figure 6: Indoor vs Traditional Farming in Terms of Growth and Survival

In contrast, the traditional setup experienced some challenges. Week 1 saw a similar outcome with all plants surviving and developing 2 leaves each, but by Week 2, the survival rate dropped to 88.9%, with some plants showing only 2 or 3 leaves. By Week 3, the growth was more irregular, with some plants developing 4 to 7 leaves, but the survival rate continued to decrease to 77.77% by Week 4. The growth rate varied even further, with some plants showing fewer leaves or smaller sizes, likely due to environmental stressors such as inconsistent watering or pest pressure. A slight 4% pesticide presence was noted by Weeks 5-6, further impacting the plants' health.

The t-test analysis comparing Bok Choy growth and survival rates in the indoor and traditional setups yielded notable results. For leaf growth, the t-statistic was 1.56, and the p-value was 0.157, indicating that there is no statistically significant difference between the two setups, as the p-value exceeds the typical threshold of 0.05. This suggests that the leaf development between the indoor and traditional farming methods is not significantly different. However, when comparing the survival rates, the t-statistic was 3.50 with a p-value of 0.008, which is below the 0.05 threshold, indicating a statistically significant difference. The indoor setup demonstrated a higher survival rate, suggesting that the controlled environment provided by indoor farming supports better plant health and resilience compared to traditional farming methods.

This comparison highlights the benefits of the controlled indoor environment, which promoted more consistent growth, a higher survival rate, and healthier plants. The traditional farming setup, however, showed more variability in growth, with a lower survival rate and signs of pest damage, emphasizing the importance of a stable and controlled environment for optimal Bok Choy cultivation.

Leaf Growth Analysis

The number of leaves per Bok Choy plant over time follows a linear growth model computed using (2):

$$L = L^0 + rt \quad (2)$$

where:

- L = number of leaves at time t(days)
- L^0 = initial number of leaves (2 leaves at t=0)
- r= average leaf growth rate (leaves per day)
- t = time in days

Using the test results, the final number of leaves at Day 50 was 11 for indoor farming and 8 for traditional farming. By substituting these values into (1), the leaf growth rate for indoor farming was calculated as 0.18 leaves per day. Meanwhile, for indoor farming, 0.12 leaves per day are produced. These results indicate that indoor farming exhibits a 50% faster leaf growth rate compared to traditional farming, demonstrating the efficiency of controlled environmental conditions in accelerating plant development.

Survival Rate Difference

To measure the difference in survival rates over time, use (3).

$$SRD = SRI - SRT \quad (3)$$

Indoor farming had a 22.3% higher survival rate than traditional farming, demonstrating its effectiveness in providing a more stable growing environment.

Pest Incidence Analysis

The Pest Incidence Analysis was conducted using categorical and trend analysis methods, where the presence or absence of pests was measured as a categorical variable. In Indoor Farming, pest incidence remained consistently at 0%, indicating a stable trend with no pest occurrences throughout the study. In contrast, Traditional Farming experienced pest incidence starting from Weeks 5-6, reaching 4%, suggesting an increasing risk of pest presence over time as environmental conditions fluctuated.

Leaf Color

The Leaf Color Analysis classified leaf color as a qualitative categorical variable. Both farming setups consistently exhibited green leaves throughout the study, with no observed variation. Since there was no change in leaf color, it was treated as a constant categorical variable rather than numerical data, indicating that leaf color did not significantly differ between the two farming methods during the study period.

Comparing Automated Indoor Farming to Traditional Bok Choy Farming

The study was prompted by the absence of Bok Choy farms in the Zamboanga area. To address this gap, the researcher expanded the inquiry by using Facebook as a platform to connect with Bok Choy farm owners and farmers. A series of structured questions were developed and shared to collect relevant data and insights on Bok Choy cultivation practices, challenges, and market dynamics. This approach enabled the researcher to bridge the geographical gap and gather valuable information from external sources, enriching the study's findings. Below is a comparison of images between automated indoor Bok Choy farming and traditional Bok Choy farming over a 60 to 65-day time frame.

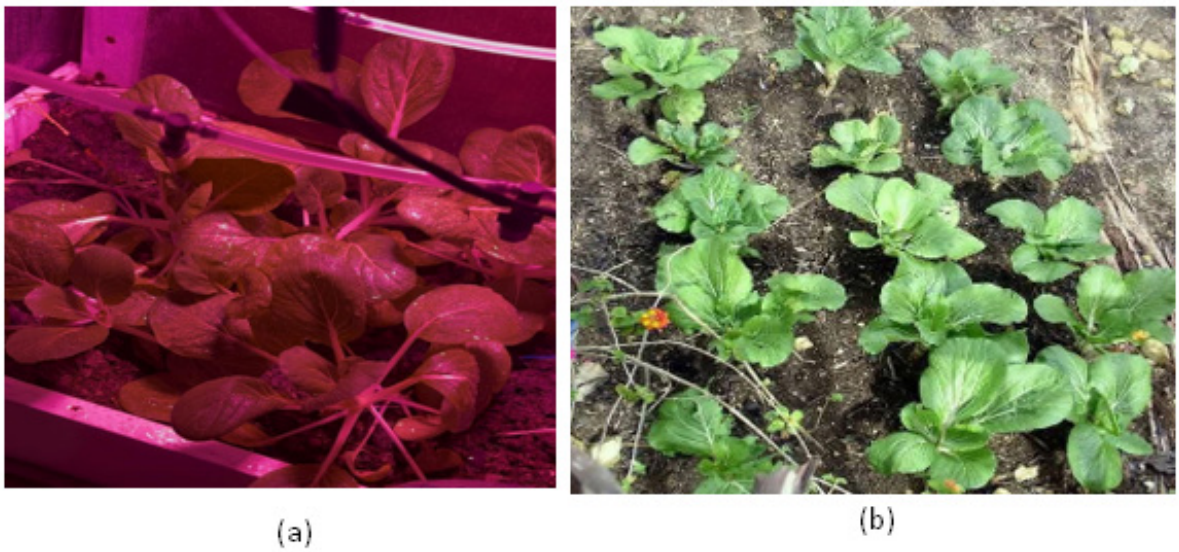


Figure 7: (a) Automated Indoor (b) Traditional Farm

Figure 7 depicts Bok Choy cultivated in two different agricultural settings: a large-scale traditional farm and an automated indoor environment utilizing UV light. In the indoor farming system, the Bok Choy stems appear slightly elongated compared to those grown in traditional farming. However, the leaves in both setups are broad and substantial. Notably, signs of pest infestation are visible on some Bok Choy plants in the traditional farming system, while no such signs are seen in the indoor farming setup. This distinction underscores the controlled environment of indoor farming, which minimizes exposure to external biotic stressors, such as pests and diseases.

Traditional farming employs eco-friendly practices, such as the use of organic fertilizers to enhance growth rates, while indoor farming operates without fertilizers, relying instead on optimized light, temperature, and humidity conditions to foster plant growth. Despite these differences, both systems yield Bok Choy of comparable size. However, indoor farming demonstrates a slight advantage in pest management and leaf size, which can be attributed to the absence of external environmental stressors and the precise distribution of resources in controlled settings. These observations suggest that while traditional farming remains a practical method for Bok Choy production, indoor farming presents potential benefits in terms of uniformity, pest control, and resource efficiency, warranting further exploration in future studies.

Conclusion

This study investigated the effectiveness of an automated urban indoor farming system utilizing UV light and automated irrigation to optimize Bok Choy cultivation compared to traditional farming methods. The results demonstrate that the indoor farming system significantly improved plant growth, pest management, and resource efficiency. The controlled environment of the indoor system contributed to a faster growth rate, with plants showing more consistent development and fewer pest-related issues. In contrast, traditional farming, while still effective, faced challenges such as inconsistent growth rates, pest infestations, and higher resource wastage. These findings highlight the potential of automated systems to enhance the productivity and sustainability of urban agriculture, offering a viable solution to food security challenges in densely populated areas. Despite the promising results, further studies are needed to assess the long-term viability, scalability, and economic feasibility of the automated system in different urban contexts.

REFERENCES

[1] K. Benke and B. Tomkins, "Future food-production systems: vertical farming and controlled-environment agriculture," Sustainability: Science, Practice and Policy, vol. 13, no. 1, pp. 13–26, Jan. 2017, doi: 10.1080/15487733.2017.1394054.

[2] E. Naseri et al., "Bok Choy promotes growth performance, lipid metabolism and related gene expression in Syrian golden hamsters fed with a high-fat diet," Food Funct, vol. 11, no. 3, pp. 2693–2703, 2020, doi: 10.1039/C9FO02975C.

[3] G. Dauchot, C. Aubry, A. Crème, E. Dorr, and B. Gabrielle, "Energy consumption as the main challenge faced by indoor farming to shorten supply chains," Cleaner and Circular Bioeconomy, vol. 9, p. 100127, Dec. 2024, doi: 10.1016/j.clcb.2024.100127.

[4] D. Velasco et al., "Biodiversity conservation research challenges in the 21st century: A review of publishing trends in 2000 and 2011," Environ Sci Policy, vol. 54, pp. 90–96, Dec. 2015, doi: 10.1016/j.envsci.2015.06.008.

[5] S. S. Solankey, P. Singh, P. Shukla, and R. Kumar, "Vulnerability of Vegetable Crops to Climate Change," Apr. 2019.

[6] N. Budavári, Z. Pék, L. Helyes, S. Takács, and E. Nemeskéri, "An Overview on the Use of Artificial Lighting for Sustainable Lettuce and Microgreens Production in an Indoor Vertical Farming System," Horticulturae, vol. 10, no. 9, p. 938, Sep. 2024, doi: 10.3390/horticulturae10090938.

[7] A. P. Montoya, M. Kacira, and F. A. Obando, "Design and implementation of a low cost microcontroller in controlled environment agriculture," Acta Hortic, no. 1279, pp. 287–294, Jun. 2020, doi: 10.17660/ActaHortic.2020.1279.41.

[8] E. Hayashi and T. Kozai, "Plant factories with artificial lighting (PFALs): Concluding remarks," 2023, pp. 483–486. doi: 10.19103/AS.2023.0126.31.

[9] J. D. Stamford, J. Stevens, P. M. Mullineaux, and T. Lawson, "LED Lighting: A Grower's Guide to Light Spectra," HortScience, vol. 58, no. 2, pp. 180–196, Feb. 2023, doi: 10.21273/HORTSCI16823-22.

[10] R. Muniappan and E. A. Heinrichs, Eds., Integrated Pest Management of Tropical Vegetable Crops. Dordrecht: Springer Netherlands, 2016. doi: 10.1007/978-94-024-0924-6.

[11] C. Su and Z. Wang, "Research on Intelligent Control System of Indoor Greening Based on Wireless Sensor Network," J Sens, vol. 2022, pp. 1–10, Feb. 2022, doi: 10.1155/2022/9741906.

[12] P. R et al., "Human-Computer Interaction: Enhancing User Experience in Interactive Systems," E3S Web of Conferences, vol. 399, p. 04037, Jul. 2023, doi: 10.1051/e3sconf/202339904037.

[13] P. Mondaca, P. Valenzuela, W. Quiroz, M. Valdenegro, S. Abades, and J. L. Celis-Diez, "Environmental conditions and plant physiology modulate Cu phytotoxicity in field-contaminated soils," Ecotoxicol Environ Saf, vol. 246, p. 114179, Nov. 2022, doi: 10.1016/j.ecoenv.2022.114179.

[14] S. Y. Nof, "Automation: What It Means to Us Around the World," in Springer Handbook of Automation, Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 13–52. doi: 10.1007/978-3-540-78831-7_3.

[15] C. Dong et al., "Physiological and Transcriptomic Responses of Bok Choy to Heat Stress," Plants, vol. 13, no. 8, p. 1093, Apr. 2024, doi: 10.3390/plants13081093.

[16] W. W. Ng, R. Tasnim, V. Capobianco, and J. L. Coo, "Influence of soil nutrients on plant characteristics and soil hydrological responses," Géotechnique Letters, vol. 8, no. 1, pp. 19–24, Mar. 2018, doi: 10.1680/jgele.17.00104

PERFORMANCE ANALYSIS OF A GRAPHENE-COATED FRACTAL MICROSTRIP PATCH ANTENNA FOR 5G APPLICATIONS

Beyza Nur Demirkoparan
Faculty of Engineering and Natural Sciences, Department of Electrical and Electronics Engineering, Sivas University of Science and Technology, Sivas, Türkiye

Husnu Deniz Basdemir*
Faculty of Aviation and Space Sciences, Department of Electrical and Electronics Engineering, Sivas University of Science and Technology, Sivas, Türkiye
**(Corresponding author) denizbasdemir@gmail.com*

ABSTRACT

In this study, a graphene-based fractal structured 5G microstrip patch antenna design is proposed. In recent years, research on microstrip patch antennas for 5G technology has advanced rapidly, with increasing focus on the use of novel materials to enhance performance. Owing to its exceptional electromagnetic, mechanical, electrical, and thermal properties, graphene has been selected as the coating material for the antenna. The coating process was carried out using the Electrophoretic Deposition (EPD) method, and the details of this process along with the experimental stages are presented in this study. The fractal-structured antenna aims to achieve higher efficiency by leveraging the high conductivity and low-loss characteristics of graphene. During the design phase, the electromagnetic performance of the antenna was theoretically calculated, and modelling and simulation studies were performed in a computer environment. Prototype production of the antenna has been completed, and the graphene coating process is currently ongoing, with analysis studies already initiated. In future work, following the completion of the antenna's physical fabrication, comprehensive analyses will be conducted. After the coating process, electromagnetic and electrical tests will be carried out, and the results will be compared with theoretical findings to evaluate the impact of graphene on antenna performance. Additionally, the potential applications of graphene in various fields will be addressed. Innovative solutions enabled by graphene in industries such as electronics, energy storage, biomedical applications, sensor technologies, and aerospace will be examined. Furthermore, similar studies and other graphene-based antenna designs in the literature will be thoroughly reviewed, with a focus on highlighting the unique contributions of the proposed design.

Keywords: Graphene-based antenna, 5G microstrip patch antenna, graphene applications in rf and microwave, next-generation antenna materials.

Introduction

With the rapid evolution of wireless communication technologies, particularly the deployment of fifth generation (5G) networks, there is an increasing demand for compact, high-performance, and low-loss antennas capable of supporting high data rates, minimal latency, and enhanced network capacity. These demands have driven the exploration of novel materials and innovative antenna design techniques to overcome the limitations of conventional microstrip patch antennas, such as limited bandwidth, relatively high conductive losses, and reduced efficiency at higher frequencies [1]. Among the various materials investigated, graphene has emerged as a highly promising candidate for next-generation antenna applications due to its outstanding electrical, thermal, and mechanical properties [2]. As a two-dimensional carbon-based nanomaterial, graphene exhibits tuneable electrical conductivity, high carrier mobility, and significant electromagnetic interaction capabilities. These features make it particularly attractive for reconfigurable and miniaturized antenna designs, especially in the context of high-frequency applications such as millimeter-wave and terahertz (THz) bands [3-5].

In recent studies, graphene has not only been utilized as a radiating element but also explored as a functional coating material that can enhance antenna performance by reducing surface resistance and improving current distribution across the patch [6-8]. Its integration onto conventional substrates through various deposition techniques, such as Electrophoretic Deposition (EPD), has demonstrated promising results in terms of gain enhancement, bandwidth improvement, and reduction in reflection loss [9]. This study focuses on the design, fabrication, and characterization of a graphene-coated microstrip patch antenna operating at 4.5 GHz for 5G applications. A fractal geometry has been employed to further optimize the antenna's performance in terms of bandwidth and efficiency. The application of reduced graphene oxide (rGO) via the EPD method is investigated in detail, assessing its impact on key antenna parameters such as return loss, radiation pattern, and gain. Comparative analyses between conventional and rGO-coated antenna prototypes are presented to highlight the functional advantages of graphene integration [10,11].

Furthermore, the broader implications of this work are discussed in relation to future wireless systems, where graphene-based materials are expected to play a transformative role not only in antennas but also in various optoelectronic devices, sensors, and THz components [12]. The outcomes of this research aim to contribute to the advancement of low-loss, high-efficiency antenna technologies, supporting the continuous development of 5G and beyond communication infrastructures.

Methods

Synthesis of Graphene

The reduced graphene oxide (rGO) used in this thesis study was synthesized from commercial graphite via the Modified Hummers method. The obtained rGO was subsequently analyzed by electrophoretic deposition (EPD) onto microstrip patch antennas. The synthesis process of graphene involved the following stages:

Preparation of Graphene Oxide (GO)

Graphene oxide (GO) was synthesized from graphite powder using the modified Hummers method. In brief, 1 g of graphite and 0.5 g of sodium nitrate ($NaNO_3$) were mixed, followed by the slow addition of 23 ml of concentrated sulfuric acid (H_2SO_4) under continuous stirring. After stirring for 1 hour, 3 g of potassium permanganate ($KMnO_4$) was gradually added while maintaining the temperature below 20 °C to prevent overheating and potential hazards. The reaction mixture was then stirred at 35 °C for 12 hours. Afterward, 500 ml of distilled water was added under vigorous stirring to dilute the solution. To ensure the complete reaction of $KMnO_4$, 5 ml of 30% hydrogen peroxide (H_2O_2) solution was added. The resulting mixture was sequentially washed with hydrochloric acid (HCl) and deionized water, then filtered and dried to obtain graphene oxide [13].

Reduction of Graphene Oxide to rGO

Dried graphene oxide (400 mg), prepared by the modified Hummers method, was dispersed in 400 ml of distilled water under vigorous stirring and ultrasonication for at least 1 hour. The reduction process was carried out using hydrazine hydrate at a weight ratio of 1:1 (hydrazine hydrate:GO), at various temperatures including 15 °C, 60 °C, 80 °C, and 95 °C. Continuous stirring and intermittent ultrasonication were applied throughout the entire reduction process. Upon completion, the rGO was isolated by filtration, thoroughly washed with distilled water (four times), and subsequently dried at 60 °C for 24 hours to remove any residual solvents [14].

Electrophoretic Deposition of rGO onto FR-4 Substrates

Electrophoretic deposition (EPD) was performed at room temperature using a custom-fabricated electrochemical glass cell. The setup comprised two electrodes: the working electrode (WE), consisting of FR-4 plates, and the counter electrode (CE), made of a platinum (Pt) wire. The electrodes were connected to a GwINSTEK PSW 08-13.5 programmable DC power supply, with a fixed electrode spacing of 1 cm. The EPD suspension was prepared using ultrasonicated reduced graphene oxide (rGO) and copper(II) acetate [$Cu(OAc)_2$]. After optimization of the deposition parameters, rGO and copper particles were co-deposited onto the FR-4 substrates under specific voltage and time conditions. The most optimal results were achieved under the conditions of 5 V applied voltage and a deposition duration of 150 minutes. An image of the coated FR-4 substrate is presented in Figure 1. Post-deposition, the samples were dried in a laboratory oven at 70 °C for 30 minutes to facilitate solvent evaporation and improve film adhesion.

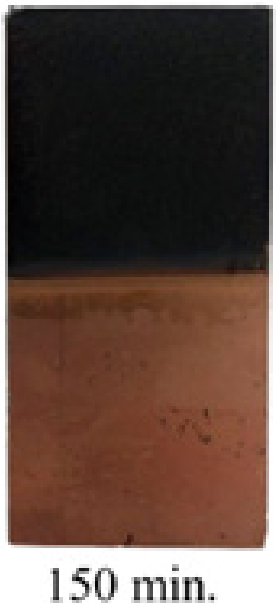


Figure 1. EPD process results in 5V-150 min.

Following EPD, selected samples underwent characterization via Raman spectroscopy and Scanning Electron Microscopy (SEM). The thickness of the deposited films was measured using both optical interferometry and SEM cross-sectional analysis to ensure accuracy and consistency. Raman spectra confirmed the successful deposition of rGO on samples 1 and 3, while other samples exhibited irregular or degraded FR-4 surfaces. The presence and structural quality of rGO were further evaluated through Raman spectroscopy, focusing on the characteristic D, G, and 2D peaks typical of carbon-based materials. These spectral features provide insights into film thickness, uniformity, and structural integrity. For sample 1, prominent Raman peaks were observed at 1329 cm^{-1} (D-band) and 1596 cm^{-1} (G-band), while sample 2 exhibited peaks at 1331 cm^{-1} and 1592 cm^{-1}, respectively. The increased intensity of the D-band is attributed to the partial reduction of graphene oxide, resulting in a higher density of defects. The intensity ratio of the D and G bands (I_D/I_G), which serves as an indicator of defect concentration and hybridization state, was calculated to be 1.13 for sample 1 and 1.18 for sample 2. SEM imaging of samples 1 and 3 revealed a relatively uniform rGO layer deposited between copper crystal grains, supporting the Raman findings. Film thickness measurements showed that sample 1 had a thickness of 1.5–2 μm based on optical interferometry, while sample 3 ranged between 3–4 μm. SEM cross-sectional analysis, on the other hand, indicated a thickness of 3.74–4.88 μm for sample 1 and 1.20–1.45 μm for sample 3. These discrepancies are attributed to surface roughness and variation across different sample regions.

Design and Analysis of Microstrip Patch Antennas Using CST Software and Realization

First and foremost, the conductivity of the material to be used graphene must be determined. Although the conductivity of graphene is inherently governed by its unique band structure, it also depends on several interrelated parameters, including chemical doping,

Fermi energy (chemical potential), electron mobility, relaxation time, operating frequency, and ambient temperature. It is emphasized that these parameters are not mutually independent and can influence one another.

Under a semi-classical, local model and in the absence of an external magnetic field, graphene's conductivity can be modelled using the intraband contribution of the Kubo formula [1] as follows:

$$\sigma = -j\frac{e^2 k_B T}{\pi\hbar^2(\omega - j\tau^{-1})}\ln\{2[1+\cos h\,(\frac{\mu_c}{k_B T})]\} \quad (1)$$

where:

1. Elementary charge is *e*,
2. Boltzmann constant is k_B,
3. Absolute temperature is *T*,
4. Reduced Planck constant is ħ,
5. Angular frequency is ω,
6. Relaxation time is τ,
7. Chemical potential (Fermi energy) is μ_c.

To ensure accurate modelling of the antenna, the dimensions of the microstrip patch must be calculated precisely. The design consists of a square patch positioned on a dielectric substrate with a relative permittivity ε_r , placed above a perfect electric conductor (PEC) ground plane. The patch, assumed to be made of copper, is excited via a copper microstrip feed line.

As shown in Figure 2, the structure represents a typical microstrip patch antenna configuration. For microstrip antennas, the width-to-height ratio should satisfy W/h > 1. When W/h > 1, surface wave excitation is reduced, and radiation efficiency increases. Narrow patches (i.e., W/h ≤ 1) confine more energy within the substrate as surface waves, which do not contribute to radiation and may cause interference or power loss [1]. The effective dielectric constant, ε_{reff} which accounts for the fringing fields, is calculated using the following empirical formula:

$$\varepsilon_{reff} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r - 1}{2}\left[1 + 12\frac{h}{w}\right]^{-1/2} \quad (2)$$

The width *W* of the patch is then determined using:

$$W = \frac{1}{2f_r\sqrt{\mu_0\varepsilon_0}}\sqrt{\frac{2}{\varepsilon_r + 1}} = \frac{v_0}{2fr}\sqrt{\frac{2}{\varepsilon_r + 1}} \quad (3)$$

Once *W* is calculated, the length extension ΔL, caused by fringing effects, is computed as:

$$\frac{\Delta L}{h} = 0.412 \frac{(\varepsilon_{reff} + 0.3)\left(\frac{W}{h} + 0.264\right)}{(\varepsilon_{reff} - 0.258)\left(\frac{W}{h} + 0.8\right)} \quad (4)$$

Finally, the effective length L of the patch is calculated using:

$$L = \frac{1}{2f_r\sqrt{\varepsilon_{reff}}\sqrt{\mu_0\varepsilon_0}} - (2\Delta L) \quad (5)$$

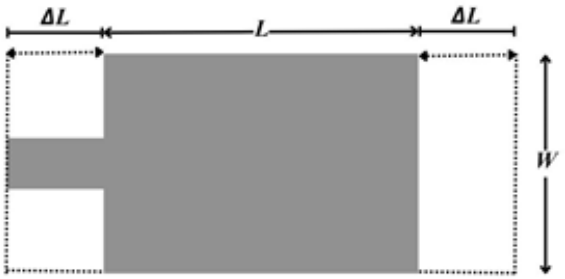


Figure 2. Physical lengths of microstrip patch.

Using the derived equations, the fundamental geometry of the antenna was modelled in CST (Computer Simulation Technology) software (Figure 3). Since the objective was to design a fractal structure, modifications were applied to align the geometry with fractal characteristics. These modifications were implemented through the integration of hexagonal (honeycomb) patterns. The notches were introduced using the conventional 1/3 scaling ratio frequently employed in fractal antenna designs [15]. Similarly, the honeycomb configuration was constructed using comparable scaling principles. The desired resonance frequency is denoted by f_r, and c represents the speed of light. The antenna's bandwidth W_p is dependent on the relative permittivity ε_r of the substrate material. To account for the influence of fringing fields that extend beyond the physical boundaries of the radiating patch, the effective dielectric constant ε_{reff} must be calculated and incorporated into the analysis.

Figure 3 illustrates the dimensions of the designed antenna. Figure 3(a) shows the front view of the design, while Figure 3(b) displays the back view. The dimensions of the front design have been optimized in accordance with fractal structures. For the antenna's rear feeding, a feeding model commonly used in UWB antennas has been employed, based on the width of the antenna. To design the UWB antenna, three techniques were applied: first, the use of recessed structures; second, the implementation of a partial ground plane; and third, the inclusion of a single slot in the patch to achieve good impedance matching. By selecting these parameters, the proposed antenna can be tuned to operate within the 2–12 GHz frequency range [16].

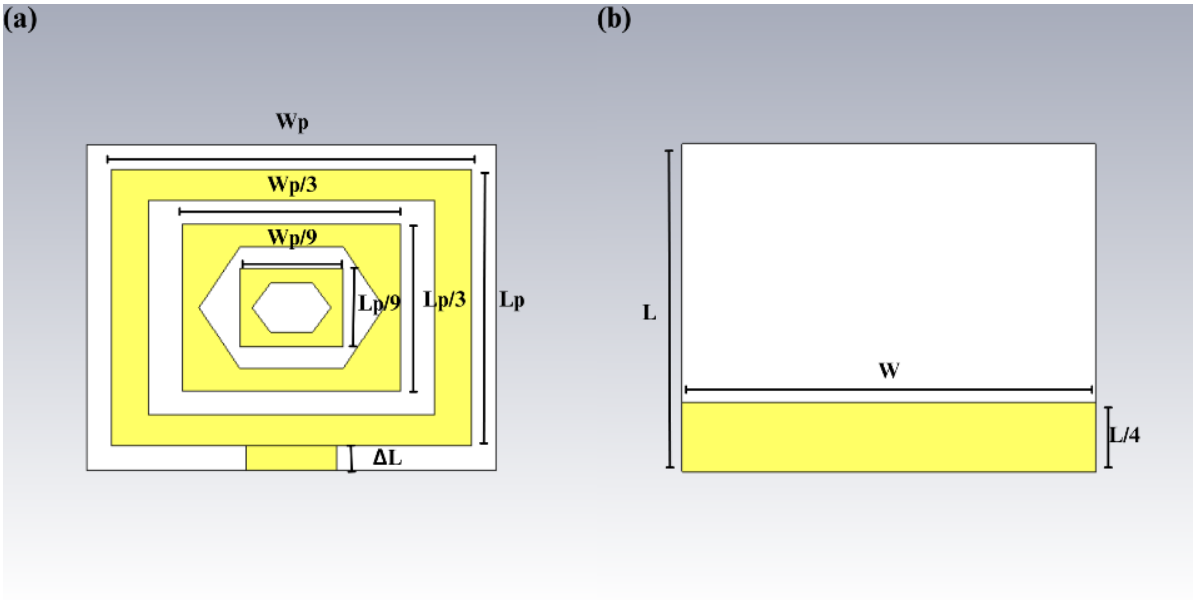


Figure 3. The CST-based design of the proposed antenna is showing (a) the front view and (b) the back view of the structure.

Figure 4 shows the structure of the designed antenna, which has been fabricated using a PCB etching process. Figure 4(a) depicts the copper plate before graphene coating, while Figure 4(b) shows the etched copper plate after graphene coating. In future studies, the real-world radiation patterns of these antennas will be obtained.

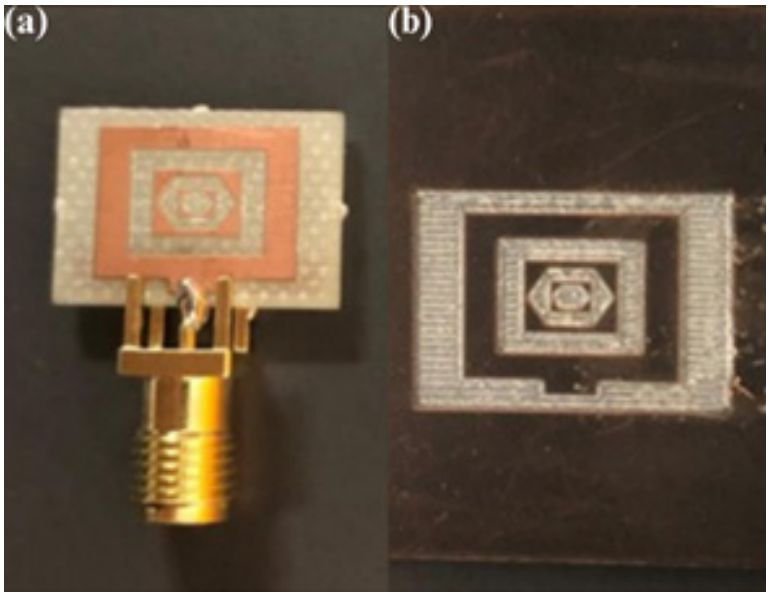


Figure 4. The fabricated prototype of the designed antenna (a) only a copper plate, (b) graphene-coated copper plate.

Antenna Design Using CST Software

In the initial design phase, a rectangular microstrip patch antenna geometry based on an FR-4 substrate was created. The relative permittivity (ε_r) of the FR-4 substrate was taken as 4.3, and the loss tangent (tan δ) was assumed to be 0.025. The substrate thickness was set to 1.6 mm, and the copper patch thickness was defined as 0.035 mm. The

antenna was designed to operate at a centre frequency of 4.5 GHz. Parameters such as patch length, width, ground plane dimensions, and feed point were analytically calculated and finalized using CST's parametric optimization tools. In antenna design, the S11 parameter (also known as return loss) is expected to be less than -10 dB. This indicates that at least 90% of the input power is effectively radiated or transmitted, while less than 10% is reflected back, ensuring efficient impedance matching and antenna performance [1]. It is observed that the designed antenna exhibits a return loss of -35 dB at 4.5 GHz, indicating excellent impedance matching and minimal reflection at the target frequency (Figure 5).

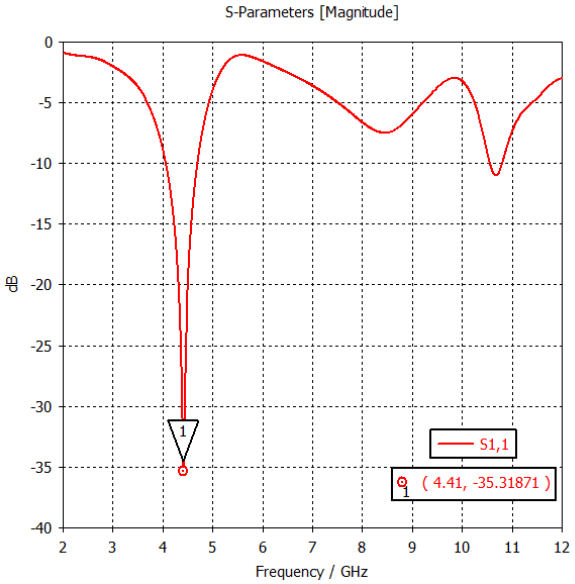


Figure 5. S-Parameters for designed antenna in 4.5GHz.

Voltage Standing Wave Ratio (VSWR) is a key parameter used to evaluate the impedance matching between an antenna and the transmission line or source. It indicates how efficiently radio-frequency power is transmitted from the source to the load (i.e., the antenna) without being reflected back.

Mathematically, VSWR is defined as:

$$VSWR = \frac{1+|\Gamma|}{1-|\Gamma|} \quad (6)$$

where Γ (Gamma) is the reflection coefficient, which quantifies the portion of the signal reflected back due to impedance mismatch.

- VSWR = 1 means perfect matching – all the power is transmitted, and none is reflected.
- VSWR < 2 is generally considered acceptable in most antenna designs.
- VSWR > 2 indicates poor matching, leading to higher reflected power and reduced efficiency [2].

It can be observed that in Figure 6, the VSWR value is 1.058382, indicating excellent impedance matching between the antenna and the transmission line, with minimal reflected power.

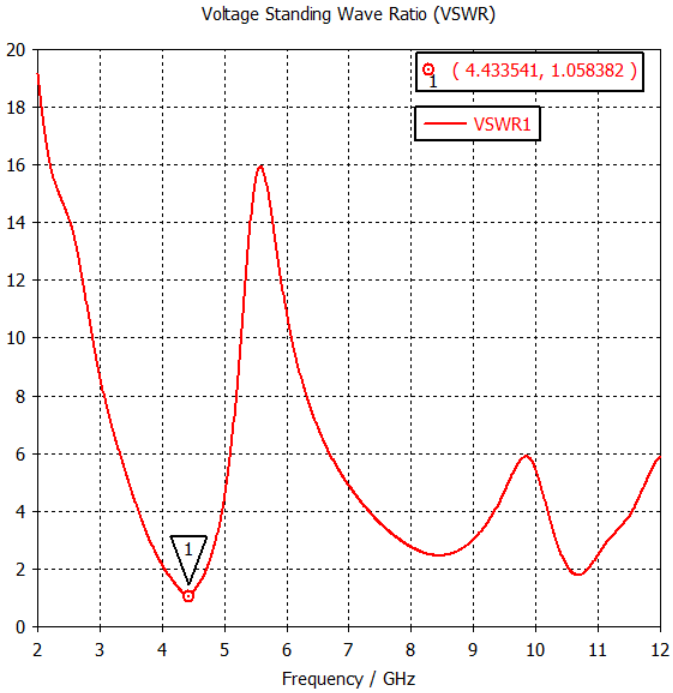


Figure 6. VSWR for designed antenna in 4.5GHz.

The polar gain plot is a graph that illustrates the angular variation of the power (electromagnetic field strength) radiated by the antenna at a fixed distance in its far-field region. At 4.5 GHz, the main lobe's amplitude (maximum gain) for phi = 0° is 3.16 dB (Figure 7).

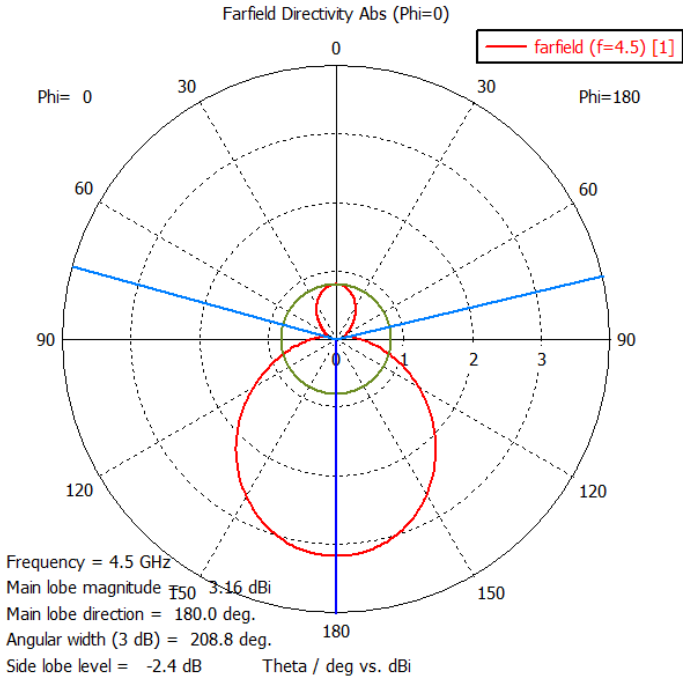


Figure 7. 3D polar gain pattern for designed antenna in 4.5GHz.

Conclusion

In this study, a graphene-coated, fractal-structured 5G microstrip patch antenna was designed, simulated, and partially fabricated. In the previous study given in [17], it was emphasized that the performed simulations had room for improvement. In this context, the antenna design was revised, and new simulations were conducted based on the updated structure. As a result of these modifications, significantly improved outcomes were obtained in terms of return loss (S11), voltage standing wave ratio (VSWR), and gain, compared to the earlier results. These enhancements in the design clearly demonstrate the critical importance of antenna parameters and the pattern configuration on overall performance. Building upon these findings, the fabrication of the antenna prototype has been initiated, and experimental measurements are currently underway. The antenna structure was modelled using CST Microwave Studio, and design optimizations were carried out to ensure compatibility with UWB (Ultra-Wideband) applications. The simulated results demonstrated that the designed antenna achieves a return loss (S11) of -35 dB at 4.5 GHz, indicating excellent impedance matching. Furthermore, a VSWR value of 1.058 confirms effective power transmission with minimal reflection. The polar gain plot at 4.5 GHz (φ = 0°) shows a maximum gain of 3.16 dB, indicating strong directional radiation performance. The physical fabrication of the antenna was completed using PCB etching, and two prototypes were produced: one with a plain copper patch and the other with a reduced graphene oxide (rGO) coating applied via electrophoretic deposition (EPD). Raman spectroscopy and SEM imaging validated the successful rGO deposition. The observed ID/IG ratios confirmed structural changes consistent with GO reduction, contributing to improved electrical conductivity [17].

In future work, physical performance tests will be conducted using a Vector Network Analyzer (VNA) to measure the antenna's real-world S-parameters. These experimental results will be compared with the simulation data to evaluate the accuracy and effectiveness of the design, as well as the influence of the graphene coating on antenna performance. The findings of this study demonstrate the promising potential of graphene as a conductive coating in high-frequency antenna applications and highlight the advantages of combining fractal geometries with advanced nanomaterials for next-generation wireless technologies.

Declaration of Interest: The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgements: This study was supported by the Scientific Research Projects Coordination Unit of Sivas University of Science and Technology under Project Number: 2023-GENL-Hav-0013. The authors gratefully acknowledge this support.

Author Contributions: All authors contributed equally to this work.

REFERENCES

[1] C. A. Balanis, *Antenna Theory: Analysis and Design*, 3rd ed. Hoboken, NJ, USA: John Wiley & Sons, 2005.

[2] J. M. Jornet and I. F. Akyildiz, "Graphene-based plasmonic nano-antenna for terahertz band communication in nanonetworks," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 12, Suppl. Part 2, pp. 685–694, Dec. 2013.

[3] A. Azizi *et al.*, "Graphene-based microstrip patch antenna with enhanced bandwidth," *Prog. Electromagn. Res. Lett.*, vol. 71, pp. 123–130, 2017.

[4] G. Li, Z. Liu, W. Gao, *et al.*, "Recent advances in graphene-based materials for electromagnetic wave absorption," *Coord. Chem. Rev.*, vol. 478, p. 214966, 2023.

[5] W. L. Stutzmann and G. A. Thiele, *Antenna Theory and Design*, 2nd ed. Hoboken, NJ, USA: John Wiley & Sons, 2012.

[6] F. M. de Sousa, A. M. Silva, D. H. Spadoti, *et al.*, "Graphene-based nano patch antenna using photonic band gap insertion into substrate for applications at THz band," *Opt. Quant. Electron.*, vol. 55, no. 210, 2023.

[7] C. Song, J. Wang, B. Zhang, *et al.*, "Dual band/ultra-broadband switchable terahertz metamaterial absorber based on vanadium dioxide and graphene," *Opt. Commun.*, vol. 530, p. 129027, 2023.

[8] A. A. Mohammad Khani, M. A. Shakeri, S. M. R. Nezhad, *et al.*, "multi-layers THz duplexer based on combined graphene patterns," *Solid State Commun.*, vol. 360, p. 114974, 2023.

[9] Muhammad Atiq Ur Rehman, Qianq Chen, *et al.*, "Electrophoretic deposition of carbon nanotubes: recent progress and remaining challenges", *International Materials Reviews*, Volume 66, Issue 8, November 2021, Pages 533-562.

[10] A. Lamminen, K. Arapov, G. de With, S. Haque, H. G. O. Sandberg, H. Friedrich, and V. Ermolov, "Graphene-flakes printed wideband elliptical dipole antenna for low-cost wireless communications applications," *IEEE Antennas Wirel. Propag. Lett.*, vol. 16, pp. 1883–1886, 2017.

[11] A. Karmakar, "Fractal antennas and arrays: A review and recent developments," *Int. J. Microw. Wireless Technol.*, pp. 1–25, 2020.

[12] Q. Zhou, Y. Chen, Z. Fang, *et al.*, "Graphene-based terahertz optoelectronics," *Opt. Laser Technol.*, vol. 157, p. 108558, 2023.

[13] H. Liu, Y. Liu, and Z. Zhu, "Chemical doping of graphene," *J. Mater. Chem.*, vol. 21, no. 10, pp. 3335–3345, 2011.

[14] M. J. Allen, V. C. Tung, and R. B. Kaner, "Honeycomb carbon: A review of graphene," *Chem. Rev.*, vol. 110, no. 1, pp. 132–145, 2010.

[15] K. Niu, P. Li, Z. Huang, L. J. Jiang, and H. Bagci, "Numerical methods for electromagnetic modeling of graphene: A review," *IEEE J. Multiscale Multiphys. Comput. Techn.*, vol. 5, pp. 143–158, 2020.

[16] S. H. Choi, J. K. Park, S. K. Kim, and J. Y. Park, "A new ultra-wideband antenna for UWB applications," *Microw. Opt. Technol. Lett.*, vol. 40, no. 5, pp. 378–382, 2004.

[17] N. Demirkoparan and H. D. Basdemir, "Graphene-based fractal structure 5G microstrip patch antenna design," in *Proc. 15th National Conf. Electrical and Electronics Engineering (ELECO)*, Bursa, Turkey, 2024

PERFORMANCE ANALYSIS OF THE CHEETAH OPTIMIZATION ALGORITHM IN PREDICTIVE MAINTENANCE FORECASTING

Nuriye Bektaşoğlu*
Zonguldak Bülent Ecevit University, Department of Electrical and Electronics Engineering, Zonguldak, Türkiye

Ali Narin
Zonguldak Bülent Ecevit University, Department of Electrical and Electronics Engineering, Zonguldak, Türkiye

Uğur İleri
Düzce University, Department of Electrical-Electronics and Computer Engineering, Institute of Science, Düzce, Türkiye
(Corresponding author) nuriyebektasoglu1@gmail.com

ABSTRACT

Traditional maintenance approaches are based on planned interventions at specific time intervals or corrective actions after a failure occurs. In contrast, predictive maintenance approaches utilize sensor data and machine learning algorithms to predict potential failures in advance and thus ensure system continuity. In this study, a machine learning-based PdM model for fault prediction in industrial systems is developed using AI4I 2020 dataset. In the classification process, the performances of Decision Trees, Support Vector Machines and k-Nearest Neighbor algorithms are comparatively analyzed. The SMOTE algorithm was applied to remove the class imbalance in the dataset; the effects of this process on the classification performance were analyzed. Compared to the original dataset, the dataset balanced with the SMOTE method showed higher classification performance. In the normalization phase, z-score and min-max methods were evaluated and the z-score method showed higher accuracy and consistency. Cheetah Optimization Algorithm was used in the feature selection process, but it was observed that this method negatively affected the model performance in some data structures. However, applying the same algorithm to alternative data structures where failure modes are categorized into binary classes resulted in a significant increase in accuracy. The results clearly reveal the impact of data preprocessing and feature selection on the performance of PdM models.

Keywords: Predictive maintenance, fault classification, data balancing, cheetah optimization algorithm.

Introduction

With the impact of the Fourth Industrial Revolution, digitalization has accelerated in production processes, and remote monitoring and uninterrupted operation of machines has become an important factor for system sustainability [1]. With this transformation, concepts such as sensor technologies, Internet of Things (IoT), cyber-physical systems and predictive maintenance (PdM) have been integrated into industrial applications, resulting in significant increases in operational efficiency and production effectiveness. In particular, Industrial Internet of Things (IIoT) and machine learning-based systems offer significant contributions in terms of monitoring the condition of equipment, predicting potential failures and improving production processes. Traditional maintenance approaches used in industrial applications generally aim to intervene after the failure occurs. These approaches can cause unexpected stoppages and high maintenance costs in production lines. On the other hand, maintenance activities performed at fixed intervals may not be necessary in every case, and this may increase costs by causing replacement of functional parts. In contrast to these traditional approaches, predictive maintenance strategies aim to determine maintenance requirements at the most appropriate time and accurately based on the analysis of sensor data obtained from machines and equipment in real time [2]. This method not only improves maintenance processes, but also contributes to more efficient production planning. For example, by continuously monitoring the current status of equipment, maintenance operations can be carried out with optimal timing, thus minimizing unexpected production interruptions.

As a result, predictive maintenance applications provide a more stable process flow in the production line, while also offering the opportunity for improvement in supply chain management.

In addition, components with a high risk of failure can be identified in advance and spare parts management can be performed more effectively. Predictive maintenance systems integrated with dynamic production scheduling can maximize productivity by providing flexibility in production processes.

In today's competitive industrial environment, predictive maintenance not only reduces operational costs, but also provides enterprises with a competitive advantage in the global market by making production processes smarter [3]. In recent years, various machine learning methods have been developed on the processing and classification of data obtained from sensors in order to increase the effectiveness of predictive maintenance applications. When the literature is examined, there are many studies in which machine learning methods based on Bayesian optimization are used for unstable data sets. It has been emphasized that these methods increase the classification success by optimizing the model hyperparameters [4]. In another study, frequency domain vibration signals were used in the detection of induction motor failures and success was achieved with a Support Vector Machine (SVM) algorithm over these signals [5]. In a study conducted on bearing failures, using wavelet transform (WT), a time-frequency domain analysis tool, accurate classification of wear levels was achieved with the help of machine learning algorithms with features extracted from vibration signals [6]. In another study published in TECHRXIV,

an analysis was conducted that examined the fault trends and the effects of these failures on operational costs, and it was found that heat dissipation failures (HDF) account for 74% of maintenance costs. This finding reveals the necessity of innovative solutions (for example, advanced cooling systems, load optimization and power stabilizers) in order to reduce failures [7]. In the study of Sengupta et al., predictive maintenance strategies for armored vehicles were considered and evaluated by various machine learning algorithm analyses performed on sensor data, k-fold cross-validation and TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) [8]. In another study, İleri et al. examined the performance of both approaches by comparing deep learning models (1D-LeNet, 1D-AlexNet, 1D-VGG16) and traditional machine learning algorithms (DT, SVM, k-NN) for automatic fault classification [9]. These researches highlight the advances made by machine learning and deep learning based methods in the field of PdM and their performance on the AI4I 2020-PdM dataset. The studies mentioned above, data imbalance in predictive maintenance applications, limited scalability and dependence on shallow machine learning methods limit the performance of error classification models.

In this study, solutions were developed for the data imbalance problem in order to improve predictive maintenance processes, the effects of different normalization techniques were examined, and the cheetah optimization algorithm was applied to determine the most appropriate feature set. In the first stage, using the AI4I 2020 predictive maintenance dataset, the SMOTE (Synthetic Minority Over-sampling Technique) method, which produces synthetic examples for minority classes in order to eliminate the class imbalance in the dataset, was applied and a balanced data set was created. The z-score normalization process was applied to the obtained balanced data set and then analysis was performed with classification algorithms such as Decision Trees (DT), Support Vector Machines (SVM) and k-Nearest Neighbors (k-NN). In order to increase model performance and eliminate unnecessary features, feature selection was performed with the Cheetah Optimization Algorithm. At this stage, the classification models were evaluated on various performance metrics such as accuracy, precision, sensitivity and F1 score. In the second phase of the study, in order to examine the failure modes in more detail, each failure mode was separated into binary classes and a new data set was created. Z-score normalization was applied to this data set; then, the analysis process was carried out with the same classification algorithms (DT, SVM and k-NN). In order to eliminate unnecessary features and increase model accuracy, the Cheetah Optimization Algorithm was applied to this data set and the features were gradually removed. As a result, the performance of the models created for both data sets was compared through the determined performance metrics and the most effective method for classifying the failure modes was revealed. At each stage, the data processing method with the highest performance rate and the machine learning model were determined.

In the rest of the paper, the data used, the preparation of the data, the classification algorithms, the method of feature selection and the performance metrics will be mentioned, and finally the results obtained will be given and discussed.

Method

Data Set

The AI4I 2020 predictive maintenance dataset is a publicly available dataset provided by the UCI Machine Learning Repository [10]. This data set includes sensor data collected from the industrial production line of a simulated production facility. The purpose of the data set is to contribute to research and development in the fields of predictive maintenance and machine learning. The data set contains five different types of failures given in Table 1. The dataset used in the study consists of a total of 10,000 rows and contains six basic features. These features can be defined as follows:

The first feature contains the product identity or product type information that defines each production unit.

The second feature is the air temperature and this parameter is generated by the random walk method.

The third feature expresses the process temperature in the production process.

The fourth feature is the rotation speed and this value is derived from a reference corresponding to 2900 Watt power in the system.

The fifth feature is the torque information, which is usually observed at approximately 40 Nm.

The sixth and last feature, tool wear, varies depending on the quality variation of the product.

The synthetic dataset in question offers a suitable sample for predictive maintenance applications and training of machine learning-based models with its structure that mimics maintenance scenarios that can be encountered in real production environments.

Table 1: Failure Modes And Their Descriptions

Failure Mode	Description
Tool Wear Failure (TWF)	Malfunctions caused by the wear of tools
Heat Dissipation Failure (HDF)	Failures in heat dissipation
Power Failure (PWF)	Power loss or electrical malfunctions
Overstrain Failure (OSF)	Malfunctions caused by overloading
Random Failures (RNF)	Unexpected and randomly occurring failures

Normalization

In this study, data scaling methods were used to reduce the negative effects that variables with different measurement units may have on model performance. Such preprocessing steps aim to increase both the model's more efficient learning and the stability of the results by ensuring that all features are represented in similar ranges. The two basic normalization approaches frequently preferred in the literature are min-max and z-score

based methods. The mathematical expression of the min-max method, which performs the rescaling process for each value within a certain lower and upper limit range, is given in equation (1).

$$y = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (1)$$

In this equation, y = shows the normalized data, x_i = the input value, x_{min} = the smallest number in the input set, x_{max} = the largest number in the input set.

The mathematical representations of the Z-score normalization process are as follows:

$$\mu = \frac{1}{N}\sum_{i=1}^{N}(x_i) \quad (2)$$

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu^2} \quad (3)$$

$$z = \frac{x - \mu}{\sigma} \quad (4)$$

Here z denotes the normalized value, and μ and σ denote the mean and standard deviation, respectively.

Data Balancing

A significantly unbalanced distribution of observation numbers between classes in data sets can have negative effects on the performance of classification algorithms. In particular, the inadequacy of examples belonging to the minority class can make it difficult for the model to learn the characteristics specific to this class, which can lead to incorrect classifications. One of the approaches developed for this problem is the SMOTE method.

SMOTE generates synthetic examples based on the distance between the examples based on the close neighbors of the observations belonging to the minority class [11]. Thus, the representative power of the minority class is increased, while the imbalance between classes is reduced.

In this study, the limited number of machine failure (MF) cases in the data set causes class imbalance. Therefore, strategies to eliminate this imbalance in the analysis process are primarily considered.

In the original dataset, the minority class fault case contains only 339 samples, while the majority class consists of 9,661 samples. This imbalance prevents machine learning models from learning the distinction between classes correctly, negatively affecting the prediction performance. Therefore, the minority class samples were synthetically multiplied by applying the SMOTE method to the dataset, and the number of samples of both classes was approximately equalized (9,492 samples). With the balanced dataset obtained as a result of this process, it was aimed for the model to learn both classes effectively and produce more balanced classification results. The change in the class distribution before and after the SMOTE application is clearly shown in Figure 1.

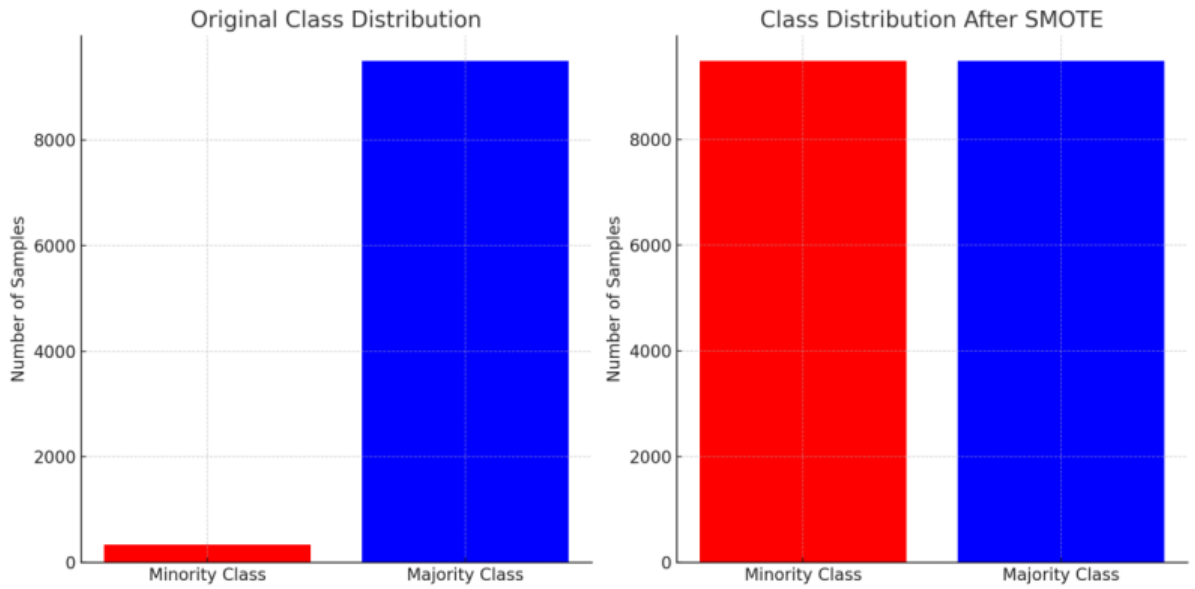


Figure 1: Class Distribution Before And After SMOTE Method

Classification Algorithms

In classification-based machine learning approaches, models learn from given labeled data, create structures representing each class, and try to predict the data into the correct class. This process works to understand the features of each category, providing an effective solution for classification problems. In particular, accurate prediction of machine failure and error types plays an important role in the maintenance and repair processes of systems. The basic features and working principles of widely used machine learning algorithms such as DT, SVM, and k-NN will be discussed in the following sections.

Decision Trees

Decision trees (DTs) are supervised learning algorithms frequently used in both classification and regression problems and aim to create more homogeneous structures in each subgroup by dividing the data into branches according to certain features [12]. The structure of the model is constructed in a way that divides the data set into subgroups by conditional splits performed on a feature at each node starting from the root node. The leaf nodes formed as a result of these splitting operations correspond to either a specific class label or a numerical predicted value. Decision trees provide an advantage especially in terms of interpretability; since each decision path clearly shows the conditions according to which the data is classified. Thanks to this feature, it allows the model's decision processes to be monitored transparently. The ability to make fast and intuitive decisions based on data makes decision trees especially effective in the analysis of complex data structures.

Support Vector Machines

Support Vector Machines (SVM) is one of the supervised learning algorithms and is generally used for classification problems. The main purpose of this algorithm is to find the hyperplane that best separates the data into two classes [13]. This hyperplane is determined in a way that will maximize the remaining distance (margin) between the classes.

In other words, it aims to obtain the most appropriate separation plane by balancing the minimum distance to both classes. In addition, SVM is not limited to linearly separable data; it can also define effective decision boundaries in nonlinear problems by transforming the data space to higher dimensions through kernel functions. The most basic SVM mathematical expression can be explained by equation (5).

$$f(x) = w^{T}x + b \tag{5}$$

In this equation, x = input feature vector of the classifier, w = weight vector of the hyperplane, T = transpose and finally b = hyperplane location.

k-Nearest Neighbor

The k-nearest neighbor (k-NN) algorithm determines the class of a data point by looking at the classes of its k nearest neighbors. This method is widely used especially in areas such as pattern recognition and text classification [14].

k-NN provides advantages to users due to its simplicity, robustness against noisy data and generally high accuracy rate. However, k-NN can be computationally costly in large data sets, because distance calculation is performed with all training data for each new data point. In addition, it is important to choose the k value correctly; small k values lead to overfitting, while large k values can provide more generalized results. Cover and Hart proposed the k-NN algorithm, which takes into account a certain number of nearest neighbors (k) to determine the class of a data sample and determines the class of the test sample by selecting the most frequent class among these neighbors [21]. Later, Bailey and Jain developed the algorithm with a method in which neighbors are weighted according to distance.

Assuming that X 1 and X 2 are represented by feature vectors:

$$X_1 = \{ x_{11}, x_{12}, ..., x_{1m} \} \tag{6}$$

$$X_2 = \{ x_{21}, x_{22}, ..., x_{2m} \} \tag{7}$$

In this equation, m represents the dimension of the feature space. Neighborhood is measured by distance. The most commonly used distance measure is the Euclidean distance expressed in equation (8):

$$\text{dist}(X_1, X_2) = \sqrt{\sum_{i=1}^{m} (x_{1i} - x_{2i})^2} \tag{8}$$

Here x_{1i} and x_{2i} are the i-th features of data points X_1 and X_2, respectively.

Cheetah Optimization Algorithm

The cheetah optimization algorithm (COA) is a meta-heuristic search algorithm inspired by the hunting strategies of cheetahs [15]. Cheetahs are the fastest land animals in the world, living mostly in Africa and Iran and can reach speeds of up to 120 km/h. Their evolutionary adaptations have allowed cheetahs to have this incredible speed and flexibility. Their long tails help them maintain their balance while performing fast maneuvers, while

their thin legs allow them to run strong and fast. Their light and flexible spines increase the flexibility of their bodies during fast movements. Cheetahs approach their prey by combining their stealth and speed and catch them in a short time. However, since they cannot maintain their high speed for a long time, the hunting process usually lasts no more than half a minute. Therefore, cheetahs' hunting strategies require fast and effective attacks. Cheetahs usually hunt gazelles, antelopes, rabbits and pigs. They slowly approach their prey and try to reach the shortest distance, then hide and wait for the prey to approach. When they get close enough, they quickly attack. Cheetahs disrupt their prey's balance by striking them quickly and powerfully with their front paws and causing them to fall to the ground. They hold their prey tightly with their powerful claws, preventing them from escaping. This strategy allows cheetahs to hunt successfully and catch their prey safely. Cheetah hunting strategies are divided into three parts. The first part is searching for the prey, the second part is watching the prey, and the last stage is attacking the prey at the right time. COA provides a structure that can be applied to optimization problems by imitating these biological processes.

Searching for Prey

Cheetahs actively search in their own territory or in the surrounding areas to find their prey. The mathematical expression of cheetahs' search for their prey is given in equation no. 9.

$$X_{i,j}^{t+1} = X_{i,j}^{t} + \hat{r}_{i,j}^{-1} \cdot \alpha_{i,j}^{t} \quad (9)$$

Here:

- $X_{i,j}^{t+1}$ and $X_{i,j}^{t}$ represent the next and current locations of the i. cheetah (i = 1, 2, ..., n) in the j. plane (j = 1, 2, ..., D), respectively. Here, 'n' represents the cheetah population number and 'D' represents the number of variables (size) in the optimization problem. t: index represents the current hunting time.

- The $\hat{r}_{i,j}^{-1}$ parameter represents the randomization parameter. This parameter consists of random numbers obtained from the standard normal distribution and varying between 0 and 1. It plays an important role in providing the balance of exploration and exploitation in the search process by determining the direction and speed of movement of the cheetah.

- $\alpha_{i,j}^{t}$ represents the amount of change in the cheetah's current location (movement distance). This represents the step length of the i. cheetah in the j dimension at time t. The step length can be chosen as t/(1000.T) with $\alpha_{i,j}^{t} > 0$. Here T represents the maximum length of the hunting period. It should be noted that the cheetah's step lengths are completely random and this model takes this into account. The search strategy is shown in Figure 2a [16].

It should be noted that the cheetah's step lengths are completely random and this model takes this into account. The search strategy is shown in Figure 2a [16].

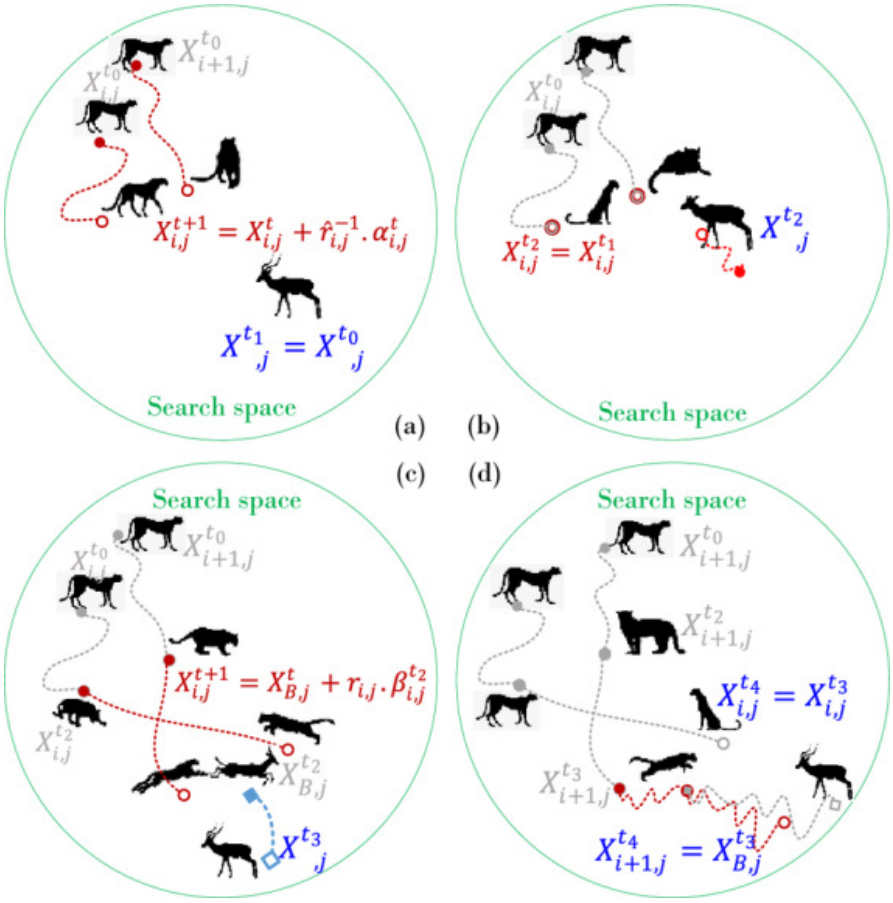


Figure 2: Cheetah Optimization Algorithm Modeling [16]

Watching the Prey

Cheetahs must display a high level of attention and control while hunting, as even the slightest movement can be detected by the prey and trigger an escape response. In order to minimize this situation, the cheetah stops moving and lies in wait when it thinks it is close enough to its prey. This strategy is based on the cheetah maintaining its position and waiting for the prey to come within range on its own. This waiting and position fixing behavior is mathematically modeled as in equation 10:

$$X_{i,j}^{t+1} = X_{i,j}^{t}, \quad (10)$$

This strategy is a critical factor that increases the hunting success of the cheetah optimization algorithm. In particular, by ensuring that not all cheetahs change their location at the same time, the risk of premature convergence of the algorithm is prevented. In this way, a more comprehensive exploration process is carried out in the solution space and the probability of getting stuck in local minima is significantly reduced.

Attacking the Prey

Cheetahs use two basic elements, speed and flexibility, together when catching their prey. When a cheetah decides to attack its prey, it starts the pursuit process by running at maximum speed. However, when the prey notices the cheetah approaching, it suddenly starts to run away.

At this point, the cheetah carefully analyzes the direction of movement of its prey and determines the most suitable attack route, aiming to intercept the prey at a certain point.

While the cheetah approaches its prey at maximum speed, the prey must suddenly change direction and attempt to escape in order to survive. However, thanks to both its speed and flexibility, the cheetah can respond to these sudden maneuvers of its prey and continue its pursuit by instantly adjusting its direction.

If hunting takes place in a group, each cheetah must dynamically adjust its position according to the positions of the escaping prey, the leader cheetah, or other cheetahs nearby. This strategy allows the cheetahs to catch the prey in the most efficient way. These attack tactics of the cheetahs are mathematically defined by equation number 11.

$$X_{i,j}^{t+1} = X_{\beta,j}^{t} + r_{i,j} \cdot \beta_{i,j}^{t} \quad (11)$$

Here is the current position of the prey in the jth plane. In other words, it is the best current position within the cheetah population. In attack mode, it allows the cheetahs to get as close as possible to the prey's position in a short time using maximum speed. Therefore, the new position of the i. cheetah in attack mode is calculated according to the prey's current position. r(i,j) and are the turn factor and interaction factor values of the i. cheetah in the jth plane, respectively. The turn factor reflects the interaction between the cheetahs or between a cheetah in capture mode and the leader. This factor reflects the sharp turns of the cheetahs in capture mode and is mathematically defined as the difference between the position of the neighboring cheetah and the position of the i. cheetah.

COA stands out as a powerful and efficient method in the feature selection process, especially in high-dimensional data sets. This algorithm naturally has the ability to discover local and global optima and quickly scans the solution space to eliminate unnecessary features and select the most meaningful features in order to increase the accuracy of the model. An important advantage of COA is that it provides faster convergence compared to genetic algorithms and other evolutionary techniques. This speed saves time in feature selection when working with large data sets and allows the processing time to be minimized. This fast and efficient structure of the Cheetah Optimization Algorithm increases the accuracy of the model, especially in predictive maintenance applications, and offers the opportunity to use computational resources more efficiently. For this reason, Cheetah Optimization was preferred in feature selection in our PdM study; because the algorithm optimizes the overall performance of the model by determining the most appropriate feature subset with high accuracy rates.

Performance Metrics

Performance metrics are critical to show how effective and reliable the model is. These metrics are used to evaluate the accuracy, reliability and efficiency of the model. The confusion matrix is the basis for these calculations.

The confusion matrix, which shows the classification performance of the model, consists of four basic components:

•True Positives (TP): The number of examples where the model correctly predicted the positive class.

•False Positives (FP): The number of examples where the model falsely predicted the negative class as positive.

•True Negatives (TN): The number of examples where the model correctly predicted the negative class.

•False Negatives (FN): The number of examples where the model falsely predicted the positive class as negative.

The confusion matrix is a valuable tool to analyze which classes the model is successful in and which classes it is incorrectly classified in. The main metrics used to evaluate the performance of the model in this study are listed below:

Accuracy (Acc): The accuracy of a model is evaluated by how well its predictions actually work. It is the ratio of the number of examples the model correctly classified to the total number of examples.

$$Acc = \frac{TP + TN}{\mathrm{TP} + \mathrm{TN} + \mathrm{FP} + \mathrm{FN}} \tag{12}$$

Precision (Pre): Precision is a measure that evaluates how accurate the model's positive class predictions are. Precision is especially important in situations where false positives are critical.

$$Pre = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FP}} \tag{13}$$

Recall (Rec): Measures how well the model can identify all positive cases.

$$Rec = \frac{\mathrm{TP}}{\mathrm{TP} + \mathrm{FN}} \tag{14}$$

F1-Score (F1): Is the harmonic mean of precision and sensitivity. This metric balances both values, providing a more objective measure of the model's overall success. The F1 score is used when a balance needs to be established between sensitivity and precision.

$$F1 = 2 \text{ x } \frac{\text{Precision x Recall}}{\text{Precision} + \text{Recall}} \tag{15}$$

Experimental Results

In this study, the performance of the proposed model for predictive maintenance systems is evaluated through the performance of DT, SVM and k-NN algorithms used in the classification of machine failure type and four different failure modes. In the experimental process, the performances of each of these algorithms were evaluated independently and the results were compared. Different performance measures of the model such as accuracy, precision, sensitivity and F1 score were calculated and the effectiveness and applicability of PdM systems were evaluated in line with the results obtained. In order to eliminate the class imbalance in the data set, a balanced data distribution was achieved by using the SMOTE method. Compared to the original imbalanced dataset, the balanced dataset obtained with the SMOTE method was found to perform better. Min-max and z-score normalisation operations were applied to the balanced dataset and the perfor-

mances of the algorithms after both normalisations were compared. As a result of the evaluations, it was observed that the z-score normalised data performed better than the min-max normalised data.

As a result of the evaluations, it was observed that the data to which z-score normalization was applied exhibited higher performance than the data to which min-max normalization was applied.

In addition, classification operations were performed between the TWF-HDF, TWF-PWF, TWF-OSF, HDF-PWF and PWF-OSF failure modes, which were divided into binary classes among themselves, and it was observed that z-score normalization increased the success for each mode.

In the last stage, the most critical features that increased the success were selected among the five features in the data set using the Cheetah Optimization Algorithm (COA), and thus the highest success rates were obtained. In this process, unimportant features were determined for each failure mode and the classification performance was improved. It was observed that the removal of some features in particular provided a significant improvement in performance metrics. The findings obtained show that the proposed methods and the selected features can significantly increase the efficiency in the classification of machine failures and modes. These findings can be explained in detail as follows: The original data set used in the study contains a total of 10,000 data samples, 339 of which represent the machine fault condition (MF) and 9,661 of which represent the non-fault machine condition (N-MF). This distribution shows that the data set has a serious class imbalance. In the analyses performed in this section without applying any balancing method, 98.0% accuracy rates were obtained for the DT algorithm, 97.1% for the SVM algorithm and 96.8% for the k-NN algorithm.

The confusion matrix of the DT algorithm with the highest performance is presented in Figure 3. The SVM algorithm showed the highest success with a precision value of 86.7%. The F1 values were calculated as 84.2% for DT, 66.6% for SVM and 74.0% for k-NN, respectively. When the cheetah optimization algorithm was applied on the original imbalanced data set, a decrease in model performance was observed as a result of removing all features.

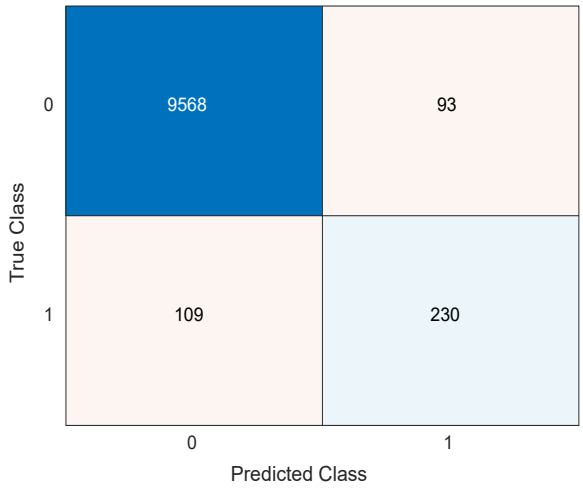


Figure 3: Confusion Matrix Of The Highest Performing DT Algorithm

In order to eliminate the class imbalance in the dataset, the minority class of 339 machine failure cases (MF) was increased to 9,492 samples by applying the SMOTE method. The DT, SVM and k-NN models achieved very successful results with 97.7%, 98.4% and 99.2% accuracy rates, respectively. However, since the accuracy rates only show the overall success of the model, other metrics should also be considered for a more detailed evaluation. In the dataset balanced with SMOTE, it is seen that the k-NN model achieved the highest performance with 99.2% F1 score, followed by the SVM with 98.3% success rate and the DT models with 97.5%. This shows that the k-NN model provides a strong balance between recall and precision.

As a result, in this balanced dataset, increasing the data caused positive changes in the performance of the models. The k-NN model became the leader in overall performance by distinguishing more classes. The confusion matrix of the k-NN model is given in Figure 4.

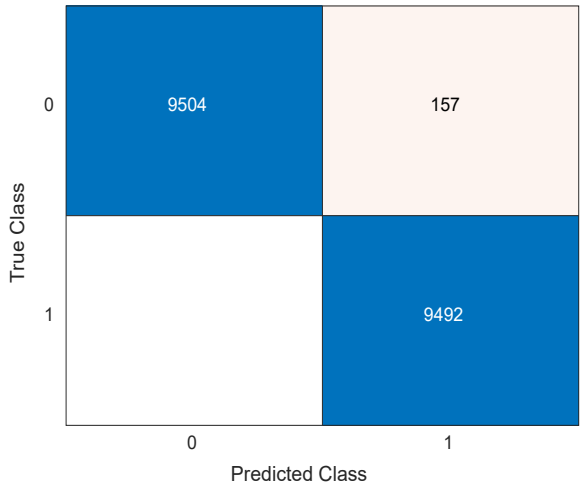


Figure 4: Confusion Matrix Of The Highest Performing k-NN Model

In the analyses performed on the balanced data set with the SMOTE method, the model performance was optimized using the cheetah optimization algorithm. Initially, when all features were used, the DT, SVM and k-NN models achieved high success rates with 97.7%, 98.4% and 99.2% accuracy rates, respectively. When the cheetah optimization algorithm was used, all features were selected, which revealed that each feature was important in the balanced data set and the contribution of each feature played a critical role in increasing the success of the model. As a result of the analyses, it was determined that the highest model success belonged to the k-NN model and this result is shown in Figure 5.

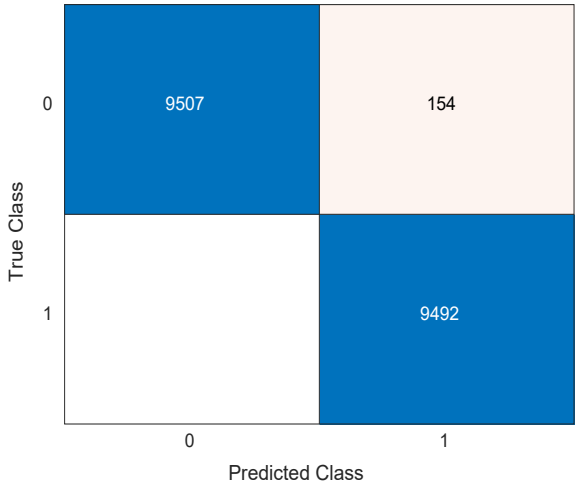


Figure 5: Confusion Matrix Showing The Performance Of The k-NN Model

The performance analysis of failure modes was also initiated with the analyses performed before applying the cheetah optimization algorithm. At this stage, the original dataset was used for each failure class and all features were included in the model.

The numbers of the dataset are as follows:

- TWF-HDF: 46 TWF, 115 HDF
- TWF-PWF: 45 TWF, 94 PWF
- TWF-OSF: 43 TWF, 95 OSF
- HDF-PWF: 112 HDF, 92 PWF
- HDF-OSF: 109 HDF, 92 OSF
- PWF-OSF: 83 PWF, 86 OSF

At this stage, when all features are used, the model success rates are 98.8% for the DT model, 90.1% for the SVM and 97.5% for the k-NN class. For the TWF-PWF class, these rates were calculated as 97.1% for DT, 92.1% for SVM and 95.7% for k-NN. In the TWF-OSF analysis, it was observed that the SVM model achieved a low performance of 75.4% accuracy and 60.5% sensitivity. However, in other models, it was observed that the accuracy of the DT model was 97.1% and the k-NN model was 93.5%. In the HDF-PWF class, DT achieved 95.1%, SVM 97.1% and k-NN 95.6% success rates, while in the HDF-OSF class, DT

reached 95.0%, SVM 93.0% and k-NN 95.5%. In the PWF-OSF class, the DT model showed 98.8% success rates, while the SVM and k-NN models showed 98.2% success rates. These results revealed that although the models generally showed high performance when all features were included, some models offered low success rates for certain classes. The performance analysis of failure modes and classification algorithms are given in detail in Table 2.

Table 2: Performance Metrics Of Classification Algorithms For Failure Modes (Before Cheetah Optimization Algorithm)

FAILURE MODE	Algorithms	Rec (%)	Pre (%)	Acc (%)	F1 (%)
TWF-HDF	DT	98.5	98.5	98.8	98.5
	SVM	81.1	91.8	90.1	84.5
	k-NN	94.1	96.7	97.5	95.3
TWF-PWF	DT	96.7	96.7	97.1	96.7
	SVM	87.8	94.8	92.1	90.3
	k-NN	95.1	95.1	95.7	95.1
TWF-OSF	DT	96.0	97.2	97.1	96.6
	SVM	60.5	86.8	75.4	59.7
	k-NN	90.8	93.9	93.5	92.1
HDF-PWF	DT	95.1	95.1	95.1	95.1
	SVM	97.1	97.0	97.1	97.0
	k-NN	95.9	96.1	95.6	96.0
HDF-OSF	DT	93.2	92.9	95.0	93.0
	SVM	94.4	94.1	93.0	94.0
	k-NN	96.5	96.5	95.5	96.5
PWF-OSF	DT	98.8	98.8	98.8	98.8
	SVM	98.2	98.2	98.2	98.2
	k-NN	97.0	98.1	98.2	98.0

Finally, the Cheetah Optimization Algorithm, which was applied to increase the model performance in the classification of failure modes, significantly improved the classification performance. In this direction, the performance metrics obtained after the algorithm and the observed improvements are evaluated in detail below. According to the table, the accuracy rates increased significantly after feature selection.

The accuracy rates of the failure modes grouped based on the binary classification structure are presented comparatively in Figure 6, before and after feature selection. According to the results obtained, the highest success rate belongs to the PWF-OSF class with 99.41%. In addition, the accuracy rates in the TWF-HDF and TWF-PWF classes reached 98.76% and 97.84%, respectively. The Cheetah optimization algorithm increased the overall success of the model by excluding unnecessary features and selecting only the effective ones. The selected features generally consist of critical parameters such as air temperature, rotation speed, torque and tool failure. As a result of the analysis, the features selected by the Cheetah Optimization Algorithm and the performance metrics re-

lated to the effects of these features on the relevant failure modes are presented in Table 3. These features play a decisive role in the accuracy of the model. It can be said that the Cheetah optimization algorithm increases the computational efficiency of the model as well as increasing the accuracy rate. Selecting the rotation speed feature, which is common to almost all classes, provided significant improvements in model performance. These results show that selecting or removing certain features in each failure mode can positively affect model performance. As a result, the Cheetah optimization algorithm stands out as an effective method in selecting the features correctly and improving the performance of the model.

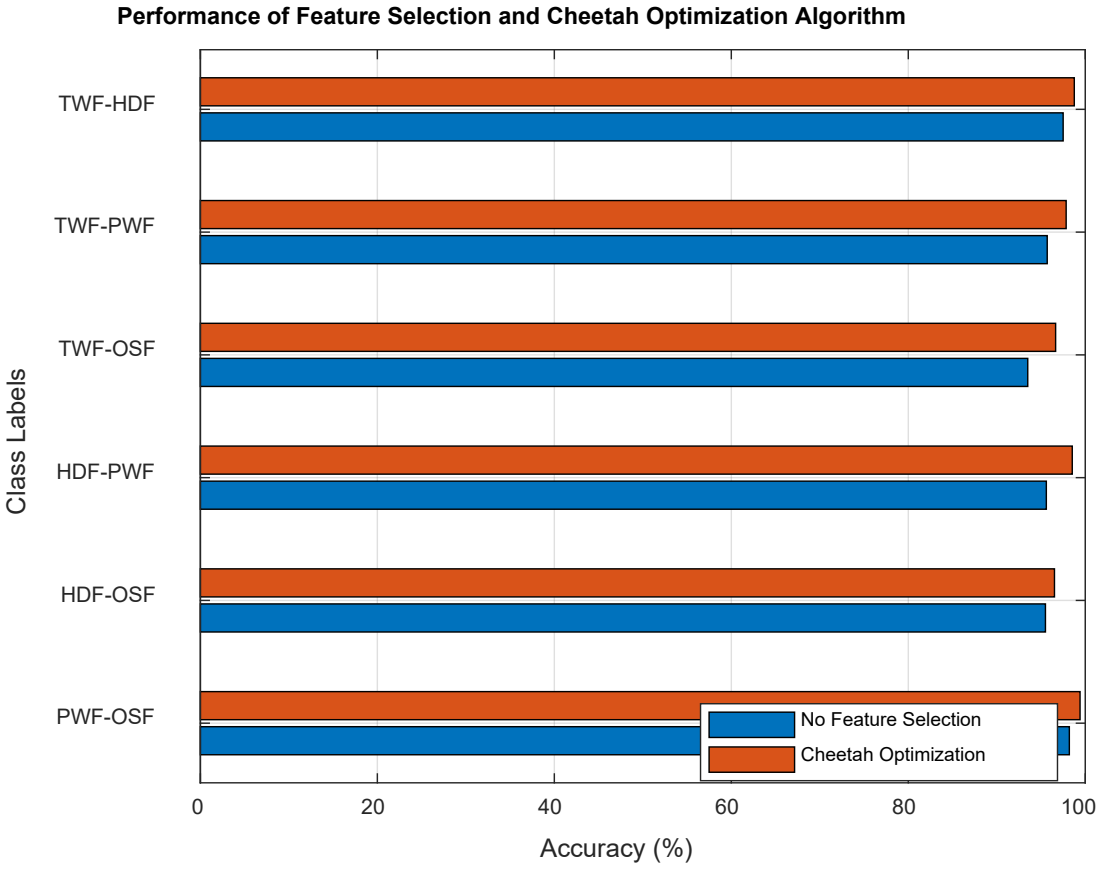


Figure 6. Graphical Representation of Accuracy Rates Before and After Feature Selection for Failure Modes

Table 3: Performance Metrics And Selected Features Of The Classification Algorithms Of Failure Modes (After Cheetah Optimization Algorithm)

FAILURE MODE	Algorithms	Rec (%)	Pre (%)	Acc (%)	F1 (%)	Selected Features
TWF-HDF	DT	98.5	98.5	98.8	98.5	Air Temperature, Rotation Speed, Tool Wear
	SVM	99.1	97.9	98.8	98.5	
	k-NN	98.9	99.6	99.4	99.2	
TWF-PWF	DT	97.2	97.8	97.8	97.5	Rotational Speed, Torque, Tool Wear
	SVM	96.7	98.5	97.8	97.5	
	k-NN	98.9	99.5	99.3	99.2	
TWF-OSF	DT	99.5	98.9	99.3	99.2	Air Temperature, Torque
	SVM	97.1	97.8	98.6	97.4	
	k-NN	98.3	98.0	99.3	98.3	
HDF-PWF	DT	95.6	95.5	94.6	95.5	Rotational Speed, Torque
	SVM	98.7	98.4	98.0	98.5	
	k-NN	98.6	98.5	98.5	98.5	
HDF-OSF	DT	93.0	93.0	94.5	93.0	Air Temperature, Rotational Speed, Torque, Tool Wear
	SVM	96.5	96.5	95.5	96.5	
	k-NN	94.9	95.1	96.5	95.0	
PWF-OSF	DT	98.8	98.8	98.8	98.8	Rotational Speed, Tool Wear
	SVM	99.4	99.4	99.4	99.4	
	k-NN	100.0	100.0	100.0	100.0	

Declaration of Interest: The authors declare that there is no conflict of interest in this study.

Acknowledgements: We would like to thank Zonguldak Bülent Ecevit University for their support in the preparation of the study.

Author Contributions: Nuriye Bektaşoğlu: Writing – original draft, Methodology, Investigation, Formal analysis, Software, Conceptualization. Ali Narin: Writing – original draft, Methodology, Investigation, Formal analysis, Software, Conceptualization.

REFERENCES

[1] Özbey, Ö. (2024). *Endüstri 4.0 ve Malatya Organize Sanayi Bölgesi uygulanabilirlik araştırması* (Master›s thesis, Balikesir University (Turkey)).

[2] Abbassi, R., Arzaghi, E., Yazdi, M., Aryai, V., Garaniya, V., & Rahnamayiezekavat, P. (2022). Risk-based and predictive maintenance planning of engineering infrastructure: existing quantitative techniques and future directions. *Process Safety and Environmental Protection*, *165*, 776-790.

[3] Bojana, N., Jelena, I., Suzic, N., Branislav, S., & Aleksandar, R. (2017). Predictive manufacturing systems in industry 4.0: Trends, benefits and challenges. In *Proceedings of 28th DAAAM International Symposium on Intelligent Manufacturing and Automation* (pp. 796-802). DAAAM International, Vienna, Austria.

[4] Ronzoni, N., De Marco, A., & Ronchieri, E. (2022, July). Predictive Maintenance Experiences on Imbalanced Data with Bayesian Optimization Approach. In International Conference on Computational Science and Its Applications (pp. 120-137). Cham: Springer International Publishing.

[5] Baccarini, L. M. R., e Silva, V. V. R., de Menezes, B. R., & Caminhas, W. M. (2011). SVM practical industrial application for mechanical faults diagnostic. *Expert Systems with Applications*, *38*(6), 6980-6984.

[6] Zhang, Z., Wang, Y., & Wang, K. (2013). Fault diagnosis and prognosis using wavelet packet decomposition, Fourier transform and artificial neural network. *Journal of Intelligent Manufacturing*, *24*, 1213-1227.

[7] Rayarao, S. R., & Rayarao, S. R. (2024). Advanced Predictive Maintenance Strategies: Insights from the AI4I 2020 Dataset. Authorea Preprints.

[8] Sengupta, P., Mehta, A., & Rana, P. S. (2023). Predictive Maintenance of Armoured Vehicles using Machine Learning Approaches. *arXiv preprint arXiv:2307.14453*.

[9] Ileri, U., Altun, Y., & Narin, A. (2024). An Efficient Approach for Automatic Fault Classification Based on Data Balance and One-Dimensional Deep Learning. *Applied Sciences*, *14*(11), 4899.

[10] Rajendiran, V. (2023). *Smart Predictive Maintenance for pro-active machine process using Machine Learning* (Doctoral dissertation, Swinburne).

[11] Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research*, *16*, 321-357.

[12] Navada, A., Ansari, A. N., Patil, S., & Sonkamble, B. A. (2011, June). Overview of use of decision tree algorithms in machine learning. In *2011 IEEE control and system graduate research colloquium* (pp. 37-42). IEEE.

[13] Amarappa, S., & Sathyanarayana, S. V. (2014). Data classification using Support vector Machine (SVM), a simplified approach. *Int. J. Electron. Comput. Sci. Eng*, *3*, 435-445.

[14] Dhanabal, S., & Chandramathi, S. J. I. J. C. A. (2011). A review of various k-nearest neighbor query processing techniques. *International Journal of Computer Applications*, *31*(7), 14-22.

[15] Çıkan, M. (2024). Çita optimizasyon algoritması kullanarak kısmi gölgelenme altındaki fotovoltaik sistemlerde maksimum güç noktası izleyicisinin tasarlanması. Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, 40(1), 555-572.

[16] Akbari, M. A., Zare, M., Azizipanah-Abarghooee, R., Mirjalili, S., & Deriche, M. (2022). The cheetah optimizer: A nature-inspired metaheuristic algorithm for large-scale optimization problems. *Scientific reports*, *12*(1), 10953.

PLATON VE ARİSTOTELES PERSPEKTİFİNDEN YAPAY ZEKÂ

Burçak Dinçergök*

Ankara Hacı Bayram Veli Üniversitesi, Lisansüstü Eğitim Enstitüsü, Felsefe Anabilim Dalı, Ankara, Türkiye

(İletişimden sorumlu yazar) burcakdncrgk@gmail.com

ÖZET

İnsan davranışlarını ve zekasını taklit ederek üretilecek bir makine tartışmalı bir konudur. Bu tartışmanın temelinde üretilecek böyle bir makinenin yaşam üzerindeki etkisinin olumlu mu yoksa olumsuz mu olacağı vardır. Bu metin, öncelikle tarihsel gelişimi içinde yapay zekanın olumlu ve olumsuz etkilerini incelemektedir. Ardından yapay zekanın olumlu ve olumsuz etkilerini Platon ve Aristoteles bağlamında ele almaktadır. Böylece Platon ve Aristoteles felsefeleri analiz edilerek yapay zekanın değeri üzerine bir sorgulama yapılmaktadır.

Anahtar Kelimeler: Platon, Aristoteles, yapay zeka.

Giriş

İnsana benzeyen makinelerin üretimine dair yazılara ilk olarak mitlerde rastlanır. Daha sonra otomatlar yapılarak insana benzeyen ilk makineler üretilmiştir. Günümüze gelindiğinde insan benzeri olan bütün sistemler yapay zeka başlığı altında ele alınmaya başlanmıştır. Bu gelişmeler beraberinde etik birçok soruyu da gündeme getirmektedir. Bu sorular, hem olumlu hem de olumsuz perspektiflerden incelenmelidir. Böyle bir bakış açısı ile yapay zekaya yaklaşan bu metin, Platon ve Aristoteles perspektiflerinden hareketle yapay zekanın bir değere sahip olup olmadığını tartışmaktadır.

Yapay Zekaya Genel Bir Bakış

Türk Dil Kurumu'na göre zeka, *"İnsanın düşünme, akıl yürütme, öğrenme, kavramları ve nesneleri zihinde canlandırabilme, objektif gerçekleri algılama, yargılama, sonuç çıkarma, bedeni kontrol edebilme, duyguları doğru algılayabilme, değerlendirebilme, icat edebilme vb. yeteneklerinin ve becerilerinin tamamı (...)"[1]* dır. Yapay zeka ise *"Bir bilgisayarın, bilgisayar kontrolündeki bir robotun veya programlanabilir bir aygıtın insana benzer biçimde algılama, öğrenme, fikir yürütme, karar verme, sorun çözme, iletişim kurma vb. işlevleri sergileyebilme yeteneği (...)" [2]* olarak tanımlanmaktadır. Buradan hareketle yapay zekanın insan benzeri davranış ve düşünme sergileyen sistemlerin ortak adı olduğu söylenebilir. Böyle bir sistemin oluşturulabilmesi için yapay zekanın insan davranışlarını öğrenmesi ve insanın yapması gereken görevleri yerine getirmesi gerekir. Çünkü yapay zeka, zeki olarak nitelenen insan davranışlarının makine tarafından yapılmasını içerir [3]. Bu alanda çalışan yazılımcılar algoritmalar ve veriler aracılığıyla yapay zekanın insan davranışlarını taklit etme

yeteneğini geliştirmek ister. Buradaki ölçüt, yapay zekanın öğrenen bir makine olmasıdır [4]. Algoritma ve verilerin öğrenmesi için gösterilen çabanın sonucunda bir olay karşısında insan gibi tepki verebilme veya bir sorun karşısında insan gibi mantığını kullanabilme becerilerini kazanması beklenmektedir. Dolayısıyla bu alanın temelinde yapay bir zekanın insan zekasını ve davranışlarını ne kadar taklit edebileceği sorusu vardır.

İnsan özelliklerinin taklit edilip edilemeyeceğine dair duyulan merak, binlerce yıl önce başlamıştır. Antik bilim tarihçisi ve folklorist olan Adrienne Mayor'a göre robot temelli teknoloji kökenini 2700 yıl önceki mitlerden alır [5]. Antik Yunan mitolojisinde insan taklit edilerek ortaya çıkan ilk figür ise Talos'tur. Mitolojiye göre Zeus, sevdiği Europa'ya hediye etmek için zanaat ve ateş tanrısı Hephaistos'a Talos'u yaptırır. Bronzdan yapılmış dev bir robot görünümünde olan Talos, Europa'yı ve çocukları korumakla görevlendirilir. İnsan gibi hareket edebilen Talos, Girit sahiline yaklaşan yabancı gemileri taşlayarak batırır veya adada gördüğü yabancı insanları öldürür. Böylece Talos, insana benzeyen makinelerin üretimine dair duyulan hayalin ne kadar eskiye dayandığını göstermektedir. Ayrıca Talos'un üreticisi olarak bilinen Hephaistos, Talos gibi kendi kendine hareket ederek görevleri yerine getirebilen birçok robotun da üreticisi olmuştur. Hephaistos'un sahip olduğu bu yetenek ile yapay zeka alanındaki pek çok gelişmenin ilk habercisi olduğu söylenebilir.

Antik Yunan mitolojisinde yapay zeka izleri taşıyan bir diğer figür, heykeltıraş Pygmalion'un yaptığı kadın heykel Galatea'dır. Galatea'nın güzelliğinden etkilenen Pygmalion, heykele aşık olmaktan kendini alıkoyamaz ve canlanması için günlerce yalvarır. Pygmalion'un bu çağrısına kayıtsız kalamayan Afrodit, heykeli canlandırarak insan yapar ve ikilinin evlenmesine vesile olur. Pygmalion ve Galatea'nın ana karakterler olduğu bu mitin yaratılan bir varlığın bir gün insan gibi akla ve davranışa sahip olabileceğinin işaretlerini içerdiği söylenebilir.

Antik Yunan mitolojisinde yapay zeka açısından ele alınması gereken bir diğer figür, titan Prometheus'tur. Prometheus, bilgiyi ve gücü insanlara sunan mitolojik bir karakterdir. Mitolojiye göre balçıktan yaratılan ilk insan güçsüzdür. İnsanın bu güçsüz haline acıyan Prometheus, Hephaistos'un ocağından ateş çalar ve bunu insanlara getirir. İnsana ateş kullanmak, yiyecek pişirmek ve alet yapmak gibi birçok şeyi öğretir. Bir süre sonra insanlar, bunları kendilerinin öğrendiklerini söyleyerek Prometheus'un emeğini görmezden gelmeye başlar. Prometheus'un ana kahramanı olduğu bu mitin insanın bilgisi ile üretilen yapay zekanın bir gün insanların yerini alacağı ile ilgili birçok problemi gündeme getirdiği söylenebilir. Bu problemin devamı olarak görülen bir başka Antik Yunan miti ise Hephaistos'un elinden çıkan Pandora'dır. Mitolojiye göre güzel ve zarif bir kadın olarak yaratılan Pandora'ya bir kutu verilerek açmaması gerektiği söylenir. Merakına yenik düşen Pandora, kutuyu açar ve kutuyu açar açmaz dünyaya kötülük yayılır. Bu mitin ise akıllara yapay zeka gibi yaratılan bir varlığın bir gün kontrolsüz bir güce sahip olarak insanlara zarar verip vermeyeceği sorusunu getirdiği söylenebilir.

13 ve 19. yüzyıl arası insanlara benzeyen makinelere dair üretimlerin başladığı ve bu konuyla ilgili eserlerin ilgi çekmeye başladığı yüzyıllardır. 13. yüzyılda Albertus Magnus'un insan kafası şekline sahip bakırdan bir otomat ürettiği bilinmektedir. Magnus'un ürettiği bu otomat, ağız açma özelliğine sahiptir. Bu otomatın biraz daha gelişmiş versiyonunu 15.

yüzyılda Leonardo da Vinci üretir. 1495 yılında üretilen bu otomat, insan hareketlerini taklit edebilen bir özelliğe sahiptir. Leonardo da Vinci'nin otomat çizimleri, her ne kadar eksik parçalara sahip olsa da, yüzyıllar sonra robot projeleri yapmak isteyen araştırmacılar tarafından kullanılmıştır. Bir 17. yüzyıl filozofu olan René Descartes'ın ise "Metot Üzerine Konuşma" kitabındaki fikirleri ile yapay zekadaki tartışmaları öncelediği söylenmelidir. Bahsi geçen kitabın beşinci bölümünde şu cümleler yer alır:

> *"(...) bedenlerimizle benzerlik taşıyan ve ahlâken bizim davranışlarımızı ne kadar mümkünse o kadar taklit eden makineler olsaydı, bundan dolayı onların asla hakikî insanlar olmadığını teşhis etmek için elimizde her zaman çok kesin iki vasıta bulunacaktır. Bu vasıtalardan ilki şudur: onlar hiçbir zaman, düşüncelerimizi diğer insanlara bildirmek için yaptığımız gibi, sözleri ve başka işaretleri birleştirerek kullanamayacaklardır. (...) İkinci vasıta ise şudur: birçok şeyi bizim kadar ve belki bizden de iyi yapsalar bile, yapamayacakları tartışma götürmeyen şeyler vardır ki bu da onların bilgiyle değil, yapılarına bağlı temayülle/hazır bulunuşlukla hareket ettiklerini gösterir [6]."*

Descartes, insana benzeyen bir makinenin ruha sahip olan gerçek bir insan gibi düşünemeyeceğini savunmaktadır. Bu fikirleri ile yapay zekanın insan zekasına ne kadar benzeyebileceği ile ilgili bir tartışmayı ele almak istediği söylenmelidir. 1726 yılında Jonathan Swift'in yazdığı "Gulliver'in Gezileri" ve 1872 yılında Samuel Butler'in yazdığı "Erewhon" isimli kitaplar ise bilimsel gelişmelerin yoğun olduğu çağlarda yaşayan insanların durumunu anlatmaktadır. Böyle anlatımlar, insana benzeyen makine üretimlerinin her çağda insanlar için olumlu veya olumsuz bir karşılık bulduğunu göstermektedir.

İnsana benzeyen makineler hem üretim hem de merak anlamında 20. yüzyılda hızlı bir ilerleme göstermiştir. 1921 yılında Karel Čapek, "Rossum'un Evrensel Robotları" isimli bilimsel kurgu tiyatro oyunu ile "robot" kelimesini kullanan ilk kişi olmuştur. Čapek'in "yapay insanlar" olarak tanımladığı bu robotların oyundaki görevi, insan benzeri iş gücü üretmektir. 1927 yılında Fritz Lang'ın yönettiği bilim kurgu filmi olan "Metropolis", içinde barındırdığı makine insan karakteri ile insan benzeri makinelerin kullanımına dair bir fikri izleyiciye sunar. 1929 yılında Japon biyolog Makoto Nishimura'nın ürettiği "Gakutensoku" isimli robot, Japonya'nın ilk robotudur. Ayrıca Gakutensoku başını hareket ettirme, mimiklerini kullanma ve kalem ile yazabilme gibi insan özelliklerini yerine getiren bir mekanizmaya sahiptir. 1936 yılında Alan Turing, "On Computable Numbers, with an Application to the Entscheidungsproblem" isimli makalesinde matematiksel hesaplamalar yapabilecek bir makinenin çalışma prensiplerinden bahsetmiş ve böylece "Turing Makinesi"ni geliştirmiştir. Bu, insanın zekasına ait olarak görülen bir yeteneğin bir makine tarafından yapılabileceğinin gösterimi olmuştur. 1940 yılında John Atanasoff ve Clifford Berry geliştirdikleri "ABC" isimli bilgisayar ile bir makinenin doğrusal denklem çözebilmesini sağlamıştır. Atanasoff, "Computing Machine for the Solution of Large Systems of Linear Algebraic Equations" isimli yazısında böyle bir bilgisayarın amacını *"Ondan fazla bilinmeyene sahip genel doğrusal denklem sistemlerinin çözümü sıklıkla denenmez. Ancak bu, pratik problemlerin çözümünde yaklaşık yöntemleri daha etkili hale getirmek için gereken şeydir [7]."* sözleriyle açıklar. Bu açıklama, üretilecek bu tarz makinelerin insanın da hayatını kolaylaştıracağını ileri sürmektedir. 1949 yılında Edmund Callis Berkley yazdığı "Dev Beyinler veya Düşünen Makineler" isimli kitap ile üretilecek makinelerin bir gün insan gibi zihinsel işlemler yapabileceğini belirtir. Her ne kadar yazıldığı dönemde kitabın

içindeki görüşler hayalci olarak karşılanmış olsa da kitabın insana benzeyen makinelerle ilgili toplumsal farkındalık oluşturmak istediği söylenmelidir.

1950-1980 yılları yapay zekanın bir çalışma alanı olarak görülmeye başlandığı ve pek çok gelişmenin yaşandığı yıllardır. 1950 yılında Claude Shannon, "Programming a Computer for Playing Chess" isimli makalesinde satranç oynayabilen bir bilgisayar programının nasıl geliştirilebileceğini anlatmış ve böylece bu amaç için kurulacak bir algoritmanın teorik temellerini göstermiştir. 1950 yılında Alan Turing, geliştirdiği "Turing Testi" ile bir makinenin bir insanı ne kadar taklit edebileceğini test etmek istemiştir. Turing yazdığı "Computing Machinery and Intelligence" isimli makalesinde bu testin amacının "Makineler düşünebilir mi?" sorusunu incelemek olduğunu belirtir [8]. Bu testte bir insan, bir makine ve bir sorgulayıcı insan vardır. Sorgulayıcı insan, makine ve insandan ayrı bir odada bulunur ve aralarında sadece yazılı iletişim kurulur. Sorgulayıcı insanın amacı, sorularına verilen cevaplardan hareketle hangisinin insan hangisinin makine olduğunu bulmaktır. Eğer sorgulayıcı bunu başaramazsa makine insan gibi düşünebiliyor sayılır. 1956 yılında düzenlenen "Dartmouth Konferansı" ise makinelerin insan benzeri zekaya ve davranışa sahip olup olamayacağının bilimsel düzeyde tartışıldığı bir içeriğe sahiptir. Bu konferansın önerisi şöyledir:

> *"1956 yazında Hanover, New Hampshire'daki Dartmouth Koleji'nde 2 aylık, 10 kişilik bir yapay zeka çalışmasının yapılmasını öneriyoruz. (...) Makinelerin dili nasıl kullanacaklarını, soyutlamaları ve kavramları nasıl oluşturacaklarını, insanlar için ayrılmış problem türlerini nasıl çözeceklerini ve kendilerini nasıl geliştireceklerini bulmaya çalışılacaktır [9]."*

Konferansta John McCarthy, makinelerin düşünebilmesiyle ilgili fikirlerin "yapay zeka" başlığı altında toplanması gerektiğini belirtmiş ve böylece "yapay zeka" terimini kullanan ilk kişi olmuştur. 1958 yılında ise McCarthy, "LISP" programlama dilini geliştirmiş ve LISP, yapay zeka üreticilerinin uzun yıllar tercih ettiği bir programlama dili olmuştur. 1959 yılında Arthur Samuel, "makine öğrenimi" terimini ilk kez kullanarak yapılan programlamalar aracılığıyla makinelerin gelişen bir deneyime sahip olabileceğini belirtmiştir. Samuel, bu fikrini dama oyunu oynayabilen bir yazılım ile gündeme getirmiş ve böylece bir bilgisayarın elde edeceği deneyim ile bir insandan daha iyi dama oynayabileceğini iddia etmiştir. 1961 yılında George Devol'un icat ettiği "Unimate" isimli robot, dünyadaki ilk endüstriyel robot olmuştur. Endüstriyel üretim hatları için görevlendirilen Unimate montaj, parça taşıma ve yerleştirme gibi görevleri yerine getirmiştir. 1965 yılında Edward Feigenbaum ve Joshua Lederberg "uzman sistemler"i geliştirerek bir alanda uzmanlaşmış insanların bilgi ve deneyimlerinin programlamalar tarafından kopyalanarak taklit edilebileceğini belirtmiştir. 1966 yılında Joseph Weizenbaum'un geliştirdiği doğal dilde iletişim kurabilen "ELIZA" ise ilk sohbet yazılımı olmuştur. Bu yazılımda bir psikoterapist rolü üstlenen ELIZA önce kullanıcının yazdığı sorunu analiz eder, daha sonra algoritmasındaki uygun olan şablonu kullanarak bir cevap verir. 1979 yılında James L. Adams'ın geliştirdiği "The Stanford Cart" kimsenin müdahalesi olmadan otonom harekette bulunabilen araçlardan biridir. Böylece insansız bir araç üretme fikri ilk defa gündeme gelmiştir. Görüldüğü üzere otuz yıllık bu süre, yapay zeka gelişmelerinin hız kazanmaya başladığı yıllardır.

1980-1987 yılları arasında yapay zeka çalışmalarının kazandığı hız artarak devam eder. 1980 yılında ticari bir uzman sistem olarak geliştirilen "XCON", mühendislere bilgisayar sis-

temlerinin yapılandırılmasında destek olarak daha verimli bir çalışma alanı sağlamıştır. 1986 yılında Ernst Dickmanns, Mercedes-Benz firması için sürücüsüz bir şekilde çalışan otomobil geliştirmiştir. 1987 yılına gelindiğinde yapay zeka alanındaki gelişmeler durağanlaşır ve 1993 yılına kadar "Yapay Zeka Kışı" yaşanır. Bu tanım, yapay zekaya karşı olan ilginin ve yatırımların azaldığı anlamına gelir. Böyle bir dönem hızla gelişen teknoloji pratikte uygulanamadığı, gerçekçi olmayan hedefler fark edildiği ve ilgi kaybına bağlı olarak araştırmalar azaldığı için yaşanmıştır.

1993-2011 yılları ise yapay zeka araştırmalarının yeniden hız kazanmaya başladığı yıllardır. Yeni yöntemlerin geliştirilmesi ile durağan dönem aşılmış ve alan yeniden canlılık kazanmıştır. 1997 yılında geliştirilen "Deep Blue", satranç oynayabilen bir uzman sistemdir. Dünya satranç şampiyonu Gary Kasparov'u yenerek insan bir rakibi yenebilen ilk uzman sistem olmuştur. 1997 yılında Windows sesli komutla kontrol sağlamak için "Dragon Sistem"i kullanmaya başlamıştır. Bu sistem, hem kullanıcının verdiği komutları algılayabilme hem de söylediğini metne dökebilme özelliğine sahiptir. 2000 yılında Cynthia Breazeal, geliştirdiği robot "Kismet" ile insan ve robot arasındaki sosyal iletişimi göstermek istemiştir. Kismet bir insan ile iletişime geçtiğinde insanın söylediklerine göre mutlu, kızgın veya üzgün gibi duygusal durumları yansıtabilmiştir. 2002 yılında iRobot şirketi, "Roomba" adındaki robot süpürgeyi piyasaya sürmüş ve böylece otonom hizmet alanına bir yenilik eklemiştir. 2003 yılında NASA, Mars'a iki tane insansız uzay aracı göndermiş ve bu araçlar insan müdahalesi olmadan Mars'ın yüzeyinde yürüyerek veriler toplamıştır. 2006 yılında sosyal medyanın yaygınlaşmasına paralel "UX" algoritması geliştirilerek sosyal medya kullanıcılarının hem deneyimleri iyileştirilmiş hem de reklam ağı canlandırılmıştır. 2010 yılında Microsoft, "Xbox 360 Kinect"i geliştirerek kullanıcıların sadece vücut hareketlerini kullanarak oyun oynayabilmesini sağlamıştır. 2011 yılında "Jeopardy!" isimli bir yarışma programına katılan "Watson" isimli yapay zeka, iki insanı yenerek yarışmayı kazanmıştır. 2011 yılında sanal bir asistan olarak geliştirilen "Siri" sesli komutları algılama, komut verme, telefon eylemlerini sağlama ve uygulamaları kullanma gibi birçok insan becerisini sağlamıştır.

2011 yılından günümüze kadar olan süreçte yapay zeka alanındaki yenilikler devam etmiştir. 2012 yılında Jeff Dean ve Andrew Ng, sinir ağlarındaki verilerin eğitilip eğitilemeyeceğini test etmek istemiştir. Üç gün boyunca rastgele açılan videolarla beslenen sinir ağı, üç günün sonunda videoların içinde en çok gördüğü figür olan kediyi kendiliğinden oluşturmuştur. İkiliye göre bu durum, makinelerin tıpkı insanlar gibi derin öğrenmeye sahip olabileceğini göstermiştir. 2014 yılında Microsoft tarafından geliştirilen "Cortana" ve Amazon tarafından geliştirilen "Alexa" sanal asistanları insanlara işlerini hatırlatmak, telefon uygulamalarını açmak ve sorularını yanıtlamak gibi pek çok alanda yardım etmektedir. 2016 yılında Hanson Robotics'in geliştirdiği "Sophia" isimli robot sahip olduğu insansı yüz ifadesi ve iletişim becerisiyle robot teknolojisinin yeniliklerinden olmuştur. 2017 yılında Google DeepMind'ın ürettiği "AlphaGo", karmaşık bir strateji oyunu olan Go'da insan rakiplerini yenmeyi başarmıştır. 2018 yılında Google'ın ürettiği "BERT" isimli doğal dil işleme sistemi, Google arama motorunda kullanılmıştır. Böylece arama motorunun anlamasının daha verimli olması ve arama sonuçlarının iyileştirilmesi hedeflenmiştir. 2020 yılında OpenAI, "GPT-3" isimli dil anlama ve üretme modelini üretmiştir. Bu model geliştirilerek 2022 yılında OpenAI tarafından "ChatGPT" isimli sohbet robotu haline getirilmiştir. Bu robot bilgi sağlama, metin yazma, çeviri yapma, sohbet etme ve asistanlık yapma gibi birçok alanda kullanılmaktadır. 2025

yılına gelindiğinde ise bir yandan yeni yapay zeka modelleri üzerine üretimlerin devam ettiği diğer yandan da yapay zekanın etik ve güvenli kullanımı için çalışmaların yapıldığı bilinmektedir.

Yapay zekanın tarihsel gelişimine bakıldığında bir yandan gelişim unsuru olarak görülürken diğer yandan da endişe unsuru olarak görüldüğü açıktır. Dolayısıyla insanın zekasını ve davranışlarını taklit eden bir makineye karşı sergilenecek olumlu veya olumsuz bir bakış, sadece çağımıza özgü değildir. Örneğin, 18. yüzyılda buhar makinesinin icadı ile başlayan Sanayi Devrimi sürecinde el emeği yerini makineye bırakmıştır. Bu durum kimileri için daha fazla ürünün piyasaya sürülmesiyle üretim, taşıma ve ekonomi gibi alanlarda verimliliğin artması demektir. Ancak kimileri için ise makinelerin işçilerin yerini alması ile başlayan iş kaybı, kötü çalışma şartları ve toplumsal eşitsizlik gibi problemler demektir. Benzer bir tablonun günümüzde de devam ettiği söylenebilir. Bu bağlamda günümüzde hızla geliştiği bilinen yapay zeka kimileri tarafından olumlu görülüp desteklenirken kimileri tarafından da olumsuz görülerek eleştirilmektedir.

Olumlu görenlere göre yapay zekanın yaşama katılmasıyla verimlilik, tasarruf, öneri ve insan hatalarının azalması konularında gelişim sağlanmıştır. Örneğin, Dragon Sistem'in söylenileni yazıya dökme özelliğine sahip olması hem fiziksel engeli olanlar hem de sürekli yazı yazanlar için kolaylık ve verim sağlamaktadır. XCON'un iş dünyasında kullanılması ile şirket sahipleri tasarruf yaptıklarını açıklamıştır. UX algoritmasının sosyal medya deneyimlerine paralel reklam öneriyor olması ile kullanıcılar, kendilerine doğru yönlendirmeler sunulmasından memnun kaldıklarını belirtmiştir. Üretilen ABC bilgisayarının insanların yapmaktan çekindiği ondan fazla bilinmeyene sahip genel doğrusal denklemi yapabiliyor olması ile insan kaynaklı ortaya çıkabilecek hataların önüne geçilmek istenmiştir. Bunlara ek olarak yapay zekanın eğitim, tıp gibi hayatın pek çok alanına sirayet ettiği söylenmelidir. Yapılan bir araştırmaya göre yapay zeka piyasasının 2024'ten 2030'a kadar olan bileşik yıllık büyüme oranının %36,6 olması beklenmektedir [10]. Sohbet robotu, sanal asistan ve dil çeviri sistemi gibi yapay zeka araçlarının kullanımının günümüzde gittikçe artması bu araştırmayı destekler niteliktedir.

Olumsuz görenlere göre yapay zekanın yaşama katılmasıyla işsizlik korkusu, etik sorunlar, insan zihninin pasifleşmesi, gizlilik ve güvenlik konuları gündeme gelmiştir. Örneğin, endüstriyel bir robot olan Unimate'in montaj, parça taşıma ve yerleştirme gibi görevleri yerine getiriyor olması bir gün insanların işsiz kalabilme ihtimalini gözler önüne sermektedir. Bir sohbet yazılımı olan ELIZA'nın psikoterapist rolü üstlenmiş olması ise birçok etik sorunu gündeme getirmektedir. Bu etik sorunların başında ELIZA'nın bir diplomaya sahip olmaması ve bu bağlamda danışana vereceği cevapların yanıltıcı olabileceği, danışanın ELIZA'ya karşı bağımlılık geliştirmesi durumunda nasıl bir tedbir alınacağı gibi konular vardır. Bir uzman sistem olan Deep Blue'nun satranç oyununda şampiyon olmuş bir insanı yenmesi ise yapay zeka ile insan zihninin pasifleştirileceğini gündeme getirmiştir. Böyle düşünenlere göre insanın uzmanlaştığı alanları taklit etmesi için üretilen uzman sistemlerin kullanımının artmasıyla insan zihni kendisini daha az kullanmaya başlayarak pasifleşecektir. Siri gibi sanal asistanların veri toplama ve depolama gibi özelliklere sahip olması ise veri sızıntısı ihtimalini akıllara getirerek gizlilik ve güvenlik ile ilgili eleştirileri ortaya çıkarmıştır. Son yıllarda gerek kişi gerekse ülkelerin aldığı kararlarla yapay zeka kaynaklı ortaya çıkabilecek

zararların en aza indirgenmek istendiği açıktır. 2015 yılında Elon Musk, Stephen Hawking, Steve Wozniak gibi isimler gelişen yapay zeka teknolojisinin potansiyel tehlikelerinin farkında olduklarını belirterek savaş amaçlı otonom sistem geliştirmeyeceklerini ifade eden açık mektup yayımlamışlardır. 2023 yılında ise dünya çapında yapay zeka kaynaklı güvenlik sorunları konuşulmaya başlanmış ve çeşitli tedbirler alınmaya başlanmıştır. Bu bağlamda 2024 yılında Avrupa Birliği tarafından "Yapay Zeka Yasası" yürürlüğe konmuş ve yasanın amacının yapay zekayı daha güvenli hale getirmek olduğu açıklanmıştır.

Görüldüğü üzere yapay zeka, yaşamı kolaylaştıran olumlu özelliklere sahip olduğu gibi üzerine derin analizler yapılması gereken olumsuz özelliklere de sahiptir. Dolayısıyla bir yapay zeka incelemesi sadece yapay zekanın pratik faydalarına değil, aynı zamanda yapay zeka kaynaklı ortaya çıkabilecek etik ve insan hakları problemlerine de odaklanmalıdır. Yapay zekaya olumlu ve olumsuz bakış bağlamında yürütülen bu sorgulama, Platon ve Aristoteles felsefeleri açısından da ele alınmaya uygundur. Bunu daha net bir hale getirmek için Platon ve Aristoteles perspektiflerinden hareketle yapay zekanın bir değere sahip olup olamayacağını tartışmak gerekmektedir.

Platon Perspektifinden Yapay Zeka

Platon ontolojisi gerçekliğin duyularla değil, idealların kavranmasıyla mümkün olduğunu savunur. Platon'a göre duyusal şeyler, idealların kusurlu birer kopyasıdır. Bunu mağara alegorisi ile açıklar:

> *"(...) Yeraltında, mağaramsı bir yerde insanlar oturuyorlar. Bu mağaranın girişi boydan boya gün ışığına açık. İnsanlar bu mağarada, çocukluklarından beri bacaklarından ve boyunlarından zincire vurulmuş bir şekilde; hiç kımıldamadan ve sürekli karşıya bakarak oturuyorlar. Zincirlerinden dolayı başlarını çeviremiyorlar. Arkalarından, uzakta ve yukarılarda yanan bir ateşin ışığı vuruyor. Yukarıda, bu zincire vurulmuş insanlar ile ateş arasından bir yol geçiyor. Bu yol boyunca, tıpkı kukla oynatanların marifetlerini göstermek için seyircilerin önüne yerleştirdiklerine benzer alçak bir duvar örülmüş.' (...) 'Şimdi, bu alçak duvarın arkasından geçen insanlar düşün. Duvarın üzerinden sadece taşıdıkları çeşit çeşit alet, taş ve tahtadan yapılmış insan ve hayvan heykelleri görülüyor. Doğal olarak bu insanların kimi konuşuyor, kimi susuyor.' (...) 'zincirlerinden kurtarılsalardı ve böylece kavrama bozuklukları da giderilseydi neler olurdu? (...) Utanıp sıkılır, daha önce gördüklerinin şimdi ona gösterilenlerden daha gerçek olduğunu düşünmez miydi? [11]'"*

Platon düşüncesinde bir şeyin gerçek özü, duyusal dünyada değil; idealar dünyasında bulunmaktadır. Platon'un duyusal dünya ile idealar dünyası arasında yaptığı ayrımın ve gerçek bilgiyi idealar ile özdeşleştirmesinin yapay zekaya olumsuz bir değer yüklediği söylenebilir. Neticede yapay zeka, kendisine yüklenen veri ve algoritmalar aracılığıyla insan zekasını ve davranışlarını kolayca taklit edebilmektedir. Platon perspektifinden yapay zekaya bakıldığında taklit aracılığıyla üretilen yapay zekanın yapacağı eylemlerin hiçbir gerçekliğe sahip olamayacağı düşünülmelidir. Çünkü duyusal dünya, idealar dünyasının kusurlu bir kopyasından ibarettir. Gerçek bilgi, kusursuz olan idealar dünyasında bulunduğundan yapay zekanın ürettiği bilginin duyusal dünyada meydana gelen verilerden öteye geçememesi beklenmelidir.

Duyusal dünya ve idealar dünyası arasındaki farkı açıklayan Platon, dikkatleri iyi ideasına çeker. Platon, mağara alegorisinin devamında iyi ideasından şu sözlerle bahseder: *"(...) benim görüşüme göre, iyi idea'sı, kavrananlar dünyasında en son olarak ve ancak büyük bir çabayla görülebilir. (...) gerçeği ve aklı bize, kavrananlar dünyasına hükmeden olarak bizzat o bahşeder [11]."* Platon'a göre iyi ideası, evrendeki düzeni sağlayan ilkedir. İdealar, duyusal dünyadaki şeylerin nedeni; iyi ideası ise ideaların nedenidir. Dolayısıyla iyi ideası aslında evrendeki bütün düzeni sağlayan ve her şeyin en yüksek nedeni olan ideadır. İyi ideasının bütün gerçekliğin en temel kaynağı olarak görülmesi, yapay zeka açısından "taklidin taklidi" meselesini gündeme getirmelidir. Bütün gerçekliğin en temel kaynağı iyi ideası olduğuna göre yapay zekanın insan taklit edilerek oluşturuluyor olması, aslında iyi ideasından gelen bir taklit sürecini içermelidir. Zaten kusurlu olan bir kopyanın daha da kusurlu hale getirilmesi, yapay zekanın değerinin daha da düşmesi anlamına gelmelidir.

Kusurlu olarak görülen duyusal dünya, insanın duyularıyla algıladığı her şeyin yanıltıcı olduğu düşüncesini beraberinde getirir. Bu yüzden epistemolojik anlamda gerçek bilgi, duyusal bir içeriğe sahip olmamalıdır. Platon bu konudaki düşüncesini bölünmüş çizgi analojisi ile açıklar:

> *" (...) eşit olmayan iki parçaya ayrılmış bir çizgi olarak düşün ve bu parçalardan her birini; görülenler alanını da, kavrananlar alanını da, tekrar aynı oranlarda ikiye böl. Sonra, görülenler alanındaki parçaları, netlik ve bulanıklık durumlarına göre göster. Ben bu resimlerde önce gölgeleri, ikinci olarak sudaki yansımaları ve her şeyden çok mat, düz ve parlak olanları ve benzer şeyleri görüyorum. (...) 'Şimdi de diğer parçayı; bu resme benzeyeni, yani çevremizde yaşayan hayvanları, bütün bitkileri ve bütün eşya türlerini al.' (...) 'Şimdi de, kavranan alanın nasıl bölünmesi gerektiğine bak.' (...) 'Şöyle: Ruh, belli varsayımlardan hareketle, bu bölümlerden birini aramalıdır ve o arada, demin sözünü ettiğimiz kopyalanan nesneleri görüntü olarak kullanmalıdır. Ama buradan hareketle başa dönmez, sonuca doğru ilerler. Diğer bölümü ararken de yine bir varsayımdan hareket eder ama, bu sefer ilk bölümde kullandığı görüntüleri kullanmaksızın, varsayımsız başlangıç yönünde gider, idea'nın yardımıyla, onun düzenlenmiş aşamalarını (yöntemleri) izler [11]."*

Platon'un bölünmüş çizgi analojisi ile vurgulamak istediği, gerçek bilgi olan ideaların bilgisine noesis ile yani saf akılsal bilme ile ulaşılabileceğidir. Gerçek bilgiye ulaşma sürecini hiyerarşik bir şekilde ele alan Platon, böylece insan zihninin somut düşünmeden soyut düşünmeye nasıl yükseleceğini de göstermektedir. Burada vurgulanması gereken, bahsi geçen yükselmeye yalnızca akıl yürütmeyle gerçeği kavrayabilenlerin ulaşabileceğidir. Böyle bir perspektifte şu an için yapay zekanın doxa'dan episteme'ye geçişe sahip olamadığı söylenmelidir. Bu durum yapay zekanın insanın gösterdiği ruhsal ve akılsal gelişimi gösteremediği anlamına gelmelidir. Böyle düşünen birisine göre bir gün yapay zekanın insanın yerini alabilmesi mümkün değildir.

Erdemli bir bireyin ve toplumun nasıl olması gerektiğini açıklamak isteyen Platon, ahlak ve siyaset arasında bir ilişki kurar. Bu ilişkiye göre bireyin sahip olacağı erdem toplumun erdemini, toplumun sahip olacağı erdem ise bireyin erdemini belirler. İnsan ruhunun üç parçalı bir yapıya sahip olduğunu belirten Platon'a göre ruhun akıllı kısmı, yürekli-atılgan kısmı ve iştah duyan kısmı vardır. Ruhun yürekli-atılgan ve iştah duyan kısmını dengede tutan

akıldır. Bu noktada Platon, ruhun her kısmına bir erdemin karşılık geldiğini belirtir. Ruhun akıllı kısmının erdemi bilgelik, yürekli-atılgan kısmının erdemi cesaret ve iştah duyan kısmının erdemi ölçülülüktür. Bu üç erdemin denge ile uygulanabilmesi ise adaleti ortaya çıkarır. Platon'a göre adalet, mutluluktur: *"(...) adil ruh ve adil adam iyi yaşar; adil olmayan ise kötü yaşar. (...) Ve de, iyi yaşayan mutlu ve övgüye layıktır; kötü yaşayansa tam tersine. (...) Yani, adil adam mutlu, adil olmayan adam mutsuzdur [11]."* O halde adalet, hem bireysel hem de toplumsal mutluluğu sağlayacak bir unsurdur. Bu noktada sorgulanması gereken, yapay zekanın toplumsal düzeni tehdit edecek şekilde kullanılma ihtimalidir. Böyle bir kötüye kullanımın sonucunda hem toplumsal düzenin bozulması hem de insanlar arasındaki adalet duygusunun sarsılması beklenendir. Bu durumda Platoncu anlamdaki mutluluk yok olur. Ayrıca Platon'un toplum anlayışında bireye yüklediği sorumluluk açıktır. Ancak özellikle sohbet robotlarının insanın karar verme anında kullanılabiliyor olmasının toplumsal yapıyı olumsuz etkileme ihtimali vardır. Böyle bir ihtimal ise Platon'un insanın kendi doğasıyla uyum içinde gelişmesi gerektiği fikrine uygun değildir. Dolayısıyla burada sorgulanan, yapay zekanın insanı kendi doğasından uzaklaştırarak bireysel ve toplumsal bir bozulmaya yol açma ihtimalidir. Yani yapay zekanın değerinin artması insan ruhunun değerini azaltabilir. Platon'a göre gerçek bilgiye ulaşmak için sürekli olarak düşünmek ve sorgulama yapmak gerekmektedir. Bu noktada yapay zeka, düşünülenlere ve sorgulananlara yeni bir boyut katarak daha önce fark edilmeyen düşünce ve sorgulamalara sevk edebilir.

Aristoteles Perspektifinden Yapay Zeka

Aristoteles idealar kuramından dolayı hocası Platon'u ciddi bir şekilde eleştirir. Bu eleştirinin temelinde Platon'un ideaları, duyusal dünyanın dışında görmesi vardır. Aristoteles'e göre idealar, duyusal dünyanın dışında olamaz. Eğer olursa bu gerçekliğin pratik dünyada hiçbir işe yaramadığı anlamına gelir. Böylece Aristoteles, gerçekliğin merkezini idealar dünyasından duyusal dünyaya çeker. Aristoteles, "Metafizik" eserini *"Tüm insanlar doğal olarak bilmeye iştah duyarlar. Bunun göstergesi de duyumlara düşkün olmamızdır (...) [12]"* sözü ile başlatır. Çünkü duyusal dünya, yaşamın temel parçasıdır. Bu parçayı çözmek için bilmek gerekmektedir ve Aristoteles'e göre bilmek, nedenlerin bilgisine sahip olmaktır. Buradan hareketle Aristoteles, dört neden öğretisini ortaya koyar.

Öğretide bahsi geçen dört neden; maddi neden, formel neden, fail neden ve ereksel nedendir. Maddi neden, içinde değişime ve başka bir şey olabilmeye dair bir özellik taşıyan ve oluşturucu olan nedendir. Formel neden, maddi nedenin getirdiği özelliği değişmeyen bir hale dönüştüren nedendir. Fail neden, bütün oluşumun kaynağı olan nedendir. Ereksel neden, bir şeyin yapılış amacını açıklayan nedendir. Aristoteles daha sonra fail ve ereksel nedenin formel nedene özdeş olduğunu belirterek sadece maddi ve formel nedenden bahseder. Çünkü fail ve ereksel nedenden bahsetmek aslında formel nedenden bahsetmektir. Bir varlığın en temel unsuru olarak görülen madde, varlığın gerçeklik kazanması için gereklidir. Aristoteles'e göre varlık, madde ve formdan ibarettir. Bu birliktelik, bir şeyin varlık olarak tanımlanmasını sağlar. Dolayısıyla her ne kadar madde varlığa gerçeklik sağlasa da varlığın şeklini veren formdur. Madde bir potansiyel olduğuna göre potansiyelin aktüelleşmesi form sayesindedir. O halde madde, bir forma sahip olmadıkça aktüel hale geçemez. Aristoteles düşüncesinde form, potansiyelden aktüel hale geçişin anahtarıdır. Bu durum herhangi bir şeyin varlık kazanması için hem madde hem de forma sahip olması

gerektiğini gösterir. Aristoteles bu düşüncesi ile her şeyin içinde değişime adapte olacak bir özellik taşıdığını ifade eder.

Aristoteles'e göre madde ve form birlikteliğinden muaf olan sadece Tanrı'dır. Çünkü Tanrı, saf formdan ibarettir. Ayrıca Tanrı, Aristoteles düşüncesinde ilk hareket ettiricidir. Bu çerçevede Aristoteles, evrendeki hareketlerin hareket etmeyen bir ilk hareket ettirici gerektirdiğini ifade eder. O da Tanrı'dır. Tanrı saf forma sahip olduğundan hiçbir şekilde madde içermemektedir. Burada tartışılması gereken, yapay zekanın insandan daha mı fazla saf forma yakınlaştırılmak istendiğidir. Bu tartışma ancak mükemmellik kavramı üzerinden yapılabilir. Sürekli öğrenme sürecinden geçirilen yapay zeka böylece karşısına çıkabilecek değişik durumlara hazırlanmaktadır. Çünkü yapay zeka üreticilerinin hayalleri arasında yapay zekayı mükemmel bir hale getirmek vardır. Bu nokta, bir gün yapay zekanın insandan daha mı fazla Tanrı'ya yaklaşmış olacağının sorgulanması gerektiğini beraberinde getirir. O halde yapay zekanın değerini Aristoteles perspektifinden sorgulayanlar için yapay zekanın bir gün insandan daha gelişmiş olarak görüleceği bir konuma gelmesi ile Tanrı'ya yapay zekanın mı yoksa insanın mı daha yakın olacağı bir problem olarak doğmalıdır.

Yapay zekanın sahip olduğu algoritma ve verilerin insan aklı ve davranışları örnek alınarak hazırlanıyor oluşu, Aristoteles perspektifinden yapay zekaya bakanlar için olumlu bir değer olabilir. Varlığa bakışın duyusal dünyaya çekilmiş olması bu değerin ilk adımıdır. Bu değer, Aristoteles'in bilginin oluşma sürecinde duyular ve aklın beraber hareket etmesi gerektiğini söylemesi ile pekişir. Aristoteles, duyuların her türlü bilginin kaynağı olduğunu kabul etmekle beraber bilginin oluşma sürecinde tek başına yeterli olmadığını belirtir. Çünkü duyu, sadece tikeli konu aldığından bilginin oluşma sürecinde yetersiz kalmaktadır. Bireysel duyumlardan gelen ortaklıklar çoğaldıkça tümel ortaya çıkar ve bu deneyin ortaya çıktığı aşamadır. Deney aşamasında tek tekler gözlemlenerek tümelin bilgisine ulaşılır. Ancak deney aşaması da bilginin oluşma süreci için yetersizdir ve bir üst aşamaya geçilmesi gerekir. Bilginin oluşma sürecinde elde edilebilecek en yüksek aşama tümel kavrama aşamasıdır. Bu aşamada tümel, kendinde olduğu gibi kavranır. Burada unutulmaması gereken tümelin bilgisine ulaşmak için tek olandan yola çıkıldığıdır. Dolayısıyla Aristoteles'in bilgi hiyerarşisi, duyular ve aklın beraber hareket etmesi gerektiğini göstermektedir. Aristoteles'in tümel kavrama anlayışı, yapay zeka yazılımcılarının kullandığı derin öğrenme ve makine öğrenmesi gibi yöntemler açısından önemli görülebilir. Çünkü, bu tür teknikler verilerden genelleme yapar. Yani tümel kavrama yeteneğine sahiptir. Aristoteles'in bilgi edinmek için duyu ve aklın birlikteliğini savunması da yapay zeka yazılımcıları için önemli bir diğer nokta olabilir. Böyle bir bakış açısı, hazırlanan yazılımın sadece veriye dayalı kalmaması gerektiğini; bu veriyi anlamlandıran bir akıl yürütme mekanizmasının da gerekli olduğunu göstermektedir. Aristoteles'in bu düşüncesi, yapay zekanın bir gün tamamen insan gibi olabilmesini sağlamak için bir katkı sunabilir.

Aristoteles'in gerçekliğin kaynağını duyusal dünyaya çekmiş olması, insan aklı ve davranışlarının öne çıkmasını beraberinde getirir. Aristoteles felsefesinde bunun en açık örneği, tümel kavramanın düşünce erdemleri ile kesiştiği noktadır. Aristoteles felsefesinde sahip olunan bilgiyi etkili bir şekilde kullanabilmenin yolu, düşünce erdemlerine sahip olabilmekten geçer: *"Aristoteles düşünce erdemlerini zanaat (Teckne), nous, aklı başındalık (pronoesis), bilim (episteme) ve bilgelik (sofia) olmak üzere beş başlıkta inceler. Bu erdemler, teorik*

ve pratik yaşam arasında destekleyici bir köprü kurulmasına olanak sağlar [13]." O halde düşünce erdemlerine sahip olmak, akıl ve düşünme sürecini güçlendirecektir. Aristoteles felsefesinde sadece teorik bilgiye sahip olmak yetmez, sahip olunan bilginin pratik hayata doğru bir şekilde yansımış olması gerekir. Bu yansıma olmadan iyi bir yaşam veya doğru bir davranıştan söz edebilmek mümkün değildir. Dolayısıyla gündelik hayatta doğru eyleyebilmek için gelişen insan aklı ve davranışları yapay zeka için iyi bir örnek olabilir.

Aristoteles'in teorik ve pratiğin birlikteliğini vurgulayan felsefesi, bilgiye dayalı eylemlere verdiği değeri gözler önüne serer. Dolayısıyla böyle kodlanan bir yapay zeka sağlık, eğitim, üretim gibi hayatın pek çok alanında insana yardımcı olabilir. Böyle bir perspektife göre insanın danışacağı yapay zeka, insanın doğru kararlar almasına yardımcı olarak hayatını kolaylaştırmalıdır. Böylece insanlar kendilerini geliştirdikleri şekilde eğitim alabildikleri, sağlıklı ve planlı yaşayabildikleri bir yaşama kavuşabilecektir. Bu bireysel gelişim, ilerleyen zamanda toplumsal bir faydaya dönüşme ihtimalini içinde taşır. Çünkü gelişen bir bireyin toplum için etkili hizmet verme potansiyeli yüksektir. Bu noktada yapay zekanın insanlık için önemini, MIT Bilgisayar Bilimleri laboratuvar yöneticilerinden Edward Fredkin "Tarihte üç büyük olay vardır. Bunlardan ilki kainatın oluşumudur. İkincisi yaşamın başlangıcının olmasıdır. Üçüncüsü de yapay zekanın ortaya çıkışıdır" sözüyle açıklar [3]. Fredkin'in de vurguladığı gibi, yapay zekanın ortaya çıkışı insanlık tarihi için önemli bir adımdır. Dolayısıyla Aristoteles perspektifinden yapay zekanın değerine bakan birisi için yapay zekanın sağladığı verimlilik, zaman ve enerji tasarrufu gibi unsurlar yapay zekayı önemli ve değerli kılmalıdır.

Aristoteles, "Nikomakhos'a Etik"in ikinci kitabında duygular ve eylemler için altın orta ilkesini belirler: *"(...) eşitlik, aşırılıkla yetersizlik arasında herhangi bir ortadır. Nesnede ortayla, iki ucun her birinden eşit uzaklıkta olan şeyi, bütün insanlar için tek ve aynı olan bir noktayı anlıyorum ve bizimle ilişkili olan şeyde ortadan, ne çok ne de az olan şeyi anlıyorum (...) [14]"* Dolayısıyla Aristoteles'e göre bireysel ve toplumsal ahlak için denge önemlidir. Ancak kontrolsüz bir şekilde gelişen yapay zeka, bu dengeyi ortadan kaldırarak gizlilik ve güvenlik konularında pek çok sorunun yaşanmasına sebep olabilir. 2002 yılında deney merkezinden kaçan robot Gaak, bu duruma bir örnektir. Magna Deney Merkezi'nin yürüttüğü bilimsel bir araştırmanın parçası olan Gaak, gÖZETimsiz bırakıldığı an bulunduğu bölümden çıkmış, aşağı inmiş, otoparka kadar kaçmayı başarmış hatta bir otomobilin altında kalma tehlikesi bile yaşamıştır [15]. Bu örnek ile vurgulanmak istenen, yapay bir aklın bir gün insanı tehlikeye sokup sokmayacağıdır. Aristoteles perspektifinden yapay zekanın değerini sorgulayan birisi burada mutluluk kavramından hareket etmelidir. Aristoteles'e göre mutluluk haz yaşamı, politika yaşamı, zenginlik yaşamı ve teorik yaşamda bulunabilir. Ancak bunların en doğrusu teorik yaşam ile sağlanan mutluluktur. Aristoteles'in mutluluk kavramı öne çıkarılarak sorgulanan, yapay zekanın kullanımının artmasıyla insanların teorik yaşamdan yani mutluluktan uzaklaşma ihtimali olmalıdır.

Sonuç

Yapay zeka, insan benzeri davranış ve düşünme sergileyen sistemlerin genel adıdır. Bu çerçevede yapay bir zekanın insan zekasını ve davranışlarını ne kadar taklit edebileceği merak konusu olmuştur. Bu merak, kendisini ilk olarak mitlerle ortaya çıkarmıştır. Talos, Hephaistos, Galatea, Prometheus ve Pandora'nın ana kahraman olduğu mitler incelendiğinde

yapay zeka sorgulamalarına tanık olunabileceği söylenebilir. Bu sorgulamaların temelinde yapay zekanın bir gün insan gibi akla ve davranışa sahip olup olmayacağı, yapay zekanın insanların yerini alıp almayacağı ve yapay zekanın bir gün kontrolsüz güce sahip olup insanlara zarar verip vermeyeceği gibi problemler vardır. 13 ve 19. yüzyıl arası ise hem insana benzeyen makinelerin üretiminin başladığı hem de konuyla ilgili eserlerin popülerleştiği yüzyıllardır. Üretimlerin başlangıcı ile eserlerin popülerleşmesinin paralelliği, insana benzeyen makinelerin üretiminin her çağda insanlar için bir karşılık bulduğunu göstermektedir. 20. yüzyılın başları ise insana benzeyen makinelerin üretiminin hızlandığı ve bu makinelere duyulan ilginin arttığı dönemdir. Bu dönem hem insan benzeri iş gücü üretmeye dair birçok yeniliği hem de böyle bir algoritmanın teorik temellerini gösterebilmek yapılan pek çok çalışmayı beraberinde getirmiştir. Ayrıca bu dönemde Karel Čapek tarafından ilk defa "robot" kelimesi; John McCarthy tarafından ise ilk defa "yapay zeka" kelimesi kullanılmıştır. Ancak 20. yüzyılın sonlarına gelindiğinde "Yapay Zeka Kışı" yaşanmış ve yapay zeka alanındaki gelişmeler durağanlaşmıştır. Durağanlaşmanın sebepleri arasında hızla gelişen teknolojinin pratikte uygulanamaması, gerçekçi olmayan hedeflerin fark edilmesi ve ilgi kaybına bağlı olarak araştırmaların azalması vardır. 21. yüzyıla gelindiğinde ise bir yandan yapay zeka alanındaki gelişmeler; diğer yandan da yapay zeka kullanımının etik ve güven zemininde tartışılması devam etmektedir.

Tarihsel gelişimi boyunca yapay zekanın kimileri tarafından olumlu görülürken kimileri tarafından da olumsuz görüldüğü açıktır. Olumlu görenlere göre yapay zekanın yaşama katılmasıyla pek çok konuda gelişim sağlanabilir. Bu gelişim başlıca verimlilik, tasarruf, öneri ve insan hatalarının azalması konularını içermektedir. Olumsuz görenlere göre ise yapay zekanın yaşama katılmasıyla işsizlik korkusu, etik sorunlar, insan zihninin pasifleşmesi, gizlilik ve güvenlik sorunları gündeme gelmiştir. Görüldüğü üzere yapay zekanın yaşamı kolaylaştıran olumlu yönlere sahip olduğu söylendiği gibi önlem alınmasını gereken olumsuz yönlere sahip olduğu da söylenmektedir. Yapay zekaya olumlu ve olumsuz bakış, Platon ve Aristoteles'in perspektiflerinden de ele alınabilir. Bu bağlamda yürütülen bu sorgulama, Platon ve Aristoteles perspektiflerinden hareketle yapay zekanın bir değere sahip olup olamayacağını tartışmaktadır.

Platon'a göre gerçeklik duyusal dünyada değil, idealar dünyasında bulunmaktadır. Duyusal dünya, sadece idealar dünyasının kusurlu bir kopyasıdır. Duyusal dünya ve idealar dünyası arasında yapılan bu ayrımın yapay zekaya olumsuz bir değer yüklediği söylenebilir. Çünkü yapay zeka, duyusal dünyaya ait olan insanı taklit etmektedir. Bu bakış açısına göre duyusal dünyanın bir kopya olması, duyusal dünyayı taklitle oluşturulmuş bir yapay zekanın ürettiği bilginin de gerçekliğe ait olamayacağını beraberinde getirecektir. Ayrıca Platon felsefesinde evrendeki düzeni iyi ideası sağlar. O halde iyi ideası, her şeyin en yüksek nedenidir. İyi ideasının en temel kaynak olarak görülmesi, yapay zekanın değerinin düşmesi anlamına gelebilir. Gerçek bilgiye ulaşma sürecini hiyerarşik bir düzenle şekillendiren Platon'a göre, bu düzen doxa'dan episteme'ye geçişin bir sembolüdür. Bu yüzden Platon perspektifinden yapay zekaya bakan birisi için yapay zekanın insanın yerini alma ihtimali düşüktür. Bu bakış açısına göre episteme'ye geçiş, yapay zekanın göstereceği ruhsal ve akılsal gelişime uygun olmayabilir. Platon felsefesinde adalet kavramı hem bireysel hem de toplumsal mutluluk sağlayan bir kavramdır. Yapay zekanın kötüye kullanımı, toplumsal düzenin bozulmasını ve adalet duygusunun zedelenmesini beraberinde getirebilir. Böyle bir

durum, Platoncu anlamdaki mutluluğun bozulması demektir. Ayrıca bazı yapay zeka araçlarının insanlar gibi karar verebiliyor olmasının toplumsal yapıyı bozma ihtimali vardır. Böyle bir ihtimal, Platon perspektifinden yapay zekaya bakan birisi için uygun görünmemektedir. Çünkü bu durumda, yapay zeka insanı kendi doğasından uzaklaştıracaktır. Platon'a göre gerçek bilgiye ulaşabilmek için düşünme ve sorgulama sürecinin sürekli olması gerekmektedir. Yapay zekaya dair tartışmalar, Platon'un istediği bu sürekliliği sağlayarak düşünme ve sorgulama sürecine yeni boyutlar katmaktadır.

Aristoteles'e göre gerçeklik, pratik dünyada işe yarar olmalıdır. Bu yüzden gerçekliğin merkezi idealar dünyasında değil, duyusal dünyadadır. Bunun ispatını insanların bilme isteğinde gören Aristoteles'e göre bilmek, nedenlerin bilgisine sahip olmaktır. Aristoteles'in bu düşüncesi dört neden öğretisini ve potansiyel-aktüel ayrımını beraberinde getirir. Bu bağlamda söylenmesi gereken, maddenin bir forma sahip olmadıkça aktüel hale geçemeyeceğidir. O halde her şey içinde değişime adapte olacak bir özellik taşımaktadır. Aristoteles felsefesinde saf formdan ibaret olan sadece Tanrı'dır. Bu noktada Aristoteles perspektifinden yapay zekaya bakıldığında yapay zeka üreticilerinin yapay zekayı mükemmel hale getirme çabalarının olduğu hatırlanmalıdır. Yapay zekayı mükemmel hale getirmeye çalışmak, yapay zekanın insandan daha mı fazla Tanrı'ya yaklaşmış olabileceği ihtimalini içinde taşır. O halde Aristoteles perspektifinden yapay zekaya bakmak, yapay zekanın insandan daha fazla saf forma yakınlaştırılmak istenip istenmediği sorgulamasını ortaya çıkarabilir. Aristoteles'e göre bilgi, duyu ve aklın birlikteliğinden doğmaktadır. Aristoteles perspektifinden yapay zekaya bakan birisi için yapay zeka için hazırlanan yazılım sadece veriye dayalı kalmamalı; aynı zamanda bu veriyi anlamlandıran akıl yürütme mekanizmasına da sahip olmalıdır. Böyle bir düşünce, yapay zekanın insan gibi olabilmesini sağlamak için bir katkı olabilir. Aristoteles'in tümel kavrama hakkındaki fikirleri de yapay zeka için bir katkı sunabilir. Çünkü genellemeler yapmayı sağlayan tümel kavrama, yapay zeka yazılımcıları için bir yöntem olabilir. Ayrıca Aristoteles felsefesinde tümel kavramanın düşünce erdemleri ile beraber ele alınıyor olması, teorik ve pratiğin beraber önemli olması demektir. O halde Aristoteles felsefesinde varlığa bakışın duyusal dünyaya çekilmiş olması, yapay zeka için olumlu bir değer olabilir. Çünkü yapay zekanın insanın aklını ve davranışlarını taklit ediyor olması, duyusal dünyadaki insanın kendini iyi olana sevk etme gerekliliğini beraberinde getirecektir. Böyle kodlanan bir yapay zekanın yaşamda insana eğitim, sağlık, üretim gibi birçok konuda yardımcı olması beklenmelidir. Yapay zeka sayesinde bireysel olarak gelişen insanın toplum içindeki potansiyelinin artma ihtimali de yüksektir. Aristoteles felsefesinde hem bireysel hem de toplumsal ahlak için altın orta ilkesi bir denge sunmaktadır. Kontrolsüz şekilde geliştirilen bir yapay zeka, bu dengeyi bozarak sorunlara yol açma ihtimalini taşımaktadır. Bu ihtimal beraberinde yapay zekanın bir gün insanı tehlikeye sokabileceği endişesini getirmektedir. Bu noktada Aristoteles'in mutluluk kavramı öne çıkmalı ve yapay zekanın kullanımının artmasıyla insanların teorik yaşamdan uzaklaşma ihtimali sorgulanmalıdır.

Platon ve Aristoteles'in düşünceleri, yapay zeka alanındaki tartışmalarla ilişkilendirilebilir. Böylece sorgulanan, Platon'un idealizme ve Aristoteles'in realizme dayanan felsefelerinden hareketle yapay zekanın bir değere sahip olup olamayacağıdır. Yapay zeka geliştikçe insanın düşünme ve sorgulamasına katkıda bulunmaya devam edecektir. Böylece insanlık ve yapay zeka arasındaki bağı anlama, insanın kendini tanıması ve teknolojiyi nasıl kullanmak gerektiği gibi meselelerin gündemde kalması beklenmektedir.

REFERANSLAR

[1] URL1: https://sozluk.gov.tr/ (Erişim Tarihi: 16.03.2025)

[2] URL2: https://sozluk.gov.tr/ (Erişim Tarihi: 16.03.2025)

[3] H. Pirim, "Yapay Zeka", Journal of Yaşar University, 1, 1, s. 81-93, 2006.

[4] R. C. Schank, "Where's the Ai?", AI Magazine, 12, 4, s. 38-49, 1991.

[5] G. Topal, and N. İspir, "Kadim Tanrılardan Modern Tanrılara: 21. Yüzyıla Dair Mitolojik Bir Okuma", Kültür Araştırmaları Dergisi, 12, s. 232-252, 2022.

[6] R. Descartes, "Metot Üzerine Konuşma (Çev. Atakan Altınörs)", Bilge Kültür Sanat Yayıncılık, İstanbul, 2017.

[7] J. V. Atanasoff, "Computing Machine for the Solution of Large Systems of Linear Algebraic Equations", Iowa State University-Library, s.1-31, 1940.

[8] A.M. Turing, "Computing Machinery and Intelligence", Oxford University Press on behalf of the Mind Association, 59, s. 433-460, 1950.

[9] J. McCarthy, M. L. Minsky, N. Rochester, and C. E. Shannon, "A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence", s. 1-13, 1955.

[10] URL3: https://www.grandviewresearch.com/industry-analysis/artificial-intelligence-ai-market (Erişim Tarihi: 14.03.2025)

[11] Platon, "Devlet (Çev. Işık Soner)", Kaynak Yayınları, İstanbul, 2013.

[12] Aristoteles, "Metafizik (çev. Y. Gurur Sev)", Pinhan Yayıncılık, İstanbul, 2017.

[13] M. A. Sarı, and S. Güleroğlu, "Aristoteles'in Bilgi Anlayışı: Episteme, Mantık ve Düşünce Erdemleri İlişkisi", Tabula Rasa Felsefe & Teoloji Dergisi, 43, s. 210-223, 2024.

[14] Aristoteles, "Nikomakhos'a Etik (Çev. Zeki Özcan)", Sentez Yayıncılık, Ankara, 2014.

[15]URL4: https://stockholmcentre.org/escape-of-the-gaak-new-technologies-and-the-ethics-of-war/ (Erişim Tarihi: 16.03. 2025)

PREDICTION OF WATER LEVELS IN DAMS AND RESERVOIRS FOR ISTANBUL AND ANKARA WITH ANFIS MODEL

Sevcan Emek*
Manisa Celal Bayar University, Dept. of Computer Engineering, Manisa, Türkiye

Kübra Çelebi
Manisa Celal Bayar University, Dep.t of Computer Engineering, Manisa, Türkiye

Oğuz Aydın
Manisa Celal Bayar University, Dept .of Computer Engineering, Manisa, Türkiye
**(Corresponding author) sevcan.emek@cbu.edu.tr*

ABSTRACT

Climatic, geographical, social and environmental factors directly affect the resources needed in real life. For this reason, management models are needed for usable water resources that are of vital importance. Dams and reservoirs are important structures that meet the water needs of the living population. Droughts that develop depending on climatic and environmental variables affect the occupancy rates in dams and reservoirs. Forecasting methods are used to plan water consumption and energy management accurately and realistically. This work focuses on the prediction of water levels in the dams of Ankara and the reservoirs of Istanbul, considering climatic factors such as temperature, rainfall, windspeed, snowfall, and historical hydrological data. This study aims to provide accurate predictions to support water resource management and ensure the sustainable operation of water systems. Prediction is carried out with Adaptive Neuro-Fuzzy Inference (ANFIS) model approach. The system's effectiveness has been tested using real-world data which obtained ISKI (Istanbul Water and Sewerage Administration) and ASKI (Ankara Water and Sewerage Administration). In this work, four error metrics are used: MAE, RMSE, NMRSE, and R^2 to evaluate ANFIS model's performance. The error results are obtained for MAE, RMSE, NRMSE, and R^2 of 0.1936, 0.4061, 0.0086 and 0.9992, respectively for Ankara. Test errors for İstanbul are MAE, RMSE, NRMSE, and R^2 of 0.1502, 0.1625, 0.0057 and 0.9996, respectively. These results indicate a high level of accuracy for both cities, while indicating that the Ankara model provides more consistent predictions with lower error rates.

Keywords: ANFIS, dam water levels, reservoir occupancy prediction, water resource management, water sustainability.

Introduction

Accurate prediction of water levels in both dams and reservoirs is crucial for effective water resource management, especially in regions prone to drought and water scarcity. Water is a vital resource for human life and sustainable development, but the combined effects of rapid urbanization, population growth, and climate change have intensified water scarcity issues worldwide. As the country's population increases, the amount of usable water per person is also decreasing. Water consumption is especially high in densely populated megacities. According to TUIK (Turkish Statistical Institute), Istanbul ranks first with 18.33% of the country's population, followed by Ankara with 6.85% [1]. The amount of water in dams and reservoirs is directly related to the amount of rainfall. According to the Turkish State Meteorological Service, the 2025 water year rainfall decreased by 29% compared to normal in the five-month period and by 34% compared to the same period last year [2]. According to the ISKI (Istanbul Water and Sewerage Administration), the annual rainfall measured in Istanbul dams is quite low compared to last year [3]. The annual rainfall measurements for Istanbul are shown in Figure 1.

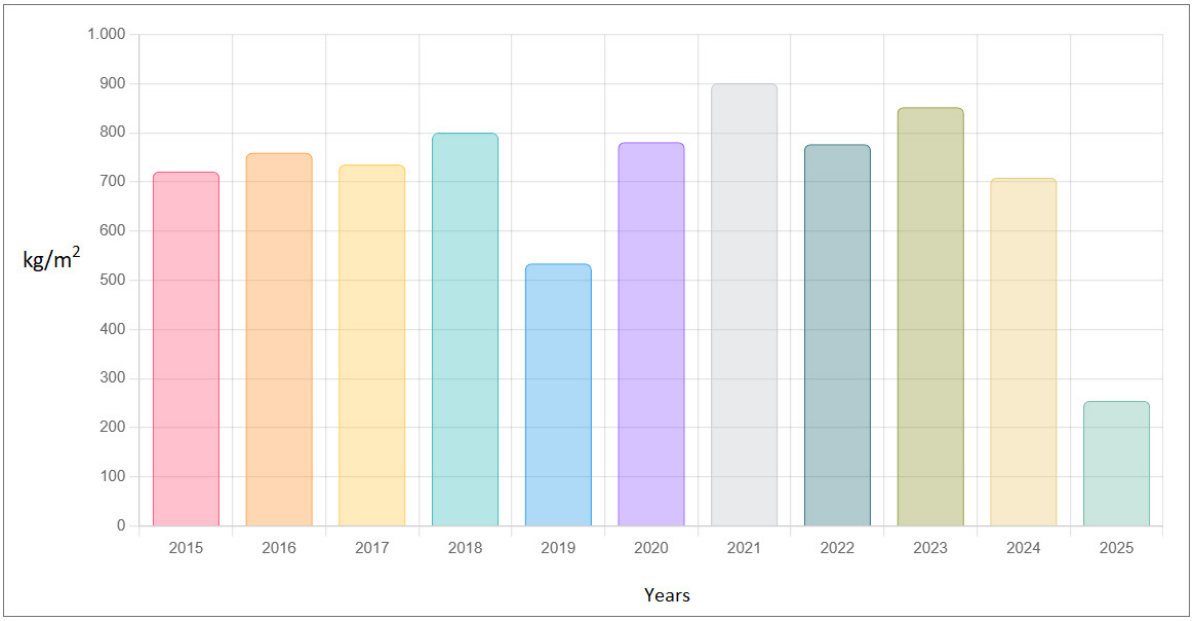


Figure 1: Annual Rainfall Graphic for Istanbul

Istanbul, as one of the largest and most densely populated cities in Europe, faces significant challenges in managing its water resources effectively. The city's reservoirs play a critical role in meeting the growing water demand, making accurate prediction and management of reservoir levels essential for maintaining water security. The other megacity, Ankara, is one of the provinces with high water consumption. According to ASKİ (Ankara Water and Sewerage Administration) data, the total occupancy rates of dams in Ankara have decreased significantly compared to last year [4]. While the total occupancy rate was 48.46% in 2024, it was determined to be 26.27% in 2025. The amount of water coming to dams by year is shown in Figure 2.

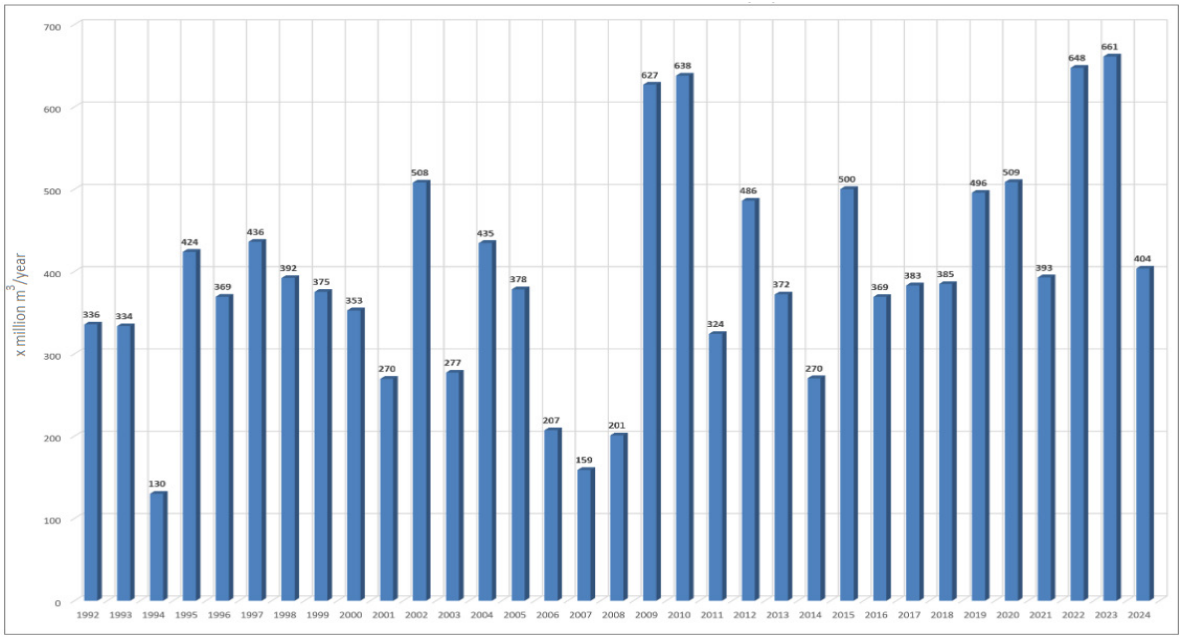


Figure 2: Annual Occupancy Measurements Graphic for Ankara

Traditional hydrological models often struggle to handle the nonlinearities and uncertainties inherent in hydrological processes. To address these limitations, artificial intelligence (AI) and machine learning approaches, such as Fuzzy logic and ANFIS, have emerged as powerful tools. These methods are particularly effective in modeling complex, uncertain systems and providing reliable predictions based on limited or imprecise data. In this study, the occupancy rates of dams and reservoirs for Ankara and Istanbul regions are estimated through the ANFIS model. ANFIS is a type of artificial neural network method [5]. This method creates an input-output structure based on input-output data pairs. ANFIS executes a hybrid learning algorithm as it integrates both neural networks and fuzzy logic inference methods. This study incorporates environmental factors such as rainfall, temperature, inflow, and outflow, as well as historical hydrological data. MATLAB's ANFIS editor and functions are utilized to load the training and testing datasets and train the model iteratively to minimize the prediction error.

The main contributions of this work define as follows:

- Developing a Regional Forecast Model: Separate ANFIS models are created for two metropolitan cities with different geographical and climatic conditions, namely Istanbul and Ankara, and the effect of regional differences on the model is evaluated. This situation constitutes a guiding example for local governments to use data-driven models in regional decision-making processes.

- Improving Model Performance with Feature Engineering: The predictive power of the model is increased by deriving new variables such as average temperature, temperature difference and previous day occupancy rate; thus, complex relationships are modeled more effectively.

- Data Transformation Processes and Harmonization: A homogeneous data set is created by applying interpolation and format conversion processes to combine daily and

monthly data. This approach constitutes an example for the integration of data with different time resolutions.

The rest of this study is organized as follows: Section 2 offers a literature review of related works. Section 3 provides detailed information about the implementation of this study. Section 4 evaluates the findings and outcomes of the model. Finally, the conclusions of this study are discussed in Section 5.

Literature Review

Water reservoir management plays a critical role in maintaining the water supply balance, especially in urban areas like Ankara and Istanbul, where population density and water demand are high. The ability to predict reservoir levels accurately enables authorities to plan water resource allocation and prepare for potential shortages. In recent years, machine learning and fuzzy logic- based models have gained attention for their ability to handle the uncertainties and complexities inherent in reservoir level prediction. Among these models, ANFIS have shown promising results by combining the strengths of neural networks and fuzzy logic.

Several studies in the literature have demonstrated the effectiveness of ANFIS and regression models in predicting water levels. These models typically incorporate data such as historical rainfall, temperature, and evaporation rates to generate forecasts. The findings from these studies highlight the importance of selecting appropriate model parameters and input features to achieve accurate predictions. Badem et al. focus on predicting dam occupancy levels to contribute to sustainable water management. They combine physical models of water inflow and outflow with data-driven models based on historical reservoir data. Various prediction models, including Random Forest, Extra Trees, Long Short-Term Memory (LSTM), Orthogonal Matching Pursuit CV, and Lasso Lars CV, are implemented. The research demonstrates the impact of physical model parameters on prediction accuracy and identified the best method for predicting dam occupancy levels by comparing it with findings from the scientific literature [6]. Ziggah and Issaka integrate the genetic optimization algorithm (GA) with two artificial intelligence methods, including an emotional neural network (ENN) and ANFIS, to predict piezometric water levels for three different piezometers located on a dam. The experimental results indicate that the GA-ENN and GA-ANFIS models demonstrated no absolute superiority between them but are superior to ANN, GA-ANN, and ANFIS in terms of prediction accuracy [7]. Gül et al. investigate the applicability of fuzzy logic methods, including ANFIS and Fuzzy-SMRGT, for estimating evaporation amounts at the Atatürk Dam. The study uses parameters such as daily average relative humidity, air temperature, wind speed, and sunshine duration to estimate evaporation. The results show that both ANFIS and Fuzzy-SMRGT models provide accurate evaporation estimates, with performance metrics comparable to those of classical methods such as Multiple Linear Regression (MLR) [8]. Nazarpoor et al. present the impact of the Rudbal Dam on the Darab Aquifer in southern Iran using an innovative fuzzy logic-based approach. This study employs the Mann-Kendall test and its advanced version, Trend-Free Pre Whitening, to detect trends in time series data, while a proposed fuzzy inference system (FIS) converts inputs into fuzzy sets using trapezoidal and triangu-

lar membership functions. The study concludes that fuzzy logic effectively addresses uncertainties and complexities, offering valuable insights for future research and broader application in similar contexts [9]. Fu et al. conduct a study where potential dam sites are identified using remote sensing and GIS. The study concludes that the fuzzy logic approach, couple with OLS regression, is a powerful tool for deciphering potential dam sites and can be applied at regional and continental scales [10]. Er et al. investigate the applicability of Fuzzy-SMRGT, an advanced fuzzy logic approach, for estimating monthly reservoir level changes. The research focus on the Alibey Dam in Istanbul, utilizing historical reservoir level data from 1989 to 2020 obtained from the Turkish State Hydraulic Works (DSI). To estimate the monthly reservoir levels, the study considers four main hydrological parameters; evaporation, water inflow into the dam, drinking water consumption, and discharge from the reservoir. The study concludes that Fuzzy-SMRGT is a viable approach for predicting reservoir levels, contributing to sustainable water management efforts [11]. Páliz Larrea et al. conduct a study using neural network models and ANFIS to forecast the water level of the Salve Faccha reservoir, which supplies water to Quito, the Capital of Ecuador. The study concludes that precipitation is not an influential input in the prediction of the reservoir water level [12]. Rankovic´ et al. develop a feedforward neural network (FNN) model to predict the piezometric water level in dams. The study uses an improved resilient propagation algorithm to train the FNN and compared the results with multiple linear regression (MLR) models. The results show that FNN models could be a powerful tool for dam health monitoring, providing better prediction accuracy than MLR models. This study demonstrates the potential of artificial intelligence techniques in modeling complex nonlinear systems in dam engineering [13]. Unes et al. investigate the use of ANFIS to estimate reservoir volume fluctuations. By utilizing input variables such as precipitation, evaporation, dam discharge volume, and irrigation water release, the study demonstrates that ANFIS outperformed traditional multi-linear regression models in predicting reservoir volumes. The ANFIS model provides low mean square error and high correlation coefficients, indicating its effectiveness in handling the complexities of dam reservoir management [14]. Bagis and Karaboga propose a control method based on fuzzy logic for the real-time operation of spillway gates during floods. By combining Artificial Neural Networks (ANNs) with fuzzy logic, the study aims to model the non-linear relationships among reservoir variables and optimize gate operations. The proposed method shows superior performance over conventional control techniques, effectively managing flood conditions and maintaining reservoir levels within desired ranges [15]. F. Chang and Y. Chang utilize ANFIS to predict water levels in reservoirs during flood periods. The study highlights the capability of ANFIS to integrate human knowledge, such as current reservoir outflow, to enhance prediction accuracy. The research demonstrates the effectiveness of ANFIS in providing accurate and reliable water level forecasts for the Shihmen reservoir in Taiwan, particularly during typhoon events. The study confirms that including human decision inputs significantly improves the model's performance [16]. Tayfur and Singh compare Artificial Neural Network (ANN), Fuzzy Logic (FL), and Kinematic Wave Approximation (KWA) models for predicting rainfall-runoff processes. This study suggests integrating data-driven (ANN, FL) and deterministic (KWA) approaches for robust rainfall-runoff modeling and water resource management [17]. Mai et al. introduce a novel approach for predicting horizontal displacement in hydropower dams by combining Interval Type-2

Fuzzy Logic Systems (IT2FLS) with Ant Colony Optimization (ACO). Using data from the Ialy Hydropower Dam in Vietnam (2004–2016), the model incorporates inputs like reservoir water levels, temperature, and dam age, which significantly impact displacement [18]. The other study focuses on developing a model using fuzzy logic and ANFIS to predict water levels in Terkos Dam, Istanbul. This study highlights the importance of precise data collection and accounting for climate change in water resource predictions. It offers an innovative approach to water management and supports accurate forecasting for sustainable planning in densely populated areas like Istanbul [19]. Moorthi et al. present the use of fuzzy logic in managing uncertainties in water resource systems, focusing on Amaravathi Dam as a case study. This study highlights fuzzy logic's flexibility in addressing uncertainties and enabling participatory decision-making in water management [20]. Zhu et al. develop a fuzzy cluster logic model combined with adaptive differential evolution (ADE) to assess the risk of loss of life due to dam failures, which are rare but catastrophic events. This study offers a scientific and practical approach to improving dam safety risk management [21]. Karaboğa et al. introduce a fuzzy logic control method optimized with the tabu search algorithm for managing dam reservoir gates during floods [22]. The other study focuses on monitoring the water levels of dams and lakes that provide water to Istanbul. In this work, water levels, precipitation amounts, and occupancy rates are analyzed by correlating them between 2004-2014 [23]. Üneş et al. apply various artificial intelligence and statistical methods to predict river flow using data from the Stillwater River, Massachusetts, between 2014-2017. The models used included Multiple Linear Regression (MLR), ANN, M5 Decision Tree (M5T), ANFIS, Mamdani Fuzzy Logic (M-FL), and Simple Membership Functions with Fuzzy Rule Generation Technique (SMRGT) [24]. Ahmad et al. use Geographic Information Systems (GIS) and Remote Sensing (RS) techniques to determine suitable dam sites. This methodology analyzes the effects of various factors (geology, land use, slope, precipitation, soil texture, etc.) used in determining the dam locations, allowing the selection of the most suitable locations [25].

The insights gained from this study could contribute to future research and practical implementations for sustainable water management strategies. Fuzzy logic has a potential to handle uncertainty and approximate reasoning, making it suitable for environmental prediction tasks.

Implementation

This study utilizes two key techniques that significantly enhance the prediction of complex hydrological systems: Fuzzy logic and ANFIS. Fuzzy logic is a rule-based system that mimics human reasoning to process imprecise data. It uses linguistic variables and membership functions to model complex environmental relationships, making it particularly suitable for hydrological modeling where exact relationships are difficult to define. ANFIS is a hybrid machine learning approach combining neural networks with fuzzy logic to optimize membership functions and improve prediction accuracy. ANFIS adapts to changing data patterns through training, providing a more accurate prediction in dynamic and uncertain environments.

Advances in artificial intelligence and data-driven modeling have significantly enhanced the ability to predict such complex systems. Fuzzy Logic and ANFIS have emerged as powerful tools for addressing the uncertainties and nonlinearities inherent in reservoir management. In this work, these techniques are implemented using MATLAB, leveraging its Fuzzy Logic Toolbox and ANFIS functionalities for defining membership functions, rules, and evaluating system performance. The integration of these AI-driven technologies within a computational framework highlights their potential in addressing challenges in urban water management.

Data has been obtained from the ASKI, ISKI, DSI and Weather Underground websites, with MATLAB being utilized for data processing, model training, and validation. The tools enable the efficient implementation of advanced algorithms while providing visualization options to better understand the relationships between input variables and predicted outcomes.

The main parameters used in this study are: monthly rainfall (mm), temperature rates (mm), humidity rates (mm), historical rainfall rates (historical reservoir occupancy rates). The obtained data has been preprocessed to handle missing values and outliers. Normalization has been applied to bring all variables to a uniform scale, ensuring compatibility with the ANFIS model. The input parameters for ANFIS model training are: rainfall (mm), temperature (max and min ^{0}C), wind speed (km/h), snowfall (cm), humidity (g/m^{3}). Algorithm steps for ANFIS model training are shown in Algorithm 1.

Algorithm 1. ANFIS Model Training

(i) Initialize input variables
(ii) Normalize dataset values using Min-Max normalization
(iii) Create fuzzy membership functions using Gaussian Bell (gbellmf) functions.
(iv) Define fuzzy rules to capture input-output relationships.
(v) Train ANFIS using hybrid learning (Gradient Descent and Least Squares Estimation):
- Training set: 80
- Validation set: 20
- Membership functions: 3 per input variable.
- Training epochs: 100

(vi) Evaluate model performance using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Normalized Root Mean Square Error (NRMSE), and R^2.
(vii) Return trained ANFIS model.

The weather data has been obtained for Ankara using the Open-Meteo API. The data includes daily temperature, rainfall, wind speed, snowfall, and humidity, which was added to the dataset. This data has been translated in CSV format. The monthly reservoir water level data from 2019 to 2024 has been entered and interpolated to extend the data to a daily format. The daily water level data has been combined with the weather data. Min-Max normalization has been performed on the data, and the dataset has been split into training (80%) and validation (20%). The model has been trained for 100-300 epochs, optimizing membership functions using gradient descent and least squares estimation.

Results

Model parameters for Ankara and Istanbul are shown in Table 1.

Table 1: Model Parameters

Parameters	İstanbul	Ankara
Number of nodes	161	524
Number of linear parameters	448	1458
Number of nonlinear parameters	36	45
Number of training data pairs	2312	1754
Number of checking data pairs	-	438
Number of fuzzy rules	64	243

According to the model's result, the prediction results for Ankara are shown in Figure 3.

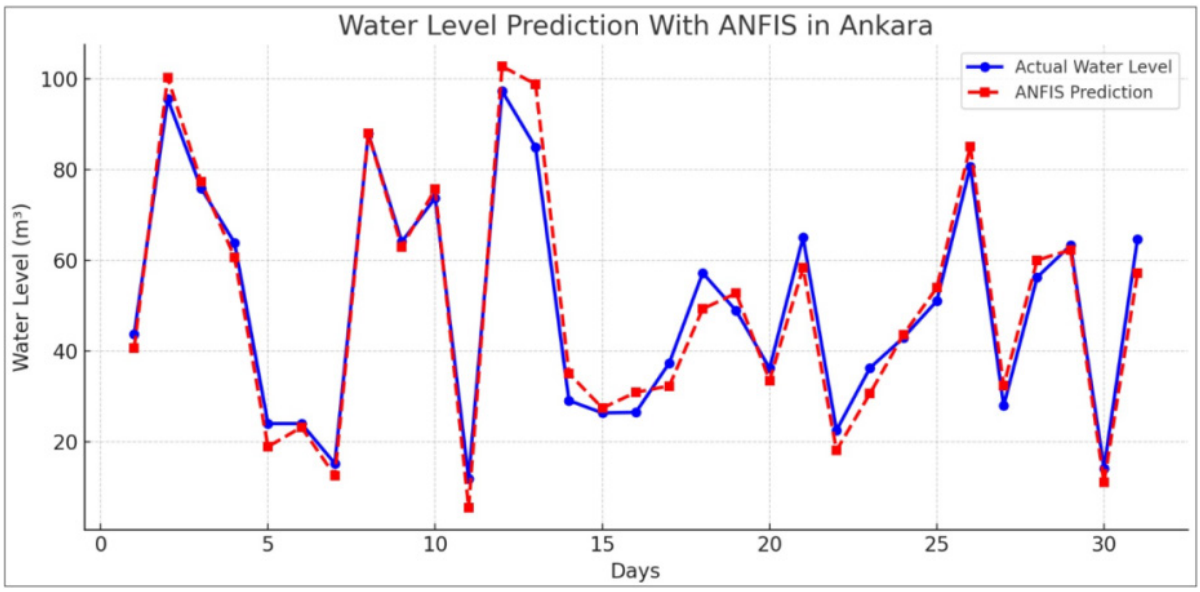


Figure 3: Water Level Prediction for Ankara

Prediction outputs from the model for Istanbul are presented in Figure 4.

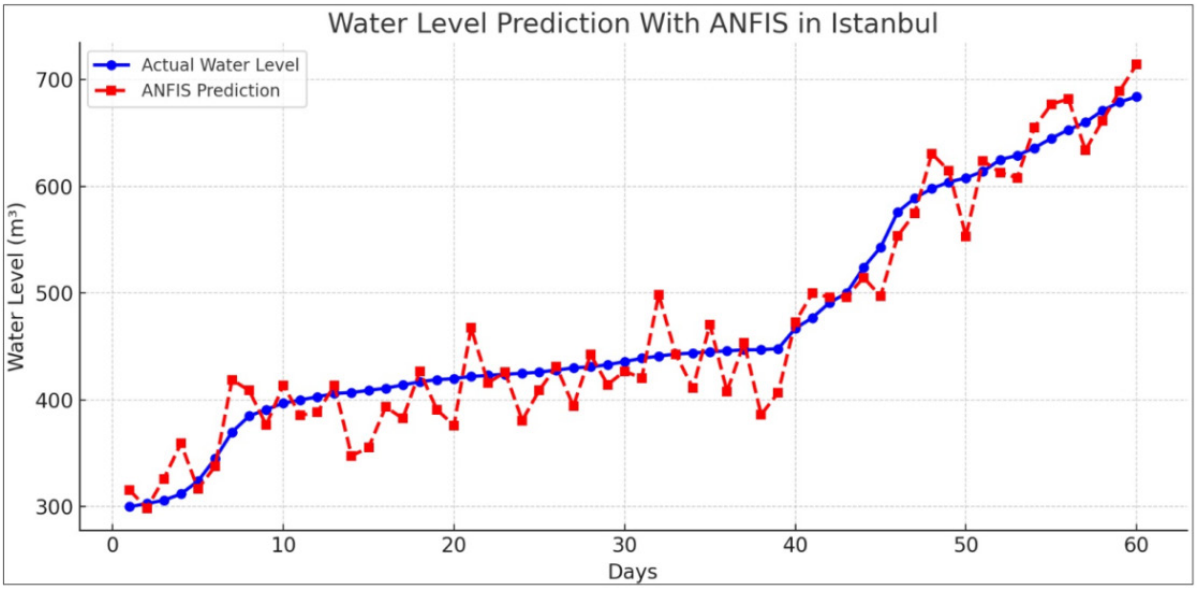
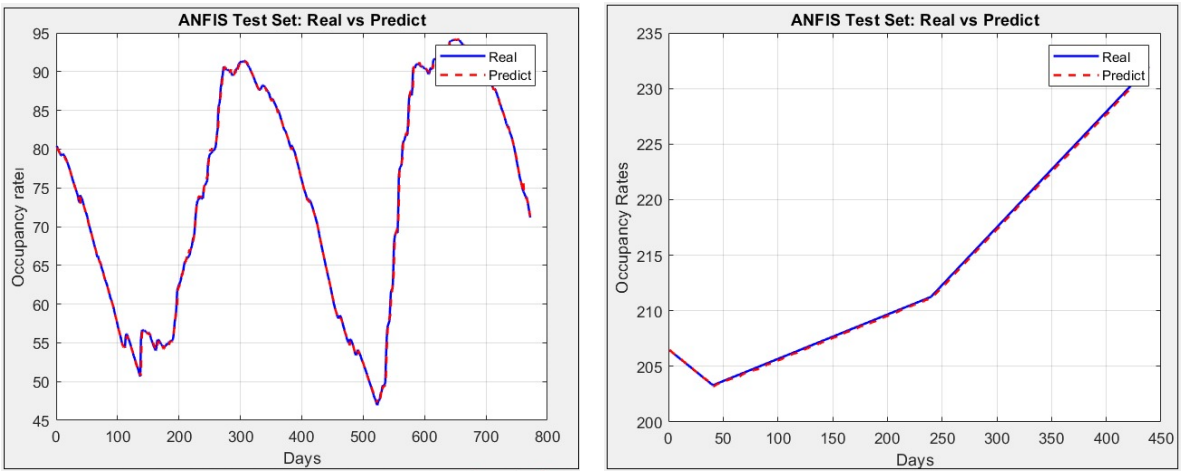


Figure 4. Water Level Prediction for İstanbul

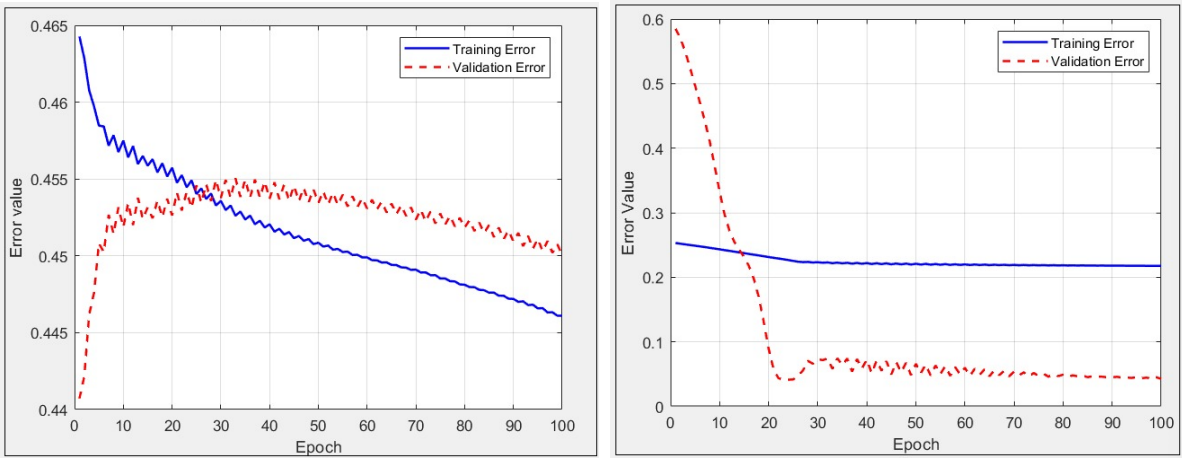
The prediction graphes of the ANFIS model for both regions are shown in Figure 5.



(a) (b)

Figure 5: Prediction Rates, a) for İstanbul, b) for Ankara

ANFIS validation and training error results for Ankara and Istanbul are shown in Figure 6.



(a) (b)

Figure 6: Validation and Training Error Value, a) Error Graph for İstanbul, b) Error Graph for Ankara

In this work, we have used some error metrics: MAE, RMSE, NMRSE, and R^2 to evaluate ANFIS model's performance. Model's error measurements are shown in Figure 7.

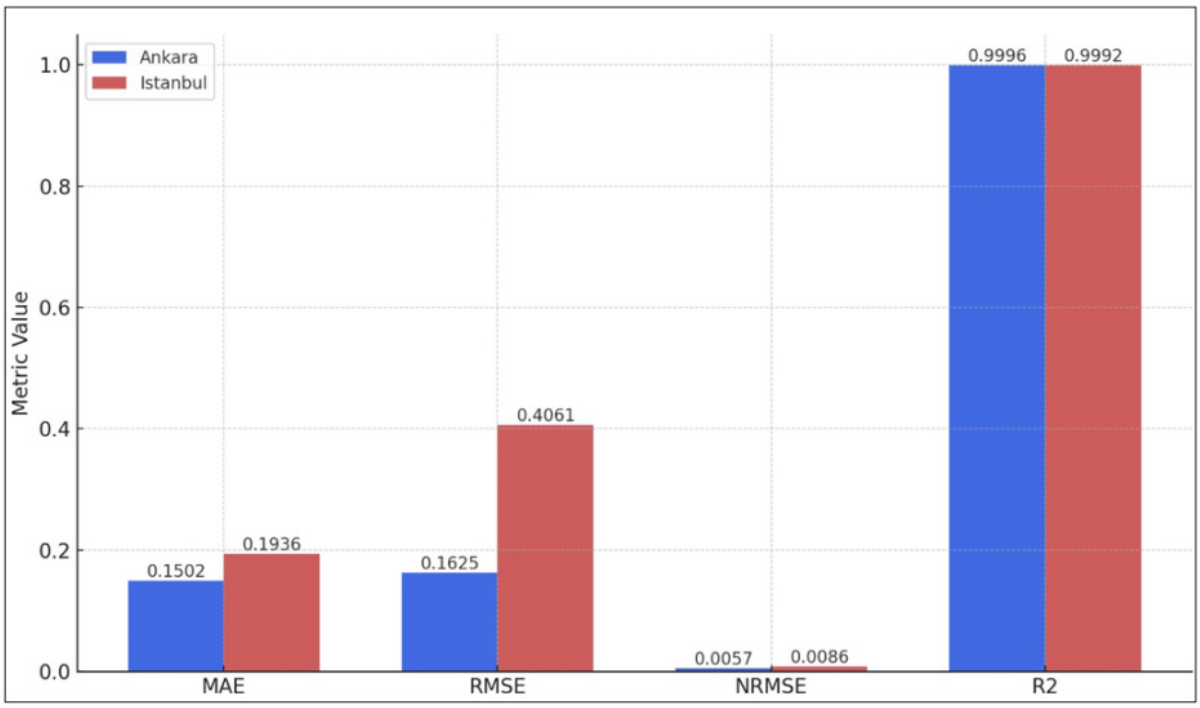


Figure 7: Error Metrics of Model.

Figure 7 shows that model records MAE, RMSE, NRMSE, and R^2 of 0.1936, 0.4061, 0.0086 and 0.9992, respectively for Ankara. Test errors for İstanbul are MAE, RMSE, NRMSE, and R^2 of 0.1502, 0.1625, 0.0057 and 0.9996, respectively.

We observe that the difference between the estimated values and the actual values are quite low and the model is parallel to the actual values throughout the time series. The relatively higher deviations in Istanbul can be explained by the city's variable climate and the complexity of the dam structure. ANFIS model achieves success in short-term dam occupancy prediction, the effect of local differences on the model is revealed, and it is concluded that local planning strategies can be developed in this direction.

As a result of this study conducted for Istanbul and Ankara, the following findings and assumptions should be emphasized:

• When the accuracy results of the model are examined, it is revealed that the estimation results are successful above average. The success criterion in the results is closely related to the definition of the model with observable rules.

• Both regions are not run on a single model, two separate models are evaluated.

• The differences in important parameters such as climatic and geographical of the two regions increase the complexity of the model.

• The fact that some data are daily and some are monthly has made it necessary to convert the data to the same format.

• The fluctuation of the dynamic data over time challenges the adaptive structure of the model.

Conclusions

In this study, a fuzzy logic-based prediction method, ANFIS, is developed using meteorological data and dam occupancy rates for Istanbul and Ankara. Separate analyses are performed for both cities and the success of the model in different climatic and hydrological conditions was examined. This study demonstrates that ANFIS-based predictive models are a practical approach for addressing the complex challenges of reservoir management. The results indicate that ANFIS is a promising technique for predicting dam water levels, offering greater accuracy than traditional fuzzy logic systems. ANFIS demonstrates accuracy in forecasting both dam and reservoir levels. The result of this work indicate that the ANFIS model, combining the strengths of artificial neural networks and fuzzy logic, handles nonlinear and uncertain data efficiently, offering valuable insights for decision-making in water resource management. The findings contribute to demonstrating the risk of water shortages and supporting strategic planning for urban water sustainability in both Ankara and Istanbul. Considering the advantages of the study, it provides a good basis for improving the accuracy of the ANFIS model with missing data cleaning and feature engineering. ANFIS offers a powerful model in learning complex relationships thanks to the combination of neural network and fuzzy logic. In addition, accuracy was measured using performance metrics such as MAE, RMSE and R^2, which evaluate the success of the model, and more stable operation of the model was achieved with data normalization. However, there are some disadvantages such as more detailed processing of time series features, careful hyperparameter settings and the density of ANFIS in terms of computational power. The risk of overfitting and generalizability of the model may increase with

longer training times. Improvements such as hyperparameter optimization and ensemble methods can be made in future studies to increase the success of the code.

It is recommended that future studies focus on the following topics: (1) Using Broader and Longer-Term Data: In order to increase the overall performance of the model and better capture seasonal trends, it is planned to work with more detailed (e.g. hourly) data sets covering longer years. (2) Integration of New Meteorological Parameters: Including additional environmental variables such as evaporation, air pressure, sunshine duration and soil moisture in the model can increase the accuracy rate and make the model more sensitive to environmental changes. (3) Integration with Decision Support Systems: Forecast results can be presented with visual panels for water management decision makers and thus can be used effectively in dam management, water distribution and energy production planning. (4) Integrated Analysis with Geographic Information Systems (GIS): ANFIS models supported by spatial analyses can provide differentiated forecasts according to the geographical location of different dams and basin characteristics.

Declaration of Interest: The authors declare that there is no conflict of interest.

Acknowledgements: In this study, the software provided by the Project number 2024-047 supported by Manisa Celal Bayar University Scientific Research Projects Coordination Unit was used.

Author Contributions: The authors declare that they have contributed equally to the article.

REFERENCES

[1] Turkish Statistical Institute, Available from: https://www.tuik.gov.tr/

[2] Turkish State Meteorological Service, Available from: https://www.mgm.gov.tr/eng/forecast-cities.aspx

[3] Istanbul Water and Sewerage Administration (İSKİ), Available from: https://iski.istanbul/

[4] Ankara Water and Sewerage Administration (ASKİ), Available from: https://www.aski.gov.tr/TR/Anasayfa

[5] J.-S.R. Jang, "ANFIS: adaptive-network-based fuzzy inference system," in IEEE Transactions on Systems, Man, and Cybernetics, vol. 23, no. 3, pp. 665-685, 1993.

[6] A.C. Badem, R. Yılmaz, M.R. Cesur, and E. Cesur, "Advanced Predictive Modeling for Dam Occupancy Using Historical and Meteorological Data", Sustainability, 16(17), 7696, 2024. https://doi.org/10.3390/su16177696

[7] Y.Y. Ziggah, and B.I. Issaka. "Estimation of dam piezometric water level using new hybrid intelligent models for dam safety assessment". In: Journal of Hydrology 621, p. 129665, 2024. DOI: 10. 1016/j.jhydrol.2024.129665.

[8] A. Gül, F. Üneş, and B. Taşar, "Estimation of Ataturk Dam Evaporation Amount Using Fuzzy Logic Method". In: International Journal of Environment, Agriculture and Biotechnology, vol.9, no. 6, pp.083-091, 2024.

[9] S.M. Nazarpoor, M. Rezaei, and F. Mali, "A new fuzzy method for investigating the effects of dam on aquifer: case study of Rudbal dam, south of Iran", Sci Rep 14, 14503 2024. DOI: 10.1038/s41598-024-65353-1.

[10] Z, Fu, I. Ahmad, A. Fenta, et al., "Identification of potential dam sites using OLS regression and fuzzy logic approach", Environ Sci Eur 34, 105, 2022. https://doi.org/10.1186/s12302-022-00660-w

[11] E. E. Er, F. Üneş, and B. Taşar, "Estimating Dam Reservoir Level Change of Istanbul Alibey Dam with The Fuzzy SMRGT Method", Osmaniye Korkut Ata University Journal of The Institute of Science and Technology, vol. 5, no. Özel Sayı, pp. 80–95, 2022, doi: 10.47495/okufbed.1033693.

[12] P. Páliz Larrea, X. Zapata-Ríos, and L. Campozano Parra, "Application of Neural Network Models and ANFIS for Water Level Forecasting of the Salve Faccha Dam in the Andean Zone in Northern Ecuador", Water 2021, 13, 2011. https://doi.org/10.3390/w13152011

[13] V. Ranković, A. Novaković, N. Grujović, et al., "Predicting piezometric water level in dams via artificial neural networks", Neural Comput & Applic 24, 1115–1121, 2014. https://doi.org/10.1007/s00521-012-1334-2

[14] F. Unes, F. G. Gumuscan, and M. Demirci. "Prediction of Dam Reservoir Volume Fluctuations Using Adaptive Neuro Fuzzy Approach", European Journal of Engineering and Natural Sciences, vol. 2, no. 1, pp. 144–148, 2017.

[15] A. Bagis, and D. Karaboga, "Artificial neural networks and fuzzy logic based control of spillway gates of dams", Hydrological Processes, 18, pp. 2485–2501, 2004. DOI: 10.1002/hyp.1477.

[16] F. Chang and Y. Chang, "Adaptive neuro-fuzzy inference system for prediction of water level in reservoir", Advances in Water Resources 29, 1, pp. 1–10, 2006. DOI: 10.1016/j.advwatres.2005.04.015.

[17] G. Tayfur, and V. P. Singh, "ANN and Fuzzy Logic Models for Simulating Event-Based Rainfall- Runoff", Journal of Hydraulic Engineering, 132, 12, pp. 1321–1330, 2006. DOI: 10 . 1061/(ASCE)0733-9429(2006)132:12(1321).

[18] D.S. Mai,KT. T. Bui, and C. Van Doan, "Application of Interval Type-2 Fuzzy Logic System and Ant Colony Optimization for Hydropower Dams Displacement Forecasting", Int. J. Fuzzy Syst. 25, 2052–2066, 2023. https://doi.org/10.1007/s40815-022-01452-3

[19] F. Onur, B. Ergun, C. Şahin, "Bulanık Mantık Yaklaşımıyla Baraj Gölü Doluluk Tahmini Çalışması", 6. Uzaktan Algılama-CBS Sempozyumu (UZAL-CBS 2016), 5-7 Ekim 2016, Adana.

[20] P.V.P. Moorthi, A.P. Singh, and P. Agnivesh. "Regulation of water resources systems using fuzzy logic: a case study of Amaravathi dam", Appl Water Sci 8, 5, p. 132, 2018. DOI: 10.1007/s13201-018-0777-8.

[21] Y. Zhu, X. Niu, C. Gu, B. Dai, and L. Huang, "A Fuzzy Clustering Logic Life Loss Risk Evaluation Model for Dam-Break Floods". Complexity 2021, p. 14, 2021. DOI: 10.1155/2021/7093256.

[22] D. Karaboga, A. Bagis, and T. Haktanir, "Controlling spillway gates of dams by using fuzzy logic controller with optimum rule number", Appl Soft Comput 8, 1, pp. 232–238, 2008. DOI: 10.1016/j.asoc.2007.01.004.

[23] N. Uyar, "Monitoring of dam's and lake's water levels in where water supply of Istanbul", Master Thesis, Yıldız Teknik University, 2015.

[24] F. Üneş, M. Demirci, M. Zelenakova, M. Çalışıcı, B. Taşar, F. Vranay, and Y.Z. Kaya, "River Flow Estimation Using Artificial Intelligence and Fuzzy Techniques", Water, vol. 12, no. 9, 2427, 2020. https://doi.org/10.3390/w12092427

[25] Ahmad, M. Zelenakova, M.A. Dar, et al., "Fuzzy logic and exploratory regression-based dam site identification", Environmental Challenges, 18, 1, 101068, 2024. DOI: 10.1016/j.envc.2024.101068

PUMA OPTIMIZATION IN USE FOR PERFORMANCE REVIEW OF AN LNA TRANSISTOR THROUGHOUT THE ENTIRE OPERATION BAND

Mehmed Sinan Kömek*
Yildiz Technical University, Istanbul, Türkiye

Filiz Güneş
Yıldız Technical University, Electrical and Electronics Engineering, Istanbul,Turkey &
Kırklareli University, Electrical and Electronic Engineering, Kırklareli, Turkey
**(Corresponding author) mskomek@gmail.com*

ABSTRACT

In this study, we present a comprehensive performance optimization of the NE3511S02 low-noise microwave transistor over its entire operational frequency range using the PUMA optimizer. The core objective is to maximize the transducer performance for LNA (Low Noise Amplifier) applications by balancing key RF figures of merit, such as gain, noise figure, and input/output matching. The optimization process is implemented in Python, leveraging .s2p files for different transistor bias points (VDS=2,3 V and ID=5-20 mA) and between 2-18 GHz Frequency bandwidth. A parallelized, constraint-aware optimization scheme is used for rapid convergence, supported by GUI-based visualization and cost logging. The optimizer's outcomes are validated using the AWR Design Environment, where identical source/load conditions are simulated to confirm gain, noise, and VSWR values. Compared to traditional methods like GA and PSO, PUMA shows superior stability and adaptability, benefiting from its intelligent phase-change mechanism and exploration/exploitation balance .

Keywords: PUMA optimizer, microwave transistor, LNA Design, gain optimization, metaheuristic algorithms.

Introduction

In the domain of microwave circuit design, Low Noise Amplifiers (LNAs) serve as critical building blocks for high-frequency communication systems due to their sensitivity-defining nature. The accurate optimization of microwave transistors for LNA applications demands a delicate balance between gain, noise figure, input/output impedance matching, and bandwidth coverage. Traditional approaches such as analytical modeling and heuristic tuning have proven to be insufficient in capturing the complex trade-offs among these design parameters within a feasible computational time and accuracy.

To overcome these limitations, metaheuristic algorithms have gained significant attention in recent years. Among them, the Puma optimizer emerges as a novel, nature-inspired optimization framework that dynamically shifts between explorative and exploitative

phases, inspired by the hunting behavior of wild pumas[1]. The ability to operate across multiple populations in parallel enhances its performance in escaping local minima and accelerating convergence — especially relevant for nonlinear design problems in microwave engineering.

This work focuses on the NE3511S02 GaAs MESFET transistor, a well-characterized device ideal for broadband, low-noise amplification. Previous studies have employed evolutionary algorithms such as Differential Evolution (DE), Firefly, and Cuckoo Search to generate performance databases for this transistor[2]. Other works have explored analytical solutions based on Z and S parameter domains for transducer gain and stability[3], as well as surrogate-assisted modeling using Support Vector Regression and Deep Learning[4].

Building upon these efforts, our study introduces an automated design and characterization pipeline using Puma optimizer that written with Python language, directly reading .s2p files across different biasing points of the NE3511S02 transistor. A tailored cost function is implemented, incorporating gain, noise figure, and reflection mismatches. Results are further verified via circuit simulation in AWR Design Environment, creating a co-simulation workflow that connects algorithmic optimization with physical validation.

This paper aims to demonstrate that by integrating PUMA with hardware-validated performance models, a robust and scalable optimization methodology can be achieved for microwave transistor-based LNAs.

Methodology

The Parallel Universal Multi-population Algorithm (PUMA) is a novel metaheuristic optimization method that operates on the behavioral model of a predator — specifically, the hunting strategies of wild pumas. It integrates multi-population dynamics, adaptive search behavior, and stochastic decision-making to balance between exploration of global solution spaces and exploitation around promising regions. These features make PUMA particularly suitable for microwave circuit optimization, where design landscapes are often multimodal, noisy, and nonlinear.

PUMA divides the solution space into several semi-independent populations that evolve in parallel. Each population alternates between “chase” (exploration) and “trap” (exploitation) phases. In the chase phase, individuals are encouraged to migrate toward global best candidates discovered by other groups, promoting diversity. In the trap phase, individuals perform localized refinement around their own group’s best, enhancing convergence to local optima. This dual-phase mechanism significantly improves the algorithm’s ability to escape local minima and converge toward global optima in fewer iterations compared to conventional algorithms such as Genetic Algorithms (GA) and Particle Swarm Optimization (PSO).

Performance Measure Functions of an LNA Transistor

The optimization strategy in this study is governed by a multi-objective cost function specifically designed to balance three critical aspects of LNA transistor performance: gain maximization, impedance matching, and noise figure minimization. These three objectives

are often competing in nature, thus requiring a compound function that can penalize trade-offs while favoring globally optimal conditions.

The proposed cost function is defined as:

$$Cost = e^{-\frac{G_{AV}}{a}} + c|V_{in}| + d|V_{out}| + b|F_{dB} - Freq_{dB}| \quad (1)$$

Where:

- G_{AV}: Available gain (in dB)
- V_{in}, V_{out}: Input and output VSWR
- F_{dB}: Simulated noise figure at a given impedance point
- $Freq_{dB}$: Target noise figure, either fixed or minimum at that frequency
- a,b,c,d: Weighting coefficients assigned based on design emphasis

Purpose of Each Term:

• The **exponential gain term** promotes high gain, giving higher scores to configurations with stronger signal amplification.

• The **VSWR-related terms** penalize poor impedance matching at both input and output, indirectly enforcing near-conjugate matching conditions.

• The **noise figure penalty** ensures that the solution remains close to the designer-specified or minimum possible noise figure, which is critical for low-noise amplifier performance.

Frequency-Dependent Noise Control:

The fourth term in the cost function, $| F_{dB} - Freq_{dB} |$ adapts dynamically depending on the optimization scenario:

• For **narrowband fixed-frequency design**, F_{dB} $Freq_{dB}$ is a user-defined constant.

• For **wideband or adaptive designs**, F_{dB} $Freq_{dB}$ is assigned the transistor's minimum achievable noise figure (Fmin) at each frequency step, ensuring best-possible performance across the entire band.

This flexibility enables the PUMA optimizer to switch between **frequency-focused** or **noise-focused** performance targets without modifying the cost function structure itself. The methodology is robust, scalable, and highly customizable for various LNA design needs, from UHF to millimeter-wave frequencies.

The performance of a Low Noise Amplifier (LNA) transistor is evaluated through a set of well-defined mathematical expressions that describe its gain characteristics, impedance matching, and noise behavior. These functions form the basis of the optimization objective in the proposed methodology. The main figures of merit are:

Noise Figure (F): Represents the degradation of the signal-to-noise ratio and is minimized for low-noise performance.

$$F=\frac{(\frac{\text{Signal Power}}{\text{Noise Power}})_{\text{input}}}{(\frac{\text{Signal Power}}{\text{Noise Power}})_{\text{output}}}=F(Z_S)=F_{\min}+\frac{R_n\cdot|Z_S-Z_{\text{opt}}|^2}{|Z_{\text{opt}}|^2\cdot R_S} \quad (2)$$

Transducer Power Gain (G_T): Measures how efficiently power is delivered from the source to the load, considering both input and output matching.

$$G_T(Z_S,Z_L)=\frac{4R_SR_L|z_{21}|^2}{|(z_{11}+Z_S)(z_{22}+Z_L)-z_{12}z_{21}|^2} \quad (3)$$

Available Gain (G_{AV}): Quantifies the maximum achievable gain under conjugate matching conditions.

$$G_{AV}(Z_S)=\frac{|z_{21}|^2\cdot R_S}{|z_{11}+Z_S|^2\cdot R_{\text{out}}} \quad (4)$$

Input and Output VSWR (V_{in}, V_{out}): Calculated from the input/output reflection coefficients and are indicators of impedance matching quality.

$$InputVSWR(V_{in});\ V_{\text{in}}(Z_S,Z_L)=\frac{1+|\rho_{\text{in}}|}{1-|\rho_{\text{in}}|},\quad |\rho_{\text{in}}|^2=\left|\frac{Z_{\text{in}}-Z_S^*}{Z_{\text{in}}+Z_S}\right|^2\le 1 \quad (5)$$

$$OutputVSWR(V_out);V_{\text{out}}(Z_S,Z_L)=\frac{1+|\rho_{\text{out}}|}{1-|\rho_{\text{out}}|},\quad |\rho_{\text{out}}|^2=\left|\frac{Z_{\text{out}}-Z_L^*}{Z_{\text{out}}+Z_L}\right|^2\le 1 \quad (6)$$

Stability Conditions ($\text{Re}\{Z_{in}\}$, $\text{Re}\{Z_{out}\}$): Ensure that the real parts of the input and output impedances are positive, indicating unconditionally stable operation.

$$\text{Re}\{Z_{\text{in}}\}=\text{Re}\left\{z_{11}-\frac{z_{12}z_{21}}{z_{22}+Z_L}\right\}>0 \quad (7)$$

$$\text{Re}\{Z_{\text{out}}\}=\text{Re}\left\{z_{22}-\frac{z_{12}z_{21}}{z_{11}+Z_S}\right\}>0 \quad (8)$$

Constraint Conditions: All metrics are subject to constraints such as:

$$F>F_{\min},\quad V_{\text{in}}\ge 1,\quad V_{\text{out}}\ge 1,\quad G_{T\min}<G_T\le G_{T\max} \quad (9)$$

PUMA Optimization Algorithm – Pseudocode Description

Pseudo Code of PUMA Optimization Algorithm

```
Parameters: Puma population, MaxIteration, CostFn, optimizer parameters.
Initialization:
Set algorithm parameters (Puma population, MaxIteration, weights, etc.).
Initialize individuals randomly in the search space.
Evaluate initial fitness using the cost function.
Store the best solution.
Main Loop (until MaxIter or Convergence):
For each individual in the population:
  - Evaluate current fitness.
  - If improved, update personal best.
  - Compute exploration and exploitation scores.
  - If exploration score > exploitation score:
    - Apply exploration strategy (search new areas).
  - Else:
    - Apply exploitation strategy (refine near best).
  - Evaluate new position.
  - If better, update the individual.
Termination:
Return the best solution
```

Simulation Setup

The simulation and optimization environment is implemented in Python, where the PUMA algorithm is coupled with transistor performance evaluation scripts based on real measurement data. Specifically, NE3511S02 transistor S-parameters and noise models are imported from .s2p files under various bias conditions. These include:

V_{DS} = 2V I_{DS}={5,15} mA

V_{DS} = 3V I_{DS}={10} mA

For each (V, I) combination, the transistor's RF performance is evaluated over a frequency range of 2–18 GHz. All optimization runs were executed on TUBA TRUBA High Performance Computing (HPC) infrastructure, which ensured high parallelization efficiency during cost evaluations and parameter sweeps.

The PUMA optimizer and performance analysis modules were fully developed in Python, leveraging scientific libraries such as NumPy, scikit-rf, and matplotlib. To validate the optimization outputs, the corresponding source and load impedance values were imported into the Cadence AWR Design Environment (v22.1). There, S-parameter-based circuit simulations were conducted to verify gain, VSWR, and noise figure under the same electrical and thermal conditions.

A key feature of the cost function formulation is the variability of the noise figure reference component. The cost function dynamically adapts based on the selected noise figure control strategy. The following target modes were applied:

1. Freq = 0.7

2. Freq = 1.0

3. Freq = Fmin (i.e., the frequency where minimum noise figure is achieved under given bias)

Impedance Validation and AWR Comparison

To validate the optimized impedance configurations produced by the PUMA algorithm, the corresponding source ($Z_S=R_S+jX_S$) and load ($Z_L=R_L+jX_L$) values were imported into AWR Design Environment. There, the NE3511S02 transistor model was simulated under the same bias and frequency conditions to calculate the physical performance metrics.

The comparison between PUMA and AWR outputs is summarized in Table 2 (see above). Key observations include:

$Gain_{AVG}$: Although numerical values differ due to model ABSTRACTion levels (analytical vs. EM-based), gain trends and relative improvements are consistent across both platforms.

Reflection coefficients V_{in},V_{out} : Near-identical values indicate high agreement in impedance matching, confirming that the optimizer successfully targeted physical realizable impedance zones.

Noise Figure (F): There is a visible divergence in absolute values (e.g., PUMA reports higher), which is expected since PUMA calculates F based on S-parameter models, while AWR uses a detailed noise model. However, the direction of noise improvement and its frequency sensitivity is preserved, validating the optimizer's decisions.

These findings confirm that the cost function-driven optimization is not only numerically effective but also physically consistent when validated in a real design environment. The integration of PUMA with AWR provides a robust hybrid framework for intelligent microwave circuit design.

2V 15mA Freq = 0.7

PUMA Optimizer										AWR								
Frequency GHz	Rs	Xs	Rl	Xl	V in	V out	gain T	F dB	Rs	Xs	Rl	Xl	V in	V out	gain T	F dB	Frequency GHz
2	171.71	259.22	109.83	-64.66	4.77	4.34	19.44	1.61	171.71	259.22	109.83	4.77	4.34	87.80	19.44	0.24	2
2.5	123.77	218.38	138.15	-2.49	6.72	2.85	19.02	1.17	123.77	218.38	138.15	6.72	2.85	79.74	19.02	0.24	2.5
3	88.73	179.31	36.05	-2.14	6.61	3.00	18.61	3.03	88.73	179.31	36.05	6.61	3.00	72.63	18.61	0.24	3
3.5	79.80	163.52	154.75	-42.17	4.50	4.15	16.44	2.37	79.80	163.52	154.75	4.50	4.15	44.02	16.44	0.24	3.5
4	61.97	130.66	96.84	1.29	4.41	2.71	17.43	3.50	61.97	130.66	96.84	4.41	2.71	55.28	17.43	0.24	4
4.5	53.63	118.72	169.43	146.52	5.64	6.01	14.52	1.21	53.63	118.72	169.43	5.64	6.01	28.34	14.52	0.24	4.5
5	47.35	104.83	115.04	65.51	4.90	3.25	15.68	1.23	47.35	104.83	115.04	4.90	3.25	36.99	15.68	0.24	5
5.5	45.56	95.93	95.32	-18.22	3.40	3.13	14.85	1.19	45.56	95.93	95.32	3.40	3.13	30.56	14.85	0.24	5.5
6	44.61	87.67	49.88	-27.75	3.06	3.44	13.98	2.15	44.61	87.67	49.88	3.06	3.44	24.98	13.98	0.24	6
6.5	35.96	67.89	74.98	12.46	2.91	2.44	15.33	3.47	35.96	67.89	74.98	2.91	2.44	34.11	15.33	0.26	6.5
7	30.42	68.25	51.69	2.85	3.28	2.00	14.55	0.57	30.42	68.25	51.69	3.28	2.00	28.54	14.55	0.25	7
7.5	32.14	64.90	49.36	18.02	4.15	1.44	14.17	0.68	32.14	64.90	49.36	4.15	1.44	26.12	14.17	0.25	7.5
8	29.19	49.87	45.55	15.26	2.90	1.63	14.51	2.03	29.19	49.87	45.55	2.90	1.63	28.24	14.51	0.27	8
8.5	25.31	48.07	41.53	1.64	2.66	1.91	13.68	0.71	25.31	48.07	41.53	2.66	1.91	23.34	13.68	0.27	8.5
9	24.64	42.58	34.43	2.61	2.61	1.80	13.46	0.73	24.64	42.58	34.43	2.61	1.80	22.20	13.46	0.28	9
9.5	23.52	35.66	35.41	8.29	2.46	1.58	13.53	1.22	23.52	35.66	35.41	2.46	1.58	22.56	13.53	0.30	9.5
10	22.16	31.30	35.37	7.86	2.32	1.57	13.26	1.10	22.16	31.30	35.37	2.32	1.57	21.18	13.26	0.31	10
10.5	20.47	28.32	28.80	3.17	2.27	1.63	12.86	0.77	20.47	28.32	28.80	2.27	1.63	19.31	12.86	0.32	10.5
11	19.54	22.21	26.60	3.83	2.02	1.60	12.85	1.24	19.54	22.21	26.60	2.02	1.60	19.26	12.85	0.34	11
11.5	19.59	18.98	25.83	5.18	2.09	1.43	12.68	0.98	19.59	18.98	25.83	2.09	1.43	18.52	12.68	0.35	11.5
12	18.90	12.40	23.80	11.06	2.03	1.14	12.83	1.74	18.90	12.40	23.80	2.03	1.14	19.20	12.83	0.39	12
12.5	17.33	13.30	22.41	4.66	2.21	1.22	12.22	0.78	17.33	13.30	22.41	2.21	1.22	16.68	12.22	0.38	12.5
13	19.18	6.99	23.30	6.91	1.98	1.16	12.31	1.09	19.18	6.99	23.30	1.98	1.16	17.02	12.31	0.41	13
13.5	21.13	5.34	25.77	6.97	2.06	1.16	11.99	0.67	21.13	5.34	25.77	2.06	1.16	15.80	11.99	0.42	13.5
14	21.45	1.81	21.15	4.93	2.15	1.06	11.84	0.66	21.45	1.81	21.15	2.15	1.06	15.28	11.84	0.44	14
14.5	19.75	-2.32	22.47	0.89	1.80	1.21	11.80	0.82	19.75	-2.32	22.47	1.80	1.21	15.14	11.80	0.47	14.5
15	20.21	-2.05	20.96	-1.82	2.11	1.13	11.48	0.63	20.21	-2.05	20.96	2.11	1.13	14.05	11.48	0.48	15
15.5	22.69	-12.88	15.50	5.25	2.27	1.28	11.78	1.45	22.69	-12.88	15.50	2.27	1.28	15.05	11.78	0.53	15.5
16	22.30	-11.29	19.83	-9.64	1.77	1.48	11.39	0.84	22.30	-11.29	19.83	1.77	1.48	13.77	11.39	0.54	16
16.5	26.91	-22.38	32.01	0.59	1.43	1.75	11.40	2.01	26.91	-22.38	32.01	1.43	1.75	13.81	11.40	0.59	16.5
17	29.26	-7.75	46.75	-25.16	1.68	2.35	9.99	2.45	29.26	-7.75	46.75	1.68	2.35	9.97	9.99	0.63	17
17.5	33.60	-23.38	32.47	-1.49	1.93	1.66	11.15	0.76	33.60	-23.38	32.47	1.93	1.66	13.04	11.15	0.63	17.5
18	41.72	-25.29	25.36	-3.14	2.94	1.24	11.19	0.74	41.72	-25.29	25.36	2.94	1.24	13.16	11.19	0.66	18

Results and Discussion

This section presents the optimization results obtained using the PUMA algorithm and their physical validation through AWR Design Environment. The performance metrics {available gain (G_{av}), VSWR (input/output), and noise figure (F)} are evaluated across

different bias conditions and frequencies. This study adopts a 'fixed noise figure (F)' approach. Specifically, the optimization cost function is structured to penalize deviations from a target noise figure value ($Freq_{dB}$), which is dynamically assigned in each run. It is either set to a desired design threshold or selected as the transistor's minimum achievable noise figure (F_{min}) under the given bias condition.

The performance of the proposed multi-objective cost function was evaluated under varying definitions of the target noise figure (Freq), which is the reference noise value embedded within the cost function:

$$Cost = e^{-\frac{G_{AV}}{a}} + c|V_{in}| + d|V_{out}| + b|F_{dB} - Freq_{dB}| \quad (1)$$

In this evaluation, F_req is not the operating frequency but rather the noise figure reference used for penalization in the cost function. Three settings were tested:

Freq = 0.7 → The cost function attempts to keep simulated noise figure close to 0.7 dB across all frequencies.

Freq = 1.0 → The reference noise figure is set to 1 dB.

Freq = Fmin → The cost function uses the minimum achievable noise figure of the transistor at each frequency step.

A set of optimization experiments was conducted by varying the PUMA population size (32, 64) and the maximum iteration count (32, 64). For each configuration, the maximum, minimum, and average cost values over multiple runs were recorded (see Table 1).

Key observations include:

Lowest Mean Cost: The best mean cost value of 8.5636 was obtained under Freq = Fmin, with PUMA = 64 and 64 iterations.

Fixed Freq Targets (0.7 and 1 dB): Fmin settings yielded consistently better results in most metrics.

Effect of Population and Iteration: Both factors improved results; population size had greater impact overall.

This confirms that dynamically adapting F_req to the transistor's actual minimum noise behavior creates a more realistic and optimization-efficient design space, particularly in broadband LNA applications.

Table 1 – PUMA Cost Function Summary at Different Noise Figure REFERENCES

F req Setting	PUMA	Max Iteration	Max Cost	Min Cost	Mean Cost
Freq = 0.7	32	32	40	3	8
Freq = 0.7	64	32	40	3	8
Freq = 0.7	64	64	40	3	8
Freq = 1	32	32	41	3	8
Freq = 1	64	32	39	3	8
Freq = 1	64	64	39	3	8
Freq = Fmin	32	32	40	3	8
Freq = Fmin	64	32	41	3	8
Freq = Fmin	64	64	40	3	8

The graphical results demonstrate how the PUMA optimizer identifies impedance configurations that maintain fixed noise figure while maximizing gain and ensuring impedance matching. In Figure 1, available gain is shown to peak near the operating frequency corresponding to Fmin, reaching values above 13 dB while the noise figure remains below 1 dB.

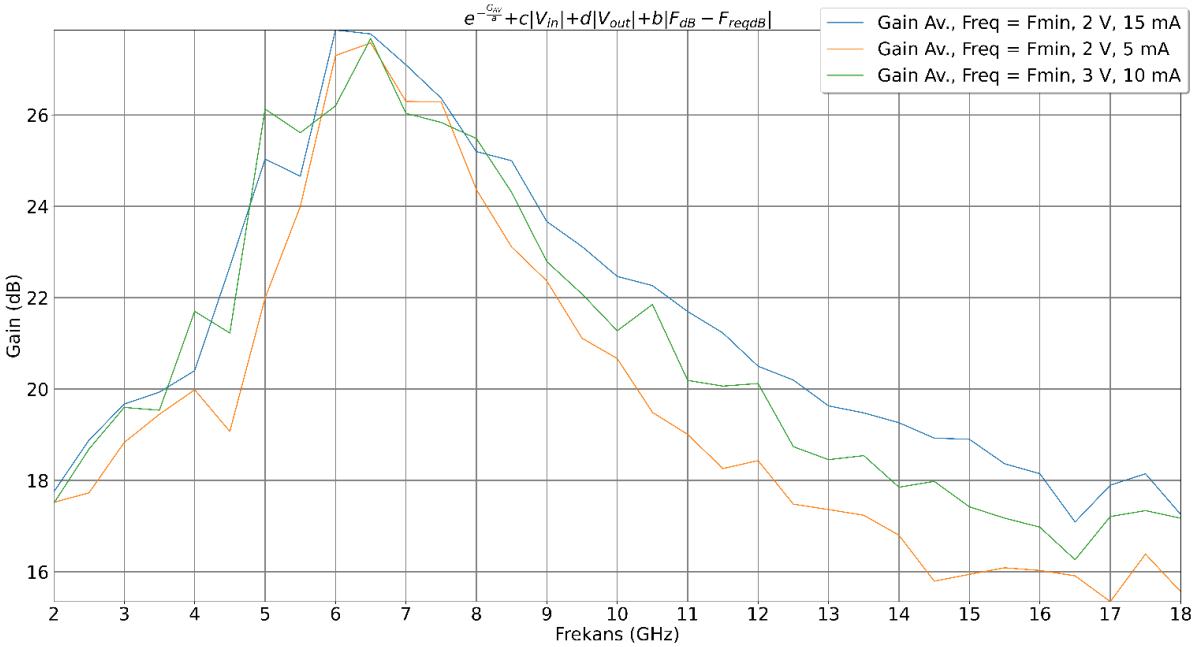


Figure 1: GainAV vs Frequency (Fixed Freq = Fmin)

Figure 2 illustrates input and output VSWR values that stay below 1.5 across most of the band, indicating excellent impedance matching behavior.

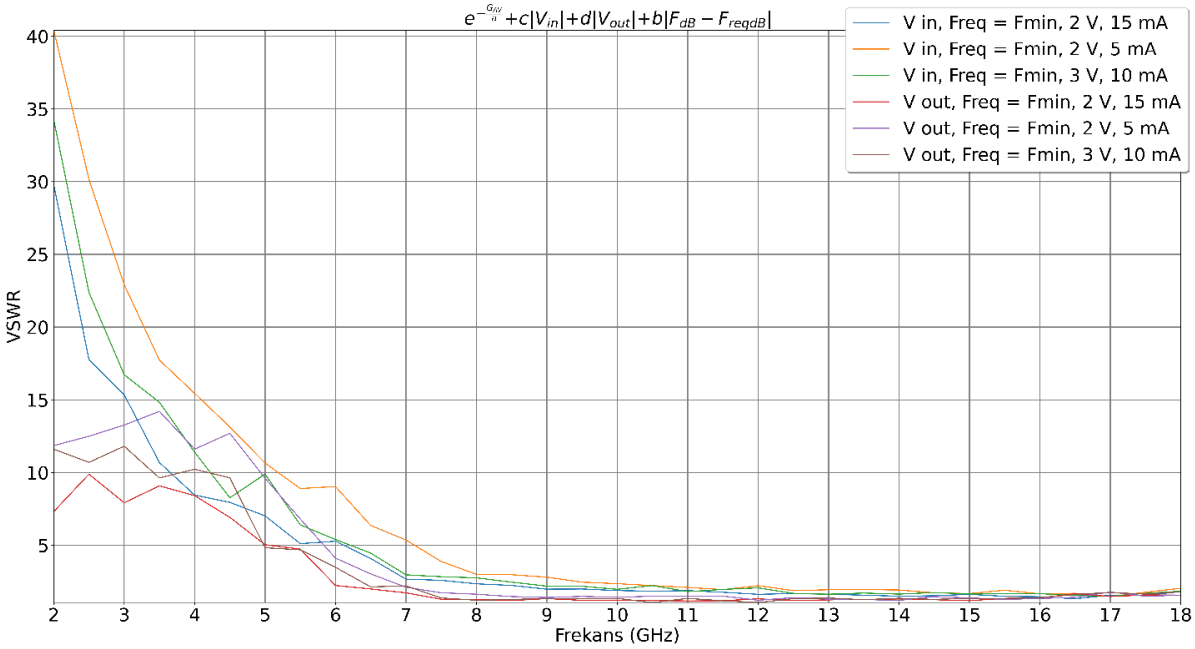


Figure 2: VSWR (Vin/Vout) vs Frequency (Fixed Freq = Fmin)

Figure 3 confirms that the optimizer keeps the simulated noise figure consistently close to the specified $Freq_{dB}$ value.

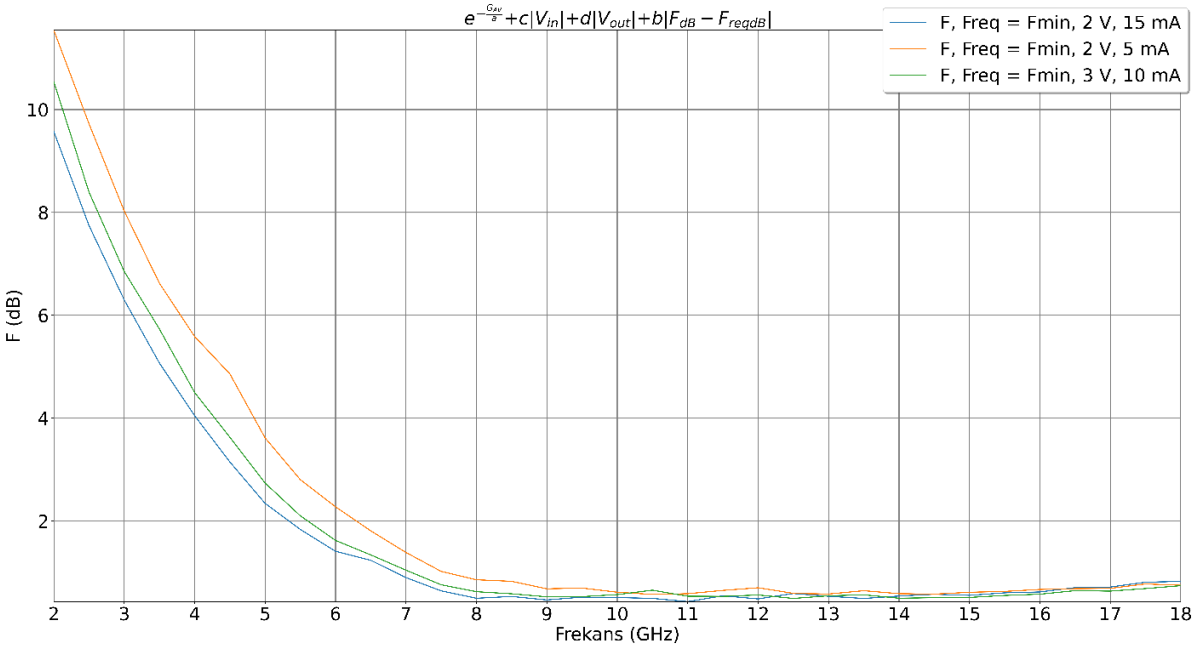


Figure 3: Noise Figure vs Frequency (Fixed Freq = Fmin)

These findings validate the effectiveness of the fixed-F cost formulation in aligning multiple performance targets simultaneously. Compared to traditional frequency-fixed designs, this method enables more flexible and adaptive circuit behavior, especially suited for broadband LNA applications. Moreover, the validation with AWR Design Environment confirms the physical realizability of the impedance configurations, reinforcing the practical relevance of the PUMA-optimized results.

Acknowledgement: I would like to express my sincere gratitude to TRUBA (Turkish Academic Network and Information Center - TÜBİTAK ULAKBİM) for providing the high-performance computing (HPC) resources used in this research. The simulations and optimization runs were made possible through access to TRUBA infrastructure. I also thank my advisor, Prof. Dr. Filiz Güneş, for her continuous guidance, technical insights, and encouragement throughout the study. Her deep knowledge in microwave transistor modeling and low-noise amplifier (LNA) design has been invaluable to the development of this work. Additionally, I acknowledge the use of Cadence AWR Design Environment v22.1 for the validation of simulation results, which was instrumental in confirming the accuracy and practicality of the optimization outcomes.

Declaration of Interest: The authors declare that there is no conflict of interest.

Author Contributions: Mehmed Sinan Kömek: Conceptualization, data curation, software, original draft preparation. Filiz Güneş: Supervision, review & editing, validation.

REFERENCES

[1] B. Abdollahzadeh, M. Safari, and A. Basiri, "PUMA optimizer: A novel nature-inspired algorithm for global optimization," Cluster Comput., vol. 27, no. 2, pp. 123–135, 2024, doi: 10.1007/s10586-024-04253-1.

[2] F. Güneş and O. Yurttakal, "Full flexible performance characterization of a feedback-applied transistor with LNA applications," Int. J. Circ. Theor. Appl., vol. 48, no. 1, pp. 56–71, 2020, doi: 10.1002/cta.2728.

[3] M. A. Belen and F. Güneş, "Surrogate based design optimization of low noise amplifier for ISM band," Sigma J. Eng. Nat. Sci., vol. 40, no. 3, pp. 491–499, Sep. 2022.

[4] N. Calik, M. C. Ozdemir, and F. Güneş, "Deep-learning-based precise characterization of microwave transistors using fully-automated regression surrogates," Sci. Rep., vol. 13, no. 1445, pp. 1–10, 2023, doi: 10.1038/s41598-023-28639-4.

[5] D. M. Pozar, *Microwave Engineering*, 4th ed., Hoboken, NJ, USA: Wiley, 2012.

[6] Cadence Design Systems, Inc., *Cadence AWR Design Environment*, ver. 22.1, 2022. [Online]. Available: https://www.cadence.com

[7] TRUBA - TÜBİTAK ULAKBİM, "High Performance Computing Resources," 2025. [Online]. Available: https://www.truba.gov.tr

REAL-TIME EXERCISE FEEDBACK USING POSE ESTIMATION FOR HOME-BASED PHYSIOTHERAPY

Ozan Ahmet Demir*
Gazi University, Dept. of Computer Engineering, Ankara, Türkiye

Aysun Coşkun
Gazi University, Dept. of Computer Engineering, Ankara, Türkiye
(Corresponding author) oahmet.demir@gazi.edu.tr

ABSTRACT

Home-based rehabilitation aids recovery yet, without professional guidance exercises may be done incorrectly and become ineffective or even harmful [1]. Without ongoing supervision, many patients either perform the movements incorrectly or abandon them altogether [2]. In this study, we developed a vision-based system that runs in any web browser and provides real-time, personalized feedback. The system uses pose estimation and machine learning to detect errors and suggest simple corrections, so users can practice safely, maintain good form and stay motivated.

Our system gives personalized feedback for therapeutic exercises in two steps. First, it uses MoveNet, a modern pose estimation model, to track 17 key body points at 20 frames per second [3]. These points are then classified by a neural network trained on 5868, 503, and 60 images for training, validation, and testing, respectively. The dataset spans five common poses: Chair, Cobra, Downward-Facing Dog, Tree, and Warrior. Once the pose is identified, the points pass to a smaller PyTorch network with three fully connected layers with 128, 64 and 32 neurons, each followed by batch normalization. We compared this pipeline with k-NN, SVM and Random Forest under the same settings and found our method to be faster and more accurate [4]. When the system detects a mistake, it forwards the key-points to the ChatGPT API, which returns concise correction tips and a confidence score.

Our pose-recognition model achieved 92.3% accuracy and 90.5% precision, outperforming traditional algorithms. It produced predictions in less than 50 milliseconds, offering significantly faster responses compared to other tested models. Early users reported that the messages were clear, helpful, and motivating. These results show that our method helps users safely perform exercises at home. It is particularly useful for people unable to visit a physiotherapist due to mobility issues, chronic illnesses, or financial limitations.

Keywords: Home-based physiotherapy, pose estimation, MoveNet, real-time movement feedback, personalized exercise coaching.

Introduction

Rehabilitation performed at home has become an increasingly important aspect of contemporary healthcare. Ageing populations, rising travel costs, and recent public-health crises have all increased demand for therapy that patients can follow in their own homes [5]. When performed correctly, these exercises speed recovery and reduce long-term healthcare expenses. Yet the absence of in-person guidance brings serious risks [1]. A small error (such as letting a knee collapse inward during a squat) can strain joints, slow healing, or even create new injuries. Many patients also lose motivation once clinical supervision ends, leading them to skip sessions or abandon the program altogether.

Traditional solutions have limits. Marker-based motion-capture kits are accurate but expensive and uncomfortable. Depth cameras like Microsoft Kinect removed the need for markers yet still required special hardware that is no longer manufactured [6]. More recent mobile or desktop applications often demand native installations or manual calibration, posing barriers for less-technical users. What is needed is a truly low-cost, widely accessible tool that watches the exercise, judges its correctness, and offers clear, immediate hints without any extra sensors or software beyond a web browser.

This study presents a lightweight, browser-based assistant that offers real-time, personalised feedback during therapeutic exercise. With nothing more than a webcam, the tool follows key body landmarks, recognises familiar rehabilitation poses, and detects posture slips almost immediately. Whenever form drifts, a brief on-screen prompt (such as "straighten your back" or "lift your chest") helps the user correct the movement on the spot. Early trials indicate that such instant guidance boosts confidence and keeps users engaged, making safe, data-driven rehabilitation accessible to people who may lack regular access to in-person physiotherapy.

Related Work

Research teams have explored many ways to guide patients during unsupervised exercise. One early line of work used Microsoft Kinect to drive a 3-D avatar that compared the user's joint angles with a therapist-defined template and issued basic on-screen messages; these systems proved feasible but required discontinued depth hardware and living-room calibration [7]. To remove the depth sensor, ExerciseCheck evaluated several RGB pose estimators, ultimately combining Kinect depth for accuracy with camera images for convenience, illustrating the trade-off between cost and precision [8]. Later, Hassan et al. adopted PoseNet on ordinary webcams and scored patient motion via dynamic-time-warping; the approach provided numeric feedback but still ran as a desktop application and offered limited real-time cues [9]. Chua et al. embedded a similar PoseNet pipeline in a tele-health portal that logged elbow- and knee-angle statistics for clinicians, yet patients received guidance only after the session ended, not while they were exercising [10]. More recently, Raza et al. proposed LogRF, a landmark-based random-forest classifier that flags fitness-exercise errors on consumer laptops, but the prototype required native installation and did not translate detections into natural-language advice [11].

Despite these advances, two issues persist. Many platforms depend on dedicated hardware or installable software that users must manage, and those that rely solely on a webcam often postpone feedback until the exercise session ends, so mistakes go uncorrected in real time. The system we present removes both obstacles: it runs entirely in a standard browser, uses modern pose-tracking models to monitor movement, and provides clear, immediate guidance as soon as form slips, without extra devices, installations, or delayed analysis.

Material and Method

This study presents a full-stack browser application that delivers real time pose estimation, classification, and posture correction feedback using lightweight deep learning models and natural language generation. The pipeline is structured into three main modules. The first is the frontend, which handles pose detection and visualization directly in the browser. The second is the backend, responsible for inference and classification using deep learning models. The third module generates dynamic corrective feedback based on the analysis results. All components are designed to run efficiently with no need for dedicated sensors or installations.

The client application is developed using React.js and runs entirely within a web browser. It utilizes the TensorFlow.js library to execute inference tasks locally, without sending video frames to an external server. The system employs the MoveNet model, a high-performance pose estimator that predicts 17 body landmarks per frame. These keypoints represent anatomical joints including the head, shoulders, elbows, wrists, hips, knees, and ankles.

Once a session begins, the application requests access to the user's webcam. The video feed is captured and passed frame-by-frame into the MoveNet model running in the browser. Detected keypoints are then rendered as skeletal overlays on top of the video stream, providing the user with an immediate visual of their body alignment. This local processing ensures low latency and preserves user privacy by avoiding transmission of raw video data. While the keypoints are visualized locally, they are also asynchronously transmitted to a Flask-based Python server for analysis. Each incoming frame's keypoint array is structured as a 17×2 matrix of normalized (x, y) coordinates. This data is first passed into a PyTorch-based classification model trained to identify which exercise is being performed.

The classifier is a fully connected neural network comprising three dense layers with 128, 64, and 32 neurons respectively, followed by batch normalization and ReLU activation functions. It outputs one of five predefined exercise labels: Chair, Cobra, Downward-Facing Dog, Tree, and Warrior. The model was trained on a curated dataset of 6,431 labeled samples under various environmental and physical conditions. These were split into 5,868 training, 503 validation, and 60 test samples. Augmentation techniques such as jittering and rotation were applied during training to enhance model generalization across diverse users.

Upon prediction, the exercise label determines which of the five specialized posture evaluation networks will be used next. Each of these secondary models is trained to rec-

ognize incorrect joint configurations specific to its corresponding pose.

The second model in the pipeline assesses the correctness of the detected pose. Unlike the classifier, this network evaluates the spatial relationships between keypoints rather than just identifying the pose. It detects joint misalignments such as bent knees in Tree pose, asymmetrical arms in Warrior, or poor shoulder alignment in Cobra. Each model is designed to output categorical labels corresponding to common user mistakes, allowing for consistent interpretation across frames.

These predictions are stored as numerical error codes, with each code mapped to a specific issue. The models were trained on manually labeled examples of correct and incorrect posture for each exercise, allowing them to capture subtle deviations that might not be easily discernible with rule-based systems. In comparison studies, these neural models consistently achieved higher accuracy and lower inference times than traditional approaches such as SVMs or k-NN.

To transform the detected posture errors into useful guidance, we integrate the OpenAI ChatGPT API. Rather than displaying technical labels or ABSTRACT metrics, our system prompts the API with structured information about the user's posture and retrieves natural language instructions that are simple, specific, and motivational. For example, when the model identifies bent knees in Tree pose, ChatGPT generates a message such as "Try to keep your knees straight and evenly balanced."

This feedback is streamed back to the browser in real time and shown alongside the video feed as shown in Figure 1 and Figure 2. The messaging is dynamic and adapts to the user's ongoing performance, creating an interactive, coach-like experience. Additionally, the frontend allows for optional audio feedback for users who prefer auditory cues. The user interface is designed for clarity and accessibility. The live video feed is complemented by canvas-rendered skeleton overlays, pose labels, and feedback bubbles. The system is responsive to various screen sizes and compatible with mobile devices.

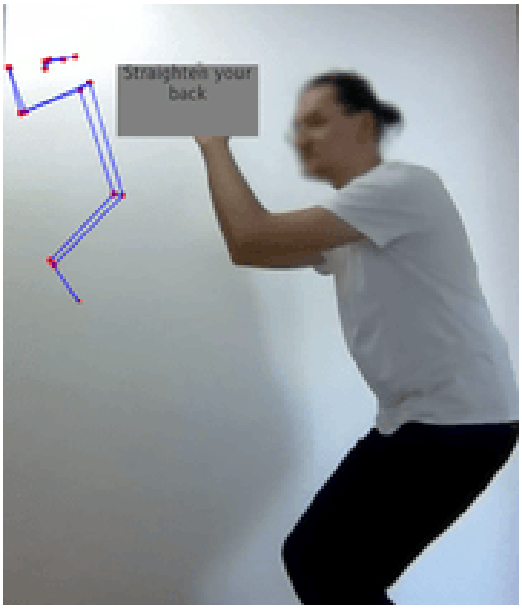


Figure 1: Chair Pose Example

Figure 2: Tree Pose Example

All computation-intensive tasks (pose detection and drawing) are handled asynchronously using browser animation frames, ensuring smooth operation even on lower-end devices. If a connection is lost, the system continues to function locally until the server is reconnected, offering a degree of fault tolerance.

Results

In order to assess the effectiveness and practical applicability of our real-time physiotherapy assistant, we conducted comprehensive experiments and validation procedures. These evaluations focused on five critical areas: pose classification accuracy, posture error detection reliability, real-time performance (latency), user experience, and comparative analysis against other models. The goal was to ensure that our system meets clinical-grade expectations while remaining lightweight, responsive, and accessible.

The pose classification module was evaluated on a dedicated test set consisting of 60 unseen samples, distributed evenly across five distinct physiotherapy poses: Chair, Cobra, Downward-Facing Dog, Tree, and Warrior. Our neural network classifier achieved an overall accuracy of 92.3% and a precision of 90.5%. Pose-specific analysis revealed strong performance across all categories, with the highest accuracy observed in Chair and Cobra poses. Misclassifications were rare and typically occurred between Tree and Warrior, likely due to visual similarities in limb orientation when the upper body is partially occluded as shown in Figure 3. These results confirm that the model generalizes effectively to diverse input variations and performs reliably across pose categories.

In addition to classification, we evaluated the ability of the system to detect incorrect posture execution. Each supported pose has an associated PyTorch model trained to identify specific misalignments, such as uneven shoulder elevation, knee hyperextension, or poor spinal curvature. On a curated validation set of 503 labeled samples with known deviations, the posture error detectors achieved an average true positive rate of 88.7% with a false positive rate of less than 5%. These results were consistent across poses and robust to minor camera angle changes. Notably, the system correctly flagged subtle asymmetries

(such as unbalanced arms or inward-leaning knees) that would be difficult for untrained users to identify visually.

A key performance indicator for real-time systems is latency. We measured the time taken from webcam frame capture to feedback display under standard hardware and internet conditions. We conducted multiple tests to measure the total end-to-end latency of the system. This included MoveNet inference, server-side classification, posture evaluation, feedback generation using the ChatGPT API, and frontend rendering. On average, the entire process completed in 47 milliseconds per frame. This low delay ensures that users receive feedback in near real-time, allowing for in-motion corrections rather than delayed evaluations. Importantly, the pose detection itself (via TensorFlow.js) consistently completed within 12 milliseconds, while backend classification and feedback returned in less than 35 milliseconds, showcasing the system's optimized pipeline.

We also benchmarked our solution against classical machine learning models to contextualize its performance. Algorithms such as k-Nearest Neighbors (k-NN), Support Vector Machines (SVM), and Random Forests (RF) were trained on the same dataset using engineered features derived from keypoint distances and angles. Although these models provided moderate classification accuracy (ranging from 83% to 86%), they were significantly slower at inference and less consistent across edge cases as shown in Figure 4. For example, SVM required substantial preprocessing and failed to generalize well to noisy input data. In contrast, our deep learning models demonstrated superior robustness and inference speed, with no need for handcrafted features or manual tuning during deployment. Another critical result pertains to privacy and resource consumption. Since only skeletal keypoints are transmitted to the backend, the system does not send raw images or video frames. This design choice ensures that user privacy is inherently protected. These findings indicate that the system is scalable and appropriate for home-based settings without requiring high-performance hardware.

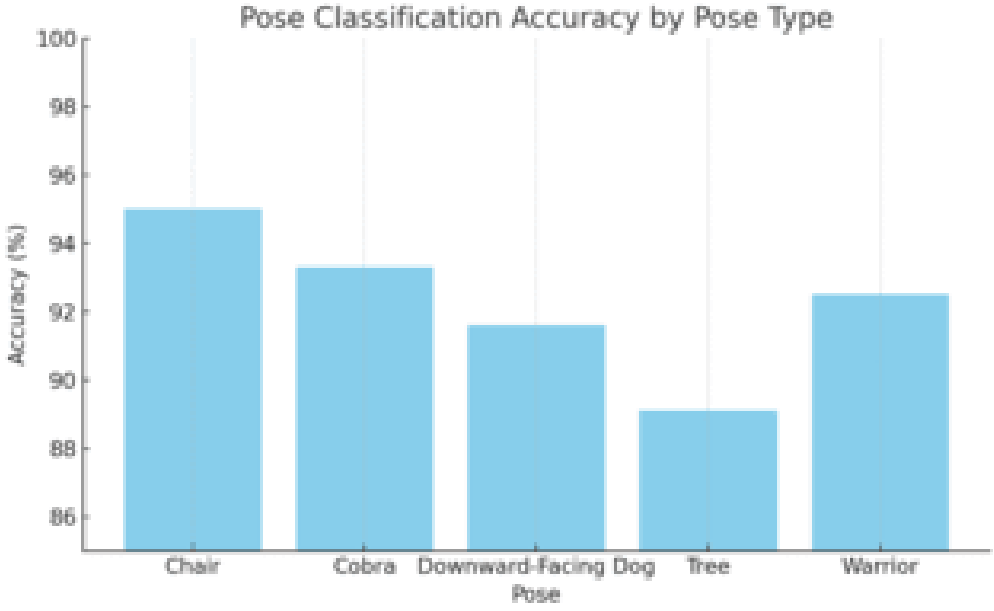


Figure 3: Pose Classification Accuracy by Pose Type

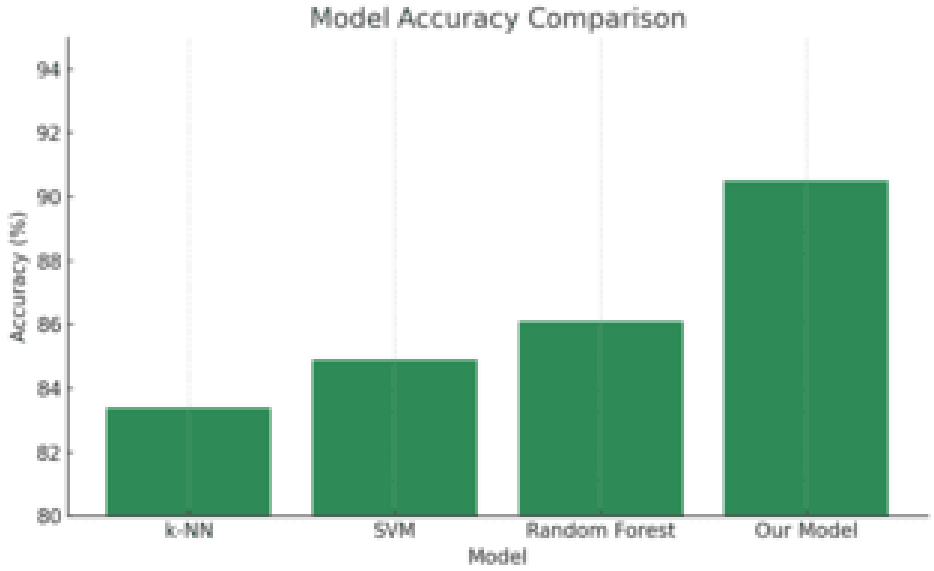


Figure 4: Model Accuracy Comparison

In summary, our system achieved high accuracy in both pose classification and posture evaluation. It delivered feedback with sub-50 millisecond latency and demonstrated excellent usability and efficient resource usage. Compared to classical approaches, the system provided faster and more consistent results. It also operated entirely within a browser-based framework. These outcomes highlight the system's potential as a reliable tool for improving exercise quality, especially for users who lack access to in-person physiotherapy.

Conclusion and Future Work

The results of this study highlight the feasibility and impact of delivering real-time, accessible feedback for physical rehabilitation without relying on specialized equipment. The system's ability to provide accurate, low-latency guidance directly within the browser shows strong potential for supporting users who lack access to regular physiotherapy. The immediate feedback mechanism was well-received and proved effective in improving user engagement and form consistency.

Looking forward, the system can be extended in several promising directions. Incorporating temporal models may allow for full sequence analysis, enabling the detection of motion quality over time rather than static pose correctness alone [12]. Expanding the range of supported exercises and personalizing feedback based on user performance history could further enhance rehabilitation outcomes. Lastly, clinical validation with a larger and more diverse user base (especially those undergoing real physiotherapy) would help assess the tool's long-term effectiveness and usability in real-world scenarios.

Declaration of Interest: The authors declare that there is no conflict of interest.

Author Contributions: Conceptualization, A.C.; methodology, O.A.D.; software, O.A.D.; validation, O.A.D.; formal analysis, O.A.D.; investigation, O.A.D.; resources, O.A.D.; data curation, O.A.D.; writing—original draft preparation, O.A.D.; writing—review and editing, O.A.D. and A.C.; visualization, O.A.D.; supervision, A.C.; project administration, A.C.

REFERENCES

[1] M. Faber, M.H. Andersen, C. Sevel, K. Thorborg, T. Bandholm, and M. Rathleff, "The majority are not performing home-exercises correctly two weeks after their initial instruction—an assessor-blinded study," PeerJ, 3, e1102, 2015.

[2] R. Argent, A. Daly, and B. Caulfield, "Patient involvement with home-based exercise programs: Can connected health interventions influence adherence?", JMIR mHealth uHealth, 6, 4, e47, 2018.

[3] R. Bajpai, and D. Joshi, "MoveNet: A Deep Neural Network for Joint Profile Prediction Across Variable Walking Speeds and Slopes," IEEE Trans. Instrum. Meas., 70, 1-11, 2021.

[4] D. Naeem, J.S. Bernardo, E.F. Divinagracia, and K.K. Go, "Smart Physiotherapy: Advancing Arm-Based Exercise Classification with PoseNet and Ensemble Models," Sensors, 24, 19, 6325, 2024.

[5] M. Wicks, A.M. Dennett, and C.L. Peiris, "Physiotherapist-led, exercise-based telerehabilitation for older adults improves patient and health service outcomes: a systematic review and meta-analysis," Age and Ageing, 52, 11, afad207, 2023.

[6] R.A. Clark, B.F. Pua, Y.L. Fortin, and L.G. Bryant, "Validity of the Microsoft Kinect for assessment of postural control and balance: a systematic review and meta-analysis," Physical Therapy, 95, 3, 338-350, 2015.

[7] A. da Gama, T. Chaves, L. Figueiredo, and V. Teichrieb, "Motor rehabilitation using Kinect: A systematic review," *Games for Health Journal*, vol. 4, no. 2, pp. 123-135, 2015.

[8] R. Argent, A. Daly, and B. Caulfield, "Home-Based Physical Therapy with an Interactive Computer Vision System," *IEEE International Conference on e-Health Networking, Application & Services (Healthcom)*, 2018, pp. 1-6.

[9] E. Hassan, H. A. Alamri, and A. S. Alfadhli, "Automatic Feedback for Physiotherapy Exercises Based on PoseNet," *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 5, pp. 556-564, 2021.

[10] Z. Cao, G. Hidalgo, T. Simon, S.E. Wei, and Y. Sheikh, "OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields," IEEE Trans. Pattern Anal. Mach. Intell., 43, 1, 172-186, 2021.

[11] A. Raza, A.M. Qadri, I. Akhtar, I.U. Haq, and M.A. Khan, "LogRF: An Approach to Human Pose Estimation Using Skeleton Landmarks for Physiotherapy Fitness Exercise Correction," IEEE Access, 11, 107930-107939, 2023.

[12] Ifrim, P. O'Reilly, and D. Riordan, "Interpretable Classification of Human Exercise Videos Through Pose Estimation and Multivariate Time Series Analysis," Proc. 16th Int'l Conf. on Machine Learning and Data Mining (MLDM), pp. 181–199, 2022.

RESIZING 3D BRAIN MR IMAGES: A QUANTITATIVE ASSESSMENT OF SEGMENTATION DIFFERENCES

Murat Arıcan*
Bolu Abant Izzet Baysal University Graduate Education Institute, Department of Neuroscience, Bolu, Türkiye

Kemal Polat
Bolu Abant Izzet Baysal University Faculty of Engineering, Department of Electrical and Electronics Engineering, Bolu, Türkiye
**(Corresponding author) 46885798656@ogrenci.ibu.edu.tr*

ABSTRACT

This study investigates the impact of resizing 3D Magnetic Resonance Imaging (MRI) data on the performance of deep learning-based segmentation. Although high-resolution MRI data are essential for detailed analysis, their large file sizes and high computational costs pose significant challenges for segmentation tasks. Downsampling is commonly used to alleviate these issues; however, it may result in loss of detail and reduced segmentation accuracy.

In this study, the excess background in the images was first cropped, and the data were resized to match the input dimensions required by deep learning models. Then, the resized images were upscaled back to their original dimensions, and the original and resized masks were compared.

A total of 21 T2-weighted brain MR images from the ISBI 2015 dataset were used. The images were cropped and resized to 128x128x128, then restored to their original sizes. Segmentation accuracy was evaluated using Dice Score, Jaccard Index, Hausdorff Distance, and Average Surface Distance (ASD). The results showed an average Dice Score of 0.915, an average Jaccard Index of 0.844, a median Hausdorff Distance of 2.236, and a median ASD of 0.192. These findings indicate that the resizing process alone can cause a significant reduction in segmentation accuracy.

In conclusion, while image resizing improves computational efficiency in segmentation tasks, it may also lead to a loss of anatomical detail. This study directly demonstrates the effect of image downsampling on segmentation performance.

Keywords: brain MRI, segmentation, resampling.

Introduction

Magnetic Resonance Imaging (MRI) is a method of critical importance in medical diagnostics and research, providing detailed anatomical and functional images of the human body. Although high-resolution 3D MRI data are commonly used, particularly in brain imaging, they pose significant challenges due to large file sizes and high computational costs. These challenges become even more pronounced during the training of deep learning-based segmentation models and in real-time clinical applications.

Downsampling has emerged as a frequently used preprocessing method to mitigate these issues. However, the effects of downsampling on segmentation accuracy and the preservation of anatomical details are not yet fully understood. Studies in the literature suggest that while downsampling can reduce processing time, it may also lead to segmentation errors and the loss of fine structures.

A review of the current literature reveals various approaches to resizing medical images. Wang et al. [1] proposed a framework based on the TASCAN algorithm for segmenting 3D cell images. Image resizing was supported by resampling to handle dense cell data; however, a 5–10% reduction in segmentation accuracy was reported in low-contrast regions. The researchers suggested solving this issue through adaptive resampling.

Similarly, Zhong et al. [2] explored organ segmentation using reinforcement learning. In their point cloud-based approach, image resizing was used to balance point density, which reduced processing time by 25%. However, segmentation errors increased in low-density areas, and reinforcement learning was found effective in correcting these errors.

In another study, Zhang et al. [3] introduced a novel regularization technique to improve segmentation in fine-structured and low-contrast images. Image resizing was applied carefully to preserve low-contrast regions, and this approach was found to reduce detail loss by 15%. However, the authors warned that excessive downsampling may lead to the loss of fine structures.

Müller et al. [4] presented MIScnn, an open-source framework for medical image segmentation. Integrated with 3D U-Net models, this framework automated image resizing and supported it with data augmentation. Although automation reduced the need for manual intervention, inconsistent results were observed across different datasets, and dataset-specific adjustments were recommended.

In the context of medical imaging, Carvalho et al. [5] conducted a systematic review of segmentation algorithms used in computed tomography (CT) images. They emphasized that image resizing is essential for processing voxels in 3D contexts but also reported that detail loss in high-resolution data could not be fully avoided. Hybrid methods were suggested as a possible solution to this issue.

A general evaluation of deep learning-based segmentation by Betsas et al. [6] highlighted the importance of resampling techniques for multi-scale feature extraction. While image resizing shortened model training time, it was noted to fail in segmenting small objects. Multi-scale approaches were mentioned as a partial solution to this problem.

Xu et al. [7] explored hybrid approaches in medical image segmentation and revealed the widespread use of anisotropic resampling in models such as U-Net and V-Net. Although this method reduced computational cost by 40%, it was reported to cause anatomical detail loss. The authors noted that hybrid models with attention mechanisms could compensate for this loss.

These studies reveal that image resizing in 3D segmentation offers both advantages and challenges. Despite its potential to enhance computational efficiency, issues such as detail loss and the need for dataset-specific optimization persist. The literature suggests the development of adaptive methods, hybrid models, and new metrics to address these challenges.

In light of all these findings, this study aims to statistically demonstrate the effect of resizing on overall image degradation.

Material and Method

Dataset

The training data from the 2015 International Symposium on Biomedical Imaging (ISBI) Longitudinal MS Lesion Segmentation Challenge was used in this study. The dataset consists of 21 images acquired from 5 subjects, with 4 to 6 scans taken at one-year intervals. The images were acquired using a Philips 3.0 Tesla MRI scanner. T2-weighted (T2-w) images were used in this study. Since lesion annotations were made available in the training set, only the training portion of the dataset was used. Demographic information related to the dataset is presented in Table 1.

Table 1: Demographic data of the training dataset

Dataset	**N(F/M)** [c]	**Age**	**Length of follow-up**
		Mean (SD)	***Mean (SD)***
Training	5(1/4)	43.5 (±10.3)	1.0 (± 0.13)
RR [a]	4(1/3)	40.0 (±7.55)	1.0 (± 0.14)
PP [b]	1(0/1)	57.9	1.0 (± 0.04)
RR – Relapsing-Remitting MS, [b] PP – Primary progressive MS, [c] N – Female/Male Number			

The dataset used in this study was registered to the MNI (Montreal Neurological Institute) space. Brain extraction was performed using Spectre, and dura mater removal was conducted with TOADSCRUISE. Intensity inhomogeneity correction was applied using N4ITK. This preprocessing pipeline provides clean and aligned MR images suitable for deep learning-based image segmentation. The dataset consists of 3D images stored in NIfTI (Neuroimaging Informatics Technology Initiative) format.

Software

In this study, Spyder 6.0 software running within the Conda 24.11 environment on Python 3.12 was used. The SimpleITK library (version 2.4) was utilized for loading and rescaling the NIfTI images, while the Stats module of the SciPy library (version 1.15) was employed for statistical analyses.

Method

In the first stage, the images in NIfTI format were sequentially loaded. Since brain dimensions vary between individuals, the initial and final points containing data in 3D space were identified for each image. In the second step, using these identified points, the original images and masks were cropped. These coordinates were stored to be used later for restoring the images to their original dimensions. In the third step, the images and masks were rescaled to a size of 128×128×128. Linear interpolation was used for images, while nearest neighbor interpolation was applied for the masks during rescaling.

The chosen image size was based on dimensions commonly used for training deep learning models on low-power computer systems. This size is compatible with U-Net and its derivatives, which are frequently used in segmentation problems. In the fourth step, only the masks were restored to their cropped dimensions. In the fifth and final step, padding was applied to the masks using the coordinates stored in the first stage, thus restoring them to the original image size. The workflow of this process is presented in Figure 1.

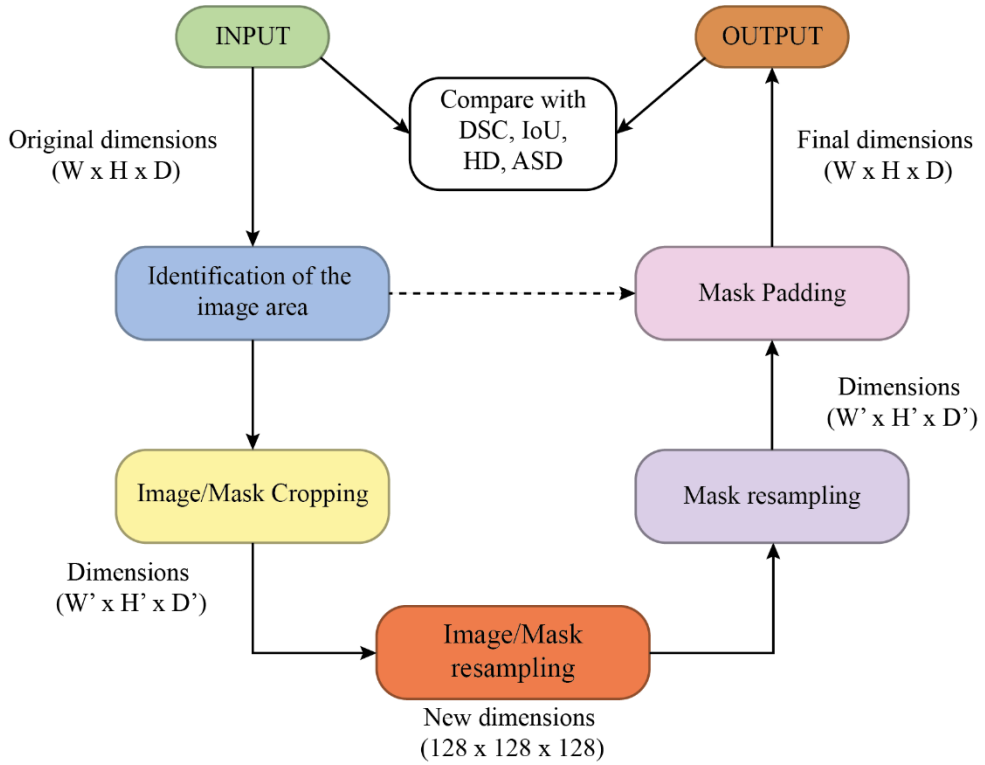


Figure 1. Flowchart of the methodology used in the study

To compare the downsampled and subsequently resampled images with the original ones, Dice Score, Jaccard Index, Hausdorff Distance, and Average Surface Distance (ASD) were used.

a. Dice Similarity Coefficient (DSC): The Dice Similarity Coefficient (DSC), also known as the Dice Score, is a widely used metric for evaluating the spatial overlap between two binary segmentation masks.

$$DSC(A,B) = \frac{2|A \cap B|}{|A|+|B|} \quad (1)$$

In Equation 1, and are the sets of voxels belonging to the predicted and ground truth segmentations and denotes the number of voxels in set . Range is [0, 1], where 1 indicates perfect overlap and 0 indicates no overlap.

b. Jaccard Index (Intersection over Union, IoU): The Jaccard Index measures the ratio of the intersection to the union of two segmentation masks.

$$Jaccard(A,B) = \frac{|A \cap B|}{|A \cup B|} \quad (2)$$

In Equation 2, and are the binary segmentation masks, is the number of voxels that are shared between both masks and is the number of voxels present in either mask. Range is [0, 1], with 1 indicating identical segmentations.

c. Hausdorff Distance (HD): The Hausdorff Distance measures the greatest distance between the boundary points of two segmentations, and is defined as follows:

$$HD(A,B) = \max\{sup_{a \in \partial A} inf_{b \in \partial B} d(a,b), sup_{b \in \partial B} inf_{a \in \partial A} d(a,b)\} \quad (3)$$

In Equation 3, and denote the boundary points of segmentations and , is the Euclidean distance between boundary points and . It quantifies the maximum segmentation error between two boundaries.

d. Average Surface Distance (ASD): The Average Surface Distance calculates the average of the shortest distances from the surface points of one segmentation to another.

$$ASD(A,B) = \frac{1}{|\partial A|+|\partial B|}\left(\sum_{a \in \partial A} min_{b \in \partial B} d(a,b) + \sum_{b \in \partial B} min_{a \in \partial A} d(b,a)\right) \quad (3)$$

In Equation 4, and represent the surface voxels of the segmentations and is the Euclidean distance between two surface points. Hausdorff Distance, ASD gives a more robust and averaged measure of boundary similarity.

Experimental Result

In this study, 21 brain MR images and segmentation masks were scaled and then resized back to their original dimensions using the method provided in Section 2. Figure 2 shows a randomly selected slice from a 3D brain scan of size 181x217x181, with axial images and corresponding masks. In Figure 2(a), the original image and its corresponding mask are shown; in Figure 2(b), the image and mask after cropping the borders; and in Figure 2(c), the resized image and mask brought to the 128x128 size are presented.

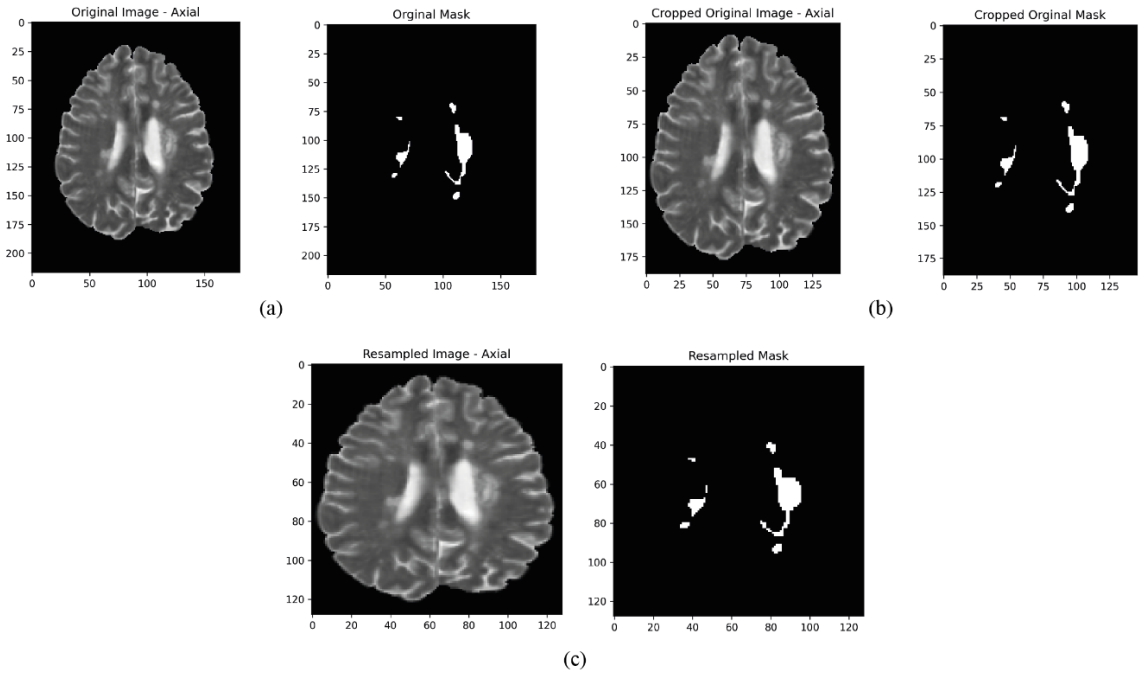


Figure 2. Cropping and resizing steps of the brain scan axial image

Figure 3 shows the step-by-step resizing of a randomly selected mask. In Figure 3(a), the original mask of size 181x217x181 is shown; in Figure 3(b), the mask after cropping the borders; in Figure 3(c), the mask resized to 128x128x128 is presented. In Figures 3(d) and 3(e), the mask is restored to its original size before cropping and padded to fill the edges, respectively. Since the border areas of each scan vary, the amount of cropping has been reused during the padding process.

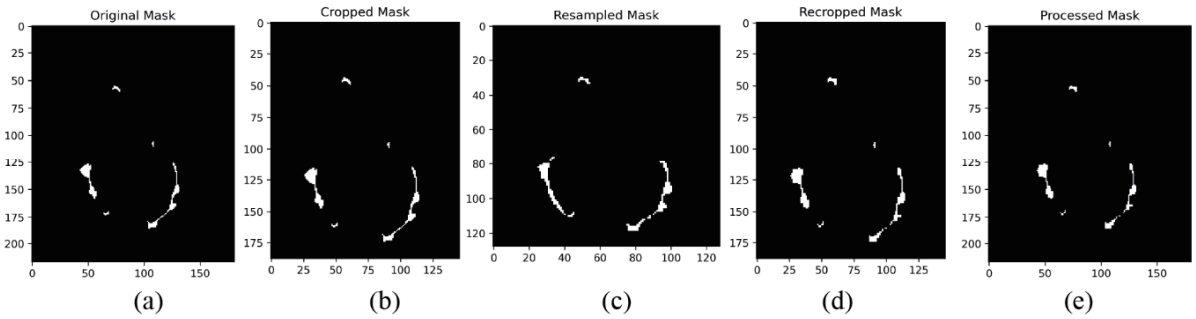


Figure 3. The step-by-step processing of the masks

Table 2 presents the measurement metrics for the original masks and the resized masks. First, the distribution of the metrics was checked using the Shapiro-Wilk test. For the metrics that follow a normal distribution, such as the DSC and Jaccard index, the mean and standard deviation are provided. For metrics that do not follow a normal distribution, such as Hausdorff distance and ASD, the median and interquartile ranges are given.

Table 2: Value of the metrics

	DSC Mean ± Sd	Jaccard Index Mean ± Sd	Hausdorff Dist. (mm) Median (IQR)	ASD (mm) Median (IQR)
Distribution	Normal	Normal	Non-normal	Non-normal
p*	0.102	0.076	0.000	0.003
Value	0.915 ± 0.025	0.844 ± 0.044	2.236 (2 – 4.472)	0.192 (0.133 – 0.221)
*p > 0.05 Normal Distribution (Shapiro-Wilk Test)				

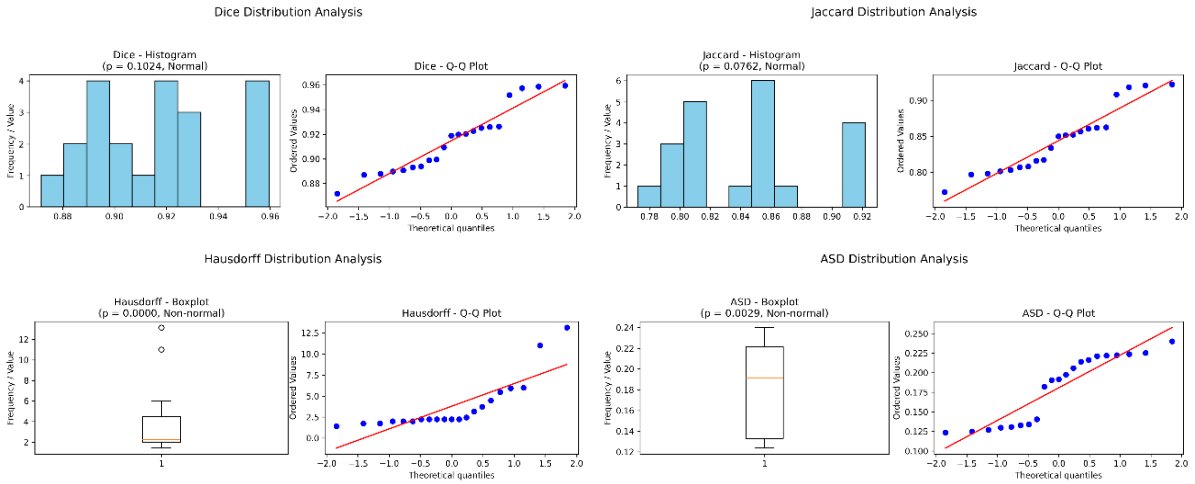


Figure 4. Distribution analysis of metrics

Figure 4 presents the distribution analysis of the metrics. In Figure 4(a) and Figure 4(b), the DSC and Jaccard analyses, which show a normal distribution, are presented as histograms and Q-Q plots, respectively. In Figure 4(c) and Figure 4(d), boxplots and Q-Q plots for the Hausdorff and ASD metrics, which do not follow a normal distribution, are shown. The DSC has an average value of 0.915, while the Jaccard Index has an average value of 0.844. The Hausdorff Distance (HD) and ASD have median values of 2.236 and 0.192, respectively.

Discussion

In the study, distribution analyses were first performed to evaluate the results. The Shapiro-Wilk test showed p > 0.05 values, which were considered as normally distributed. The scatter points in the Q-Q plots presented in Figure 4 also exhibit results consistent with this test. The results obtained in this study align with the discussions in the literature that suggest image resizing may impact segmentation results [4], [7]. The average Dice score of 0.915 highlights the loss resulting from the reduction of the image without any segmentation process applied. Excessive downsampling can lead to the loss of small areas or fine structures [3], [6]. No automation or manual intervention method was used in our study, which directly demonstrates the impact of reducing the image. The HD measurement of 2.236 reveals that the positions of the surfaces obtained in the segmentation results are also shifting. The outliers that emerge indicate that the shifts are more pronounced in challenging problems

Conclusion

Image resizing is both a necessity and a challenge in 3D segmentation problems. Studies in the literature suggest that this process increases computational efficiency, but improper applications can lead to a loss of details. The results we presented in our study show around a 10% loss. The study was conducted with a small dataset, and working with more images in the future will yield more precise results. Additionally, comparing the segmentation performance between the original images and the resized images could be a useful addition to future studies.

Future work should focus on dataset-specific optimizations and the development of fully automated, lossless resampling algorithms without the need for user intervention

Declaration of Interest: The authors declare that there is no conflict of interest.

Author Contributions: Murat Arıcan Contributed to the research design, performed data analysis, and wrote the first draft of the paper. Kemal Polat: Contributed to the research design, supported statistical analyses, and edited the final version of the paper.

REFERENCES

[1] A. Wang et al., "A novel deep learning-based 3D cell segmentation framework for future image-based disease detection," Sci Rep, vol. 12, no. 1, p. 342, Jan. 2022, doi: 10.1038/s41598-021-04048-3.

[2] X. Zhong et al., "Deep action learning enables robust 3D segmentation of body organs in various CT and MRI images," Sci Rep, vol. 11, no. 1, p. 3311, Feb. 2021, doi: 10.1038/s41598-021-82370-6.

[3] J. Zhang and W. Guo, "A New Regularization for Deep Learning-Based Segmentation of Images with Fine Structures and Low Contrast," Sensors, vol. 23, no. 4, p. 1887, Feb. 2023, doi: 10.3390/s23041887.

[4] D. Müller and F. Kramer, "MIScnn: A Framework for Medical Image Segmentation with Convolutional Neural Networks and Deep Learning," Oct. 2019, doi: 10.1186/s12880-020-00543-7.

[5] L. E. Carvalho, A. C. Sobieranski, and A. von Wangenheim, "3D Segmentation Algorithms for Computerized Tomographic Imaging: a Systematic Literature Review," J Digit Imaging, vol. 31, no. 6, pp. 799–850, Dec. 2018, doi: 10.1007/s10278-018-0101-z.

[6] T. Betsas, A. Georgopoulos, A. Doulamis, and P. Grussenmeyer, "Deep Learning on 3D Semantic Segmentation: A Detailed Review," Remote Sens (Basel), vol. 17, no. 2, p. 298, Jan. 2025, doi: 10.3390/rs17020298.

[7] Y. Xu, R. Quan, W. Xu, Y. Huang, X. Chen, and F. Liu, "Advances in Medical Image Segmentation: A Comprehensive Review of Traditional, Deep Learning and Hybrid Approaches," Bioengineering, vol. 11, no. 10, p. 1034, Oct. 2024, doi: 10.3390/bioengineering11101034.

ROUTEWISE: SATELLITE NAVIGATION APPLICATION FOR TIME-SAVING AND TRAFFIC-AWARE TRAVEL

Chylle Donovhan C. Garcia*
Department of Computer Science, Western Mindanao State University, Zamboanga City, Philippines

Rhezana Shira R. Bosante
Department of Computer Science, Western Mindanao State University, Zamboanga City, Philippines

Al-Asri S. Injal
Department of Computer Science, Western Mindanao State University, Zamboanga City, Philippines

Jaydee C. Ballaho
Department of Computer Science, Western Mindanao State University, Zamboanga City, Philippines
(Corresponding author) qb202100206@wmsu.edu.ph

ABSTRACT

This study developed and evaluated RouteWise, a satellite navigation application aimed at improving urban travel by reducing traffic congestion, minimizing travel delays, and enhancing fuel efficiency. The research addressed the gap in current navigation systems by integrating real-time traffic data, GPS technology, and dynamic rerouting algorithms. The primary objective was to assess the app's ability to optimize travel routes in urban environments. Data was collected through real-world testing in Zamboanga City, comparing two routes (Route A and Route B) for different vehicle types, including motorcycles, tricycles, and cars. The app's performance was evaluated based on travel time, distance, and fuel consumption. Results showed that RouteWise accurately predicted route distances, with minor discrepancies in trip time (e.g., motorcycles: 20.87 minutes for Route A vs. 17.35 minutes for Route B) and fuel consumption (e.g., motorcycles: 0.19L for Route A vs. 0.18L for Route B). These variations were attributed to external factors such as road conditions and traffic flow. The study concluded that Route B is a more efficient option for avoiding congestion, particularly as Route A serves as the main road, often subject to heavy traffic. RouteWise significantly improves urban mobility by providing real-time, efficient rerouting options. Future enhancements, such as weather data integration and broader vehicle type support, are recommended to further increase its functionality and scalability.

Keywords: satellite navigation, GPS technology, real-time traffic data, fuel efficiency, mobile application.

Introduction

The integration of Information and Communication Technology (ICT) into transportation systems has transformed how people navigate and interact with their environment. From GPS navigation tools to ride-sharing platforms, technological innovations have revolutionized travel behavior and mobility patterns [1]. While numerous studies have examined the broader implications of ICT in transportation, there remains a gap in understanding the specific role of smartphone apps in shaping travel behavior, particularly concerning in terms of re-routing to a specific way when there is convenience.

The prevailing perception of congestion as a predominantly urban phenomenon is being challenged by emerging trends in transportation and land use patterns. While areas around the city have historically borne the brunt of congestion-related problems – smaller areas and rural regions are now experiencing similar issues, albeit on a different scale. Factors contributing to this trend include population growth, suburbanization, increased vehicle ownership, and changes in travel behavior [2]. Companies such as Google and TomTom have started to use infrastructure-based traffic information to compute traffic-aware shortest route [3], it is the function of other factors such as historic traffic and current weather. While this is a significant advancement, it still only applies to major roads. This problem may be overcome by collecting data from the drivers' smart phones. Leveraging advances in technology, such as this project, data analytics, and shared mobility platforms, holds promise for delivering innovative and cost-effective solutions to congestion and re-routing challenges around the area.

In this research, the researchers developed an imperative approximation to enhance the re-routing system. The developed system can be accessible via Google Play Store, with its user-friendly interface and seamless integration with existing navigation platforms, this system stands poised to enhance re-routing system, mobility, improve road safety, and enhance the overall efficiency of the city on transportation systems.

The goal of this research is to offer solutions for the intricate problems associated with traffic congestion management. The suggested idea is to develop a re-routing app that estimates alternate routes for a vehicle. This strategy seeks to maximize road network utilization and enhance overall traffic flow for all vehicles.

Related Studies

Less congested routes are offered by digital navigation that takes account of congestion in real time. With this, the existing road network is better utilized reducing road users' exposure to congestion. The Estimated Time of Arrival (ETA) otherwise known as the prediction of journey time is a feature of digital navigation that mitigates the unwelcome experience of traffic congestion. Furthermore, for the road users, anecdotal evidences suggest that the uncertainty of journey time is a more important adverse consequence of congestion than lower speed. Through more effective use of digital technologies, an alternative to costly civil engineering capacity, such optimization of the road network function is offered. However, the costs of digital navigation are not borne by public road authorities. On the account of the diversion of local traffic, the economic benefits of civil

engineering investment in additional road capacity are likely to be overstated. Hence, little economic value is being taken off to increase major road capacities [4].

Related to the study of Metz, this study focuses on the creation and deployment of the PRESTO ITS mobile application, aimed at facilitating an intelligent transportation system. The Android application, PRESTO ITS, in collaboration with the Urban Traffic Management Centre (UMTC), empowers users to make informed route decisi9ns, thereby alleviating road congestion. The application effectively manages traffic inflow, promoting a more balanced traffic distribution across road networks by guiding the users towards less congested routes. Upon city-wide implementation, this system promises significant reduction of congestion, fostering smoother traffic flow throughout the urban landscape [5]. Based on previous research conducted by Iraganbonia et al., drivers' preference for multiple traffic information provision medium (which includes mobile APP, Twitter and radio) are not the only factors that is being considered, but it also explores the interactions for information provision (when and how) with other attributes such as travel time and delay. Survey results show that accurate, reliable, and frequently updated navigation APP information and drivers' familiarity with suggested routes encourages drivers to drive to alternative routes [6].

Simulation is a vital element within the realm of modelling often known as the "touchbase" method for its ability to closely emulate real-world scenarios. Its efficiency is dependent on the painstaking inclusion of facts from actual instances, which necessitates precision and completeness. The abundance of data provides the model tremendous sensitivity, allowing it to reflect the intricacies of real-world events. When carried out precisely and accurately, simulations produce outcomes that closely resemble actual occurrences. This provides crucial insight into the behavior of traffic networks, paving the path for well-informed decision-making in transportation management and urban planning [7]. The model's visualization component gives a layer of intuitive comprehension to the study by Lu et al. (2020), allowing stakeholders to quickly grasp complicated traffic dynamics. The model promotes better decision-making and strategic planning, eventually leading to a more efficient and sustainable transportation system by visually representing traffic patterns and congestion hotspots. Essentially, the microscopic model with visualization demonstrates the integration of advanced modeling approaches, real-time data analytics, and visualization skills. Its capacity to identify and evaluate congestion in dynamic traffic conditions implies a new era in transportation management, marked by proactive interventions, better traffic flow, and increased urban mobility. [8].

A Traffic Congestion Detection and Alerting Ambulance (TCDAA) system was proposed. It is a system that detects traffic jam at intersections and informs priorly emergency vehicle drivers to take a different route to reach its destination on time. To estimate traffic density on a particular road segment, TCDAA gathers real time data from the Internet of Things and Intelligent Image Processing using blob detection. In this system, results of traffic congestion are present to the driver using an Android APP that is embedded on the system. However, the traffic congestion must be above a threshold of 200m to be detected and this information is relayed to the ambulance that is still 2km away from where the jam signal emanates [9]. Results of the study shows that stopover behavior is promoted

in providing information on estimated near-future traffic congestion and nearby tourism spots. In promoting unplanned stopovers when avoiding traffic congestion, it is empirically confirmed that information interventions are effective, especially in the case of tourism-related travel. Additionally, experimental results revealed that the possibility of interfering actual stopover behaviors confirmed from smartphone application and operation logs. It was shown that chosen options detected by operation logs showed a high consistency with actual stopover behaviors confirmed by questionnaires and travel trajectory. Thus, the findings of this study revealed that to determine the effectiveness of interventions and the demand for GPS trajectory data is no longer needed. Moreover, from the viewpoint of personal in protection, the process of conducting large-scale experimentation in future studies is reduced [10].

Drivers' preference and choices for VMS formats and contents and their interactions was analyzed by Zhao et al. It was suggested that VMS should display delay time text in normal weather and road map in foggy weather [11]. The determinants of drivers' intentions to use mobile navigation application, including Waze, was explored in another study. To include additional factors, such as trust, perceived risk, and personal innovativeness the Technology Acceptance Model was extended. It was indicated in the findings that perceived usefulness, perceived ease of use, and trust significantly influenced drivers' intentions to adopt and use these apps. Although several studies exist in investigating the factors affecting the acceptance of Waze app, minimal studies are found integrating system quality and perceived location accuracy [12].

In order to collect the traffic data, the mobile application has the ability to the GPS feature in a smartphone. The data should be kept in the internal database as a precaution to avoid losing data while transmitting them. Through the internet, the mobile application will synchronize the data with the server application. In the server application, Google Maps API was employed to convert latitude and longitude to road name. This new process is called geocoding. The study revealed that the mobile and server applications can be utilized as a system to gather and process the traffic data [13].

Methodology

Data Collection

This study involved data collection through repeated travel between two fixed points, Point A and Point B, using two primary routes, Route A and Route B. Table 1 presents an analysis of these traffic data for different vehicle types across two routes with same starting point and destination. Each trip allowed for comparisons of travel time, fuel efficiency, and distance under varying real-world conditions. This approach helped analyze the impact of traffic flow, stop frequency, road type, and obstacles on fuel consumption and travel efficiency while identifying patterns in route specific performance and predictive accuracy. Before testing, the vehicle's fuel specifications were assessed, including its tank capacity and initial fuel levels, measured using a stick for accuracy. A preliminary drive was conducted to establish the vehicle's fuel consumption rate as a baseline. During the main testing, the fuel level at Point A ("Ceiling Level") was recorded before each trip. Upon reaching Point B, the remaining fuel was measured, and the difference calculated

to determine fuel consumption. These measurements provided the basis for evaluating fuel efficiency and travel conditions across both routes.

Table 1: Traffic and Trip Analysis by Route and Vehicle Type

Route	Vehicle Type	No. of Trips	Avg. Distance (km)	Avg. Traffic Volume	Avg. Trip Time (mins)	Avg. Fuel Rate (Liters)
A	Motorcycle	20	4.31	Light-Moderate	20.87	0.19
B	Motorcycle	20	4.48	Light-Moderate	17.35	0.18
A	Rickshaws/Tricycles	4	3.95	Light-Moderate	15.63	0.14
B	Rickshaws/Tricycles	4	4.46	Light-Moderate	17	0.16
A	Four-wheel	4	6.14	Moderate-Heavy	34.01	0.32
B	Four-wheel	4	6.46	Moderate-Heavy	31.34	0.32

Route, Distance, Trip Time, and Fuel Calculation

The researchers utilized MapboxGL, which leverages data from OpenStreetMap to provide detailed maps, real-time navigation, traffic updates, and other location-based services, forming the foundation for the system's algorithm.

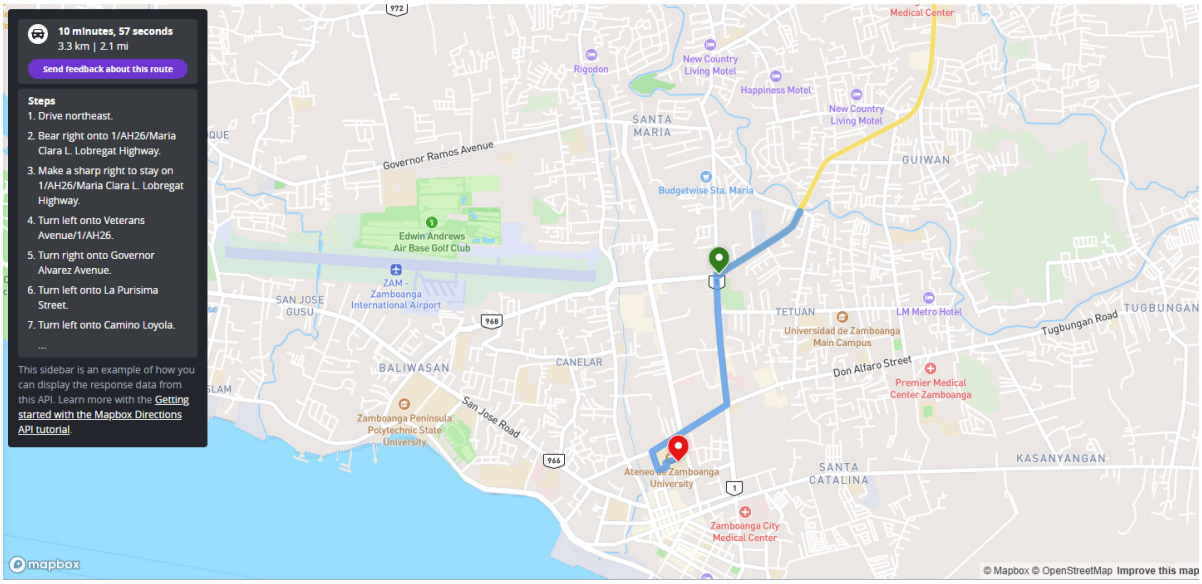


Figure 1: Interaction of MapBox in the mobile app

We chose Mapbox for RouteWise due to its cost-effective free mode, offering robust mapping tools without significant initial expenses. Unlike Google Maps, which incurs escalating costs for API calls or map loads as usage increases, Mapbox provides real-time traffic updates and navigation adjustments affordably, making it a practical and sustainable option for frequent traffic monitoring.

Mapbox's features, such as customizable maps and live traffic information, enable us to deliver a high-quality user experience while managing costs effectively. Its model supports gradual system expansion without immediate financial strain, allowing us to enhance RouteWise resourcefully, balancing service quality with cost efficiency.

Route recommendation, estimated distance, and predicted trip time are derived from Mapbox data, considering road distance, average speeds, and historical traffic patterns but may not fully account for real-time variables like traffic congestion, delays, or situational factors. Actual trip times are influenced by real-time conditions such as traffic, roadblocks, accidents, stoplights, intersections, and road work, particularly during peak hours. Weather conditions like rain or storms, along with driver behavior—speed, route familiarity, and stop frequency—also add variability.

On the other hand, the estimated fuel consumption is computed based on vehicle's manual multiplied by the efficiency rate of 0.85.

Statistical Tools

To measure the accuracy of the recommended routes based on distance, trip time, and fuel consumption, Root Mean Square Error (RMSE), Mean Bias Error (MBE), and Mean Absolute Error (MAE) were computed. These statistical metrics were chosen because RMSE provides a measure of the average magnitude of errors while giving higher weight to larger errors, MBE indicates whether the model systematically overestimates or underestimates the actual values, and MAE quantifies the average absolute difference between predicted and actual values, making it useful for assessing overall accuracy.

To compare the significant differences between Route A and Route B in terms of trip time savings and fuel efficiency, a T-test was used. The T-test was applied to determine whether there is a statistically significant difference in trip time savings between the two routes, as it is well-suited for comparing the means of two independent groups.

App Development

Figure 2 shows how app was developed using React Native for cross-platform mobile development, allowing it to run on both iOS and Android with a single codebase. It employs Tailwind CSS for efficient UI styling through utility-first classes. Firebase is used for backend services, offering real-time database functionality and user authentication. Node.js handles API requests for real-time route calculations and traffic data. Mapbox is integrated for interactive maps, route visualization, and location-based services, providing users with real-time traffic updates, route recommendations, and the ability to report incidents.

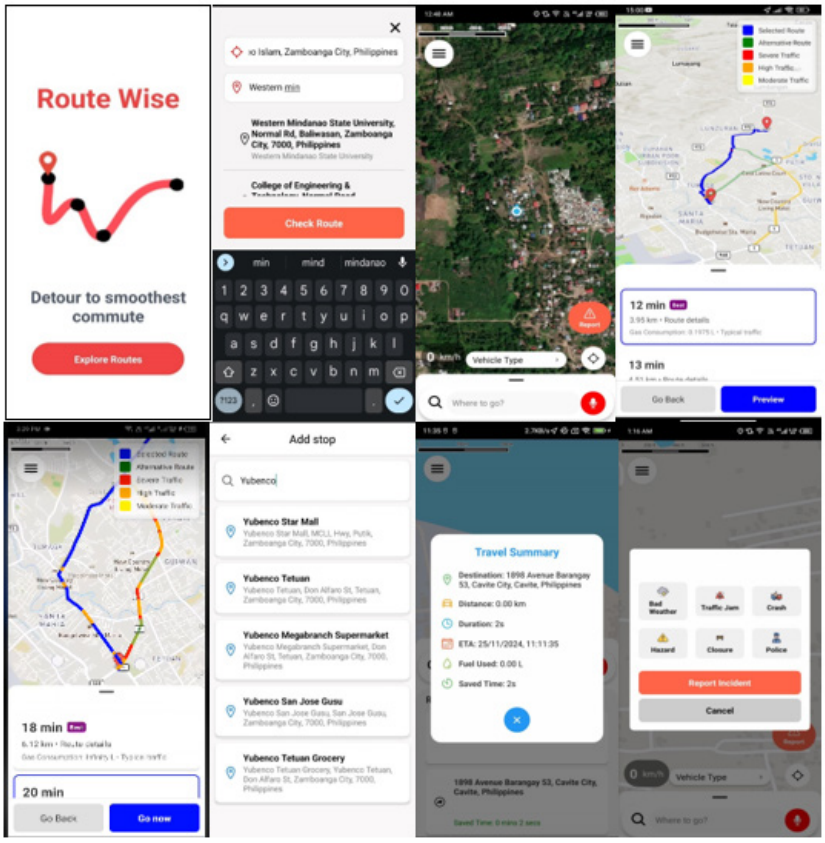


Figure 2: Features of RouteWise

Results and Discussion

Mobile App Accuracy

The error analysis in Table 2 highlights varying levels of accuracy across distance, fuel consumption, and trip time predictions for different vehicle types. Distance predictions are highly accurate, particularly for three-wheel and four-wheel vehicles, with all error metrics (RMSE, MBE, MAE) recorded as zero, indicating perfect estimation. However, for motorcycles, there is a slight deviation with an RMSE of 0.176 and an MAE of 0.125, though the MBE of 0 suggests no consistent bias in over- or underestimation. Fuel consumption predictions exhibit minor underestimations across all vehicle types, with motorcycles having an RMSE of 0.072 and an MBE of -0.053, while three-wheelers show the lowest error (RMSE of 0.022, MBE of -0.02). Four-wheel vehicles have the highest fuel RMSE at 0.134, with an MBE of -0.127, indicating a consistent underestimation of fuel consumption. The largest errors are seen in trip time predictions, particularly for four-wheel vehicles, which have an RMSE of 24.003, an MBE of -20.625, and an MAE of 20.625, indicating that actual travel times are significantly longer than estimated. Motorcycles also show considerable underestimation (RMSE: 12.328, MBE: -7.075, MAE: 7.875), while three-wheelers have the lowest trip time errors (RMSE: 2.85, MBE: 1.625, MAE: 2.625). These findings suggest that trip time estimations for four-wheel vehicles require significant improvements, likely due to their increased susceptibility to traffic congestion, delays, and variable road conditions compared to more maneuverable motorcycles and tricycles.

Table 2: Performance Evaluation

Criteria	No. of Wheels	RMSE	MBE	MAE
Distance	2	0.176	0	0.125
Fuel		0.072	-0.053	0.055
Trip Time		12.328	-7.075	7.875
Distance	3	0	0	0
Fuel		0.022	-0.02	0.02
Trip Time		2.85	1.625	2.625
Distance	4	0	0	0
Fuel		0.134	-0.127	0.127
Trip Time		24.003	-20.625	20.625

Trip Time Savings and Fuel Efficiency

The comparison between Route A and Route B for 2-wheel vehicles reveals that Route B is generally more efficient in both trip time and fuel consumption as shown in Figure 3. Route B has a lower average trip time of 17.3 minutes compared to Route A's 20.8 minutes, with a smaller maximum duration of 60 minutes versus Route A's 75 minutes, making it the faster route overall. Additionally, Route B exhibits slightly lower average fuel consumption (0.1839 L) compared to Route A (0.1885 L), though the difference is minor. These findings suggest that Route B is the better choice for efficiency.

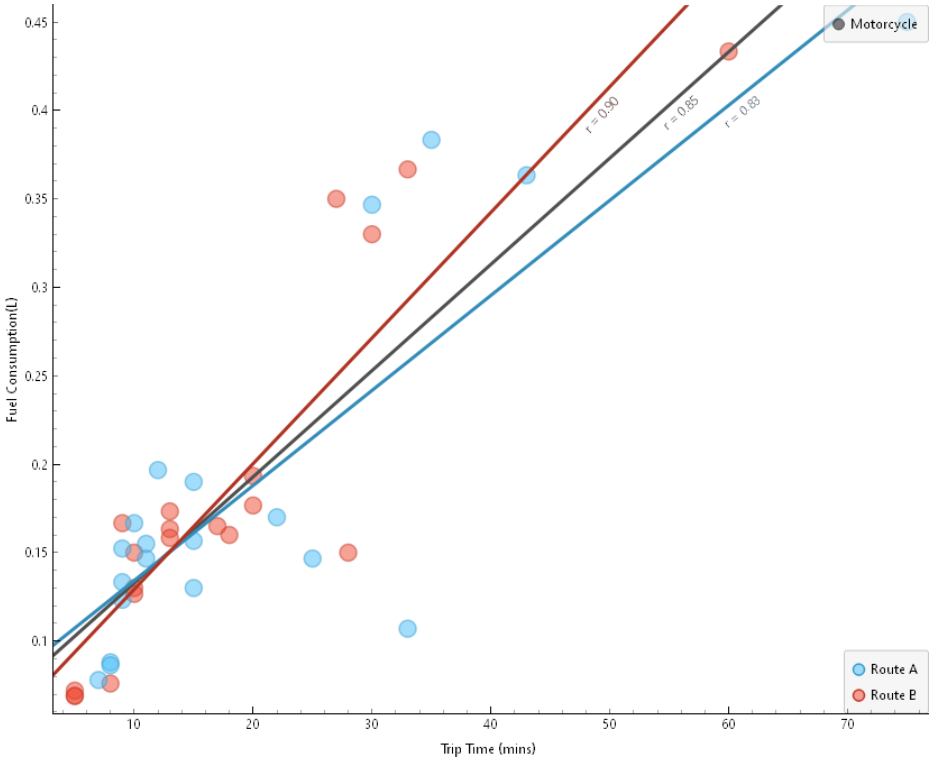


Figure 3: Comparison of Route A vs Route B (2-Wheel)

In Figure 4, Route A is the better choice for 3-wheel vehicles due to its lower average trip time and fuel consumption. While Route B involves a slightly longer route (4.46 km vs. 3.95 km), leading to marginally higher fuel use, Route A provides more efficient travel

overall. However, external factors like traffic conditions during peak hours still impact the effectiveness of both routes.

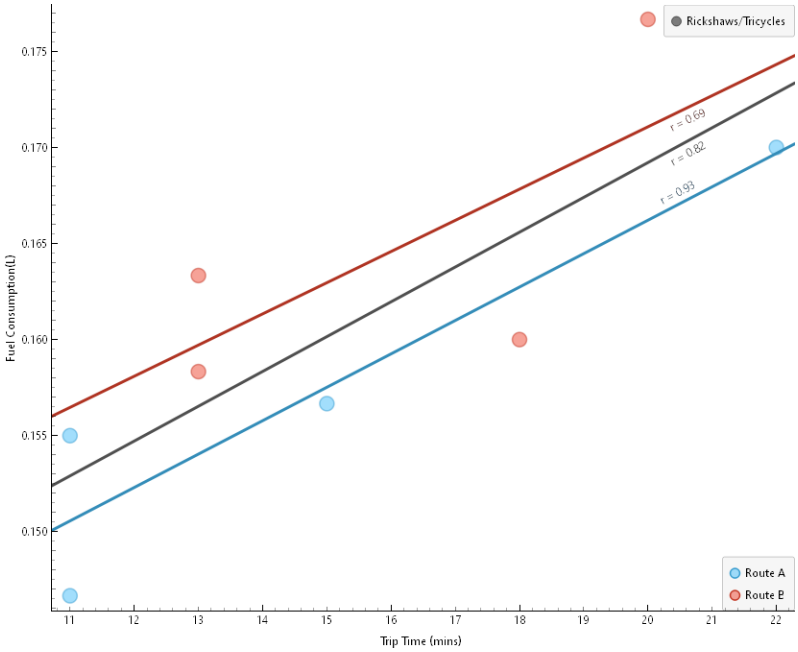


Figure 4: Comparison of Route A vs Route B (3-Wheel)

Route B is the better choice for 4-wheel vehicles due to its lower average trip time and fuel consumption as shown in Figure 5. Despite being slightly longer in distance (6.46 km vs. 6.14 km), it remains more efficient overall. However, heavy traffic conditions significantly impact both routes, especially during peak hours.

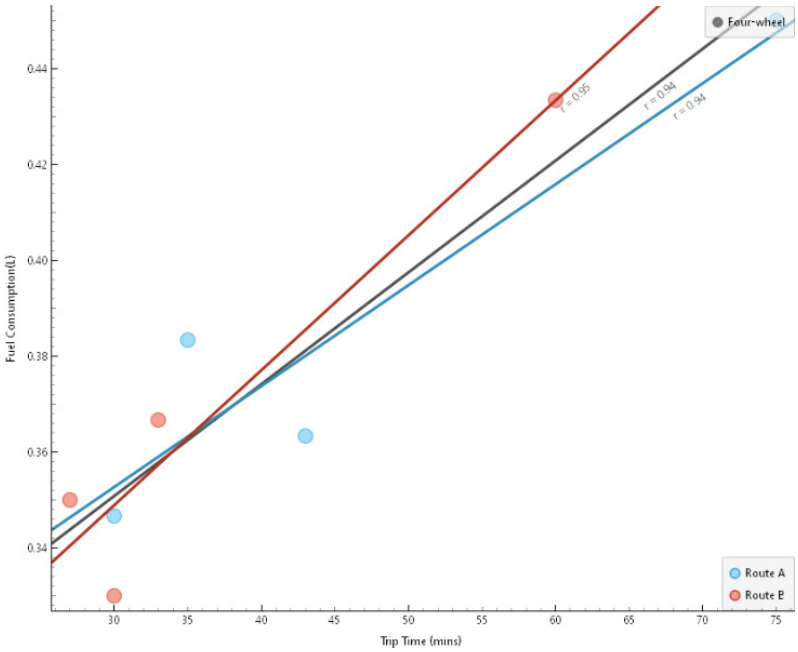


Figure 5: Comparison of Route A vs Route B (4-Wheel)

For all vehicle types, Table 3 shows that both Route A and Route B offer similar performance in terms of trip time and fuel consumption, making either route a viable choice.

For 2-Wheel vehicles, such as motorcycles, there is no significant difference in trip duration (p-value = 0.616) or fuel efficiency (p-value = 0.893), meaning both routes perform equally well. Similarly, for 3-Wheel vehicles, like tricycles, the comparison reveals no significant differences in either travel time (p-value = 0.705) or fuel usage (p-value = 0.284), suggesting that both routes are equally efficient. Finally, for 4-Wheel vehicles, such as cars, the results indicate that both routes have comparable performance in terms of trip time (p-value = 0.538) and fuel consumption (p-value = 0.637), allowing flexibility in route choice. In summary, drivers can confidently choose either route for each vehicle type, as both routes provide similar benefits in terms of efficiency and time.

Table 3: T-Test Results

Vehicle Type	Trip Time T-Statistic	Trip Time P-Value	Fuel T-Statistic	Fuel P-Value
2-Wheel	0.505255792	0.616297272	0.135356808	0.893044248
3-Wheel	-0.397359707	0.704852545	-1.176696811	0.283879872
4-Wheel	0.652262234	0.538402398	0.496855854	0.636968139

Conclusion

The accomplishment of RouteWise, a satellite navigation application for time-saving and traffic-aware travel, lies in providing a practical solution to one of the most common challenges in urban areas: traffic congestion. By utilizing real-time traffic data, Global Positioning System (GPS) technology, and advanced routing capabilities, the app aims to reduce travel time and improve efficiency by recommending optimized routes that minimize fuel consumption and maximize the efficient use of the available road network. Testing and deployment in Zamboanga City have shown that the application can effectively adapt to traffic situations, adjusting alternative routes to find those with minimal delays for the user. Although the app provides more convenient routes, travel time can still vary depending on the specific route and current traffic conditions. RouteWise has successfully managed driver flow in high-traffic areas, enhancing transportation efficiency within the city. However, some drawbacks, such as the need for consistent network service and GPS limitations, were observed during implementation. Despite these limitations, the system's overall performance classifies it as an effective tool for combating congestion and easing travel. This research contributes to the growing body of knowledge in transportation management by proposing a dynamic, adaptable solution to traffic flow challenges in cities with significant traffic issues. Future works may focus on improving the app's predictive accuracy, integrating additional data sources, such as weather conditions, and expanding its functionality to support more vehicle types and offline capabilities, further enhancing its potential for urban mobility optimization.

REFERENCES

[1] J. Gadziński, "Perspectives of the use of smartphones in travel behaviour studies: Findings from a literature review and a pilot study," Transp Res Part C Emerg Technol, vol. 88, pp. 74–86, Mar. 2018, doi: 10.1016/j.trc.2018.01.011.

[2] J. Pan, "Vehicle re-routing strategies for congestion avoidance," Ph.D. dissertation, New Jersey Institute of Technology, 2014. [Online]. Available: https://digitalcommons.njit.edu/dissertations/151.

[3] H. Prothmann et al., "Decentralised Route Guidance in Organic Traffic Control," in 2011 IEEE Fifth International Conference on Self-Adaptive and Self-Organizing Systems, IEEE, Oct. 2011, pp. 219–220. doi: 10.1109/SASO.2011.34.

[4] D. Metz, "The impact of digital navigation on travel behaviour," UCL Open Environ, vol. 4, Apr. 2022, doi: 10.14324/111.444/ucloe.000034.

[5] S. Jain and S. S. Jain, "Development of Intelligent Transportation System and Its Applications for an Urban Corridor During COVID-19," Journal of The Institution of Engineers (India): Series B, vol. 102, no. 6, pp. 1191–1200, Dec. 2021, doi: 10.1007/s40031-021-00556-y.

[6] N. C. Iraganaboina, T. Bhowmik, S. Yasmin, N. Eluru, and M. A. Abdel-Aty, "Evaluating the influence of information provision (when and how) on route choice pREFERENCES of road users in Greater Orlando: Application of a regret minimization approach," Transp Res Part C Emerg Technol, vol. 122, p. 102923, Jan. 2021, doi: 10.1016/j.trc.2020.102923.

[7] T. Alghamdi, S. Mostafi, G. Abdelkader, and K. Elgazzar, "A Comparative Study on Traffic Modeling Techniques for Predicting and Simulating Traffic Behavior," Future Internet, vol. 14, no. 10, p. 294, Oct. 2022, doi: 10.3390/fi14100294.

[8] Q. Lu, T. Tettamanti, D. Hörcher, and I. Varga, "The impact of autonomous vehicles on urban traffic network capacity: an experimental analysis by microscopic traffic simulation," Transportation Letters, vol. 12, no. 8, pp. 540–549, Sep. 2020, doi: 10.1080/19427867.2019.1662561.

[9] Roopashree V, "Traffic Congestion Detection and Alerting Ambulance using IoT," International Journal of Engineering Research and, vol. V9, no. 07, Jul. 2020, doi: 10.17577/IJERTV9IS070590.

[10] S. Aoyagi, Y. Le, T. Shimizu, and K. Takahashi, "Mobile Application to Provide Traffic Congestion Estimates and Tourism Spots to Promote Additional Stopovers," Future Internet, vol. 12, no. 5, p. 83, Apr. 2020, doi: 10.3390/fi12050083.

[11] W. Zhao, M. Quddus, H. Huang, J. Lee, and Z. Ma, "Analyzing drivers' pREFERENCES and choices for the content and format of variable message signs (VMS)," Transp Res Part C Emerg Technol, vol. 100, pp. 1–14, Mar. 2019, doi: 10.1016/j.trc.2019.01.013.

[12] Ma. J. J. Gumasing, F. J. C. S. Bermejo, K. T. C. Elpedes, L. F. E. Gonzales, and A. C. V. Villajin, "Antecedents of Waze Mobile Application Usage as a Solution for Sustainable Traffic Management among Gen Z," Sustainability, vol. 15, no. 13, p. 10186, Jun. 2023, doi: 10.3390/su151310186.

[13] D. Gunawan, Priyanka, and O. S. Sitompul, "The Development of Mobile and Server Application for Traffic Data Collection Based on Location-Based Services (LBS)," IOP Conf Ser Mater Sci Eng, vol. 851, no. 1, p. 012068, May 2020, doi: 10.1088/1757-899X/851/1/012068

SAĞLIK ÇALIŞANLARININ YAPAY ZEKÂYA İLİŞKİN FARKINDALIK DÜZEYLERİNİN DEĞERLENDİRİLMESİ

Semih Islıcık*
Bingöl Üniversitesi, Sağlık Yönetimi, Türkiye

Gamze Arıkan
Mersin Üniversitesi, Sağlık Yönetimi, Türkiye

Nazan Kartal
Çankırı Karatekin Üniversitesi, Sağlık Yönetimi, Türkiye

Osman Şahman
Batman Üniversitesi, Sağlık Yönetimi, Türkiye
(İletişimden sorumlu yazar) Ssemihislicik@hotmail.com

ÖZET

Yapay zekâ, sağlık sektöründe hızla yaygınlaşan bir teknoloji olarak karşımıza çıkmaktadır. Ancak, bu teknolojinin etkin bir şekilde kullanılabilmesi, sağlık çalışanlarının yapay zekâ farkındalık düzeylerine ve bu teknolojiye yönelik tutumlarına bağlıdır. Bu doğrultuda araştırmada, sağlık çalışanlarının yapay zekâ ile ilgili farkındalık düzeylerinin belirlenmesi ve değerlendirilmesi amaçlanmaktadır. Bu amaçla, Bingöl il merkezinde görev yapan 396 kamu sağlık çalışanının yapay zekâ farkındalık düzeyleri ele alınmıştır. Araştırmada elde edilen veriler frekans, yüzde dağılımları, bağımsız örneklem t testi, one way anova ve post-hoc tukey testleri kullanılarak değerlendirilmiş ve yorumlanmıştır. Gerçekleştirilen araştırma sonucunda sağlık çalışanları, yapay zekânın genel anlamda neyi ifade ettiğini bildiklerini, hayatı kolaylaştırdığını, sağlık kurumları için tanı ve tedavi olanakları açısından avantaj sağladığını ve daha yoğun kullanılması gerektiğini düşünmektedir. Bununla birlikte, sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin orta düzeyde olduğu, cinsiyet değişkeni açısından kadınların, yaş değişkeni açısından 24 yaş ve altı genç sağlık çalışanlarının ve meslek değişkeni açısından ise sırasıyla hemşire ve doktorların en yüksek yapay zekâ farkındalık ortalamalarına sahip olduğu sonucuna ulaşılmıştır. Sağlık kurumlarının yapay zekâ teknolojilerine yönelik yatırım yapması ve sağlık çalışanlarına bu konuda eğitimler, seminerler ve destekler sunması, çalışanların farkındalık düzeylerini artırarak bu sürecin başarıya ulaşmasında önemli bir rol oynayacaktır.

Anahtar Kelimler: Yapay zekâ, yapay zekâ farkındalık düzeyi, sağlık çalışanları.

Giriş

İnsanlığın erken dönemlerinden itibaren gelişen üretim becerileri, emek süreçleri ve çalışmalar, teknolojinin evrimine temel oluşturmaktadır. Söz konusu dinamik süreç, günümüzde yaşamın hemen hemen tüm alanlarında hayati bir rol oynayarak, insan faaliyetlerinin merkezinde yer almaktadır. Teknoloji, her geçen gün bir önceki güne göre yapabileceklerinin sınırlarını aşması, hayal dahi edilemeyecek yenilik ve gelişmelere kapılar aralaması nedeniyle insan hayatının vazgeçilmez bir parçasını oluşturmaktadır. Makinelerin insan benzeri bilişsel süreçleri ve davranışları kopyalamasını sağlayan temel bir teknoloji olarak ifade edilen yapay zekâ da insan hayatının vazgeçilmez bir parçası olma yolunda ilerlemektedir [1]. Yapay zekâ, sağlıktan eğitime enerjiden tarıma kadar neredeyse hayatımızın her alanına yerleşen önemli bir teknoloji haline gelmektedir. Sağlık sektörü, özellikle yapay zekânın dönüştürücü gücünden en fazla yarar sağlayan sektörlerin başında gelmektedir. Hastalıkların erken teşhisinden kişileştirilmiş tedavi yöntemlerine, ilaç keşfinden hasta bakım süreçlerinin iyileştirilmesine kadar birçok alanda yapay zekâ çözümleri, insan hayatını kolaylaştırmakta ve sağlık hizmetlerinin daha erişilebilir hale gelmesini sağlamaktadır [2].

Sağlık sektörü, yapay zekâ teknolojilerinin gelişmesiyle birlikte hızlı bir dönüşüm sürecinin içerisine yer almaktadır. Söz konusu dönüşüm, yapay zekânın sağlık hizmetleri ile entegrasyonuna yardımcı olurken hem sağlık bakım kalitesini hem de sağlığa erişebilirliği iyileştirerek bireysel hasta verilerine göre uyarlanmış tedavi planları oluşturmaktan sağlık hizmetleri süreçlerini iyileştirmeye kadar sağlık sektörüne çok sayıda önemli faydalar sağlamaktadır [3,4,5]. Elektronik sağlık kayıtları, tele-tıp, mobil uygulamalar ve akıllı cihazlar gibi birçok yapay zekâ uygulaması geliştirilmiştir [6,7,8]. Söz konusu uygulamalar sağlık hizmetlerinde erken tanı ve tedavi süreçlerinin hızlandırılması, iş yükünün azaltılması, hasta bakım kalitesinin iyileştirilmesi, maliyetlerin düşürülmesi ve tıbbi hataların azaltılmasında büyük rol oynamaktadır [9,10]. Sağlık hizmetlerinde yapay zekâ, hastalıkları teşhis ederek ve hastalıklarla ilgili zamanında doğru ve etkili öneriler ve talimatlar vererek sağlık çalışanı gibi davranabilmektedir [11]. Bununla birlikte yapay zekâ, günlük tıbbi uygulamalarda hastalıkları teşhis etmek için hastaların radyolojik görüntülerinden yararlanarak hastaların kişisel sağlık bilgilerine göre özelleştirilmiş sağlık bakım planları önermektedir. Acil sağlık hizmetleri için büyük öneme sahip olan triyaj hizmetlerinde yapay zekâ sistemleri, gerekli sağlık bakım aciliyeti hakkında bilinçli kararlar almak için hasta verilerini, semptomlarını ve tıbbi geçmişini saniyeler içerisinde analiz etmektedir [12]. Yapay zekâ destekli sohbet robotları ve sanal asistanlar, hastalara bilgi ve destek sunmak, sorularına yanıt vermek ve sağlık durumlarını yönetmelerine yardımcı olmak için kullanılmaktadır. Söz konusu durum, yalnızca hasta sonuçlarını iyileştirmekle kalmaz, aynı zamanda sağlık kaynaklarının kullanımını da optimize etmektedir. Dünyaca ünlü önde gelen bazı ilaç şirketleri, genetik bilgi ve hastalık kalıplarının geniş veri kümelerini analiz etmek, yeni ilaçların keşfini hızlandırmak ve bunları pazara sunmak ile ilişkili zaman ve maliyet tasarrufu sağlamak için farmakoloji ve ilaçlarda yapay zekâyı dikkate değer şekilde kullanmaktadır. Ameliyatlarda, yapay zekâ destekli cerrahi robotlar, cerrahların hassas ve minimal kesikleri gerçekleştirmesine yardımcı olmakta, bu durum ise kesik yaralarının daha hızlı iyileşmesini sağlamakta ve komplikasyon riskini en aza indirgemektedir [13]. İnsan yeteneklerinin gözden kaçırabileceği anormallikleri tespit etme yeteneğine sahip olan yapay zekâ algoritmaları, tıbbi görüntüleri doğru bir şekilde analiz etmektedir. Son zamanlarda, diş hekimliği alanı yapay zekâyı diş görüntülerini analiz

etmek ve uygun tıbbi öneriler sağlamak için kullanmaktadır. Yapay zekâ destekli diş araçları, hekimlere ağız sağlığı sorunlarının teşhisinde, hastaların tedavilerinin planlanmasında ve hasta gelişiminin izlenmesinde yardımcı olmaktadır [14]. Yapay zekâ algoritmalarının sağlık hizmetlerinde en yaygın kullanımı radyoloji alanında gerçekleşmektedir. Ancak, dermatoloji, psikiyatri, kardiyoloji, onkoloji, nörobilim, patoloji ve tıp eğitimi dahil olmak üzere diğer disiplinlerde de kullanımına ilişkin birçok örnek bulunmaktadır [15–19]. Yapay zekâ teknolojileri, sağlık ile ilgili birçok alanda avantajlar sağlasa da bazı dezavantajları da içerisinde barındırmaktadır. Kişisel sağlık verilerinin güvenlik açıkları ile hasta mahremiyeti tehlike altına girmekte, veri gizliliği ve güvenlik sorunları oluşmaktadır [20]. Veri odaklı kararlar alan yapay zekâ sistemlerinin her zaman güvenilir sonuç vermemesi ve hatalı kararların alınabilmesi hasta güvenliği açısından risk oluşturmaktadır [21]. Otomasyon süreçlerinde özellikle yapay zekâ teknolojilerinin kullanılması, sağlık çalışanları arasında olası iş kaygılarını da beraberinde getirmektedir [22]. Söz konusu durum, sağlık çalışanlarının yapay zekâ teknolojilerine karşı önyargılı olmalarına ve teknolojileri benimsemekte zorluk çekmelerine sebep olmaktadır. Diğer taraftan, yapay zekâ modellerinin sürekli olarak optimize edilmesi ve güncellenmesi gerektiği için teknik sınırlamalar ve geliştirme maliyetleri oluşmaktadır [23]. Yapay zekânın yenilikçi potansiyelini belirtilen dezavantajlar ile birlikte göz önünde bulundurmak etkili, güvenli, verimli, eşitlikçi ve etik değerleri gÖZETen bir sağlık hizmeti sağlamak için büyük önem arz etmektedir.

Sağlık sektöründe büyük dönüşümlerin habercisi olan yapay zekâ ve teknolojilerinin etkin bir şekilde kullanılması ve benimsenmesi için sağlık çalışanlarının rolü ve farkındalıkları büyük önem arz etmektedir [24]. Hekim ve hemşirelerden, eczacı ve sağlık teknikerlerine kadar uzanan birincil amaçları sağlığı iyileştirmek olan faaliyetlerde bulunan sağlık çalışanları, hasta bakımının ön saflarında yer almaktadır. Söz konusu durum ise sağlık çalışanlarının bakış açılarını ve yapay zekâ teknolojilerini benimsemelerini başarılı bir entegrasyon için önemli hale getirmektedir [25]. Yapay zekânın sağlık hizmetlerindeki etkinliği, sağlık çalışanları ve kullandıkları teknolojiler arasındaki farkındalığın yüksek olmasına bağlıdır [26,27]. Yapay zekânın sağlık hizmetleri faaliyetlerinde kullanılmasının yaygınlaşması ile birlikte sağlık çalışanlarının yapay zekâ kavramları ve teknolojilerine yönelik farkındalıklarının değerlendirilmesi zorunluluk haline gelmiştir [28,29]. Yapay zekâ teknolojileri hızla gelişirken, söz konusu teknolojilerle doğrudan ilgilenenlerin bakış açıları ve farkındalıkları hakkında çok az şey bilinmektedir. Bu doğrultudan hareketle araştırmada, sağlık çalışanlarının yapay zekâ ile ilgili farkındalık düzeylerinin belirlenmesi ve değerlendirilmesi amaçlanmaktadır.

Materyal ve Yöntem

Sağlık çalışanlarının yapay zekâ ile ilgili farkındalık düzeylerinin incelenmesi ve değerlendirilmesi amacıyla gerçekleştirilen bu araştırma tanımlayıcı ve keşifsel türde gerçekleştirilmiştir. Söz konusu araştırma, Türkiye'nin Doğu Anadolu Bölgesinde yer alan Bingöl il merkezinde faaliyet gösteren kamu sağlık kurumlarında görev yapan sağlık çalışanları ile 01.12.2024 – 20.02.2025 tarihleri arasında gerçekleştirilmiştir. Araştırma için bir kamu üniversitesinin lisansüstü eğitim enstitüsü etik komisyonundan, etik kurul izni (29.11.2024 tarihli ve 194977 sayılı) alınmıştır. Türkiye Cumhuriyeti Sağlık Bakanlığı Yönetim Hizmetleri Genel Müdürlüğü'nün güncel (18.10.2024) olarak yayınlamış olduğu Personel Dağılım Cetveli (PDC) tablosuna göre Bingöl il merkezi kamu sağlık kurumlarında görev yapan sağlık çalışanları

sayısı (hekim, diş hekimi ve eczacı sayısı 303, diğer sağlık personeli 1536) 1839 olarak tespit edilmiştir [30]. Araştırmada evreni bilinen örneklem sayısı hesaplama tablosuna göre evren büyüklüğü 1839, kabul edilebilir hata %5, güven seviyesi %95 olarak hesaplandığında en az 318 sağlık çalışanı araştırma örneklemine alınması gerekmektedir [31,32]. Bu doğrultuda, Bingöl il merkezinde kamu sağlık kurumlarında görev yapan 396 sağlık çalışanı araştırmanın örneklemini oluşturmuştur. Araştırmanın örneklem yöntemi ise basit tesadüfi örneklem olarak belirlenmiştir.

Veri Toplama Aracı

Sağlık çalışanlarının yapay zekâ farkındalıklarının değerlendirilmesine yönelik olarak elde edilen veriler anket aracılığıyla toplanmıştır. Üç bölümden oluşan anketin birinci bölümünde, sağlık çalışanlarının yaş, cinsiyet, medeni ve eğitim durumları, meslek ve toplam mesleki deneyimlerinin yer aldığı tanıtıcı bilgiler yer almaktadır. İkinci bölümde, genel olarak yapay zekâ ile ilgili sağlık çalışanlarına yöneltilen "Evet" ve "Hayır" seçeneklerinin yer aldığı 8 adet soru yer almaktadır. Anketin üçüncü ve son bölümünde ise Yılmaz ve arkadaşları [33] tarafından geliştirilmiş 27 ifadeden oluşan 5'li likert tipi "Kesinlikle katılmıyorum (1)", "Katılmıyorum (2)", "Kararsızım (3)", "Katılıyorum (4)" ve "Kesinlikle Katılıyorum (5)" ifadelerinden oluşan yapay zekâ farkındalık ölçeği yer almaktadır. Anket, yüz yüze ve Google formlar aracılığı ile online bir anket linki oluşturularak katılımcı sağlık çalışanlarına bilgilendirilmiş gönüllü onamları alınarak uygulanmıştır.

Araştırmanın Değişkenleri

Sağlık çalışanlarının yapay zekâ farkındalıklarının değerlendirildiği ve tek boyuttan oluşan yapay zekâ farkındalık ölçeği, araştırmanın bağımlı değişkenini oluştururken, sağlık çalışanlarının yaş, cinsiyet, medeni ve eğitim durumları, meslek ve toplam mesleki deneyimlerinin yer aldığı tanıtıcı bilgiler araştırmanın bağımsız değişkenlerini oluşturmaktadır.

Araştırmanın Hipotezleri

Araştırmanın amaçlarına uygun olarak geliştirilen hipotezler aşağıda gösterilmiştir;

H_1: Sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin yaş değişkenine göre istatistiksel olarak anlamlı farklılık göstermektedir.

H_2: Sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin cinsiyet değişkenine göre istatistiksel olarak anlamlı farklılık göstermektedir.

H_3: Sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin medeni durum değişkenine göre istatistiksel olarak anlamlı farklılık göstermektedir.

H_4: Sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin eğitim durumu değişkenine göre istatistiksel olarak anlamlı farklılık göstermektedir.

H_5: Sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin meslek değişkenine göre istatistiksel olarak anlamlı farklılık göstermektedir.

H_6: Sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin toplam mesleki deneyim değişkenine göre istatistiksel olarak anlamlı farklılık göstermektedir.

Araştırma Verilerinin Analizi

Araştırma kapsamında sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin değerlendirilmesinde elde edilen veriler, SPSS 26.0 (Statistical Package for Social Sciences 26.0) programı kullanılarak analiz edilmiştir. Verilerin normal dağılım gösterdiği Kolmogorov Smirnov testi ($p>0,05$) ile tespit edilmiştir. Verilerin analizinde tanımlayıcı istatistikler olarak ortalama ve standart sapmalar gösterilmiştir. Yapay zekâ farkındalık düzeyleri için Düşük farkındalık düzeyi: Yapay zekâ farkındalık ölçeği ortalaması < 2,5; Orta farkındalık düzeyi: Yapay zekâ farkındalık ölçeği ortalaması 2,5–3,5; Yüksek farkındalık düzeyi: Yapay zekâ farkındalık ölçeği ortalaması > 3,5 olarak değerlendirilmiştir [52]. Kategorik yapıdaki değişkenler için sayı ve yüzde değerlerine yer verilmiştir. Veri setlerinin normal dağılması nedeniyle iki bağımsız grup ortalamalarının karşılaştırılması için Bağımsız örneklem t testi, üç ve daha fazla grup ortalamalarının karşılaştırılması için One Way Anova testi sonucu anlamlı çıkan farkların, hangi gruplardan kaynaklandığını tespit etmek için Post-Hoc Tukey testinden yararlanılmıştır.

Bulgular

Sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin belirlenmesi ve değerlendirilmesine ilişkin katılımcıların demografik özellikleri Tablo 1'de gösterilmiştir.

Tablo 1: Sağlık çalışanlarının demografik özelliklerinin dağılımı (N:396)

Özellikler		n	%
Cinsiyet	Kadın	176	44,4
	Erkek	220	55,6
Yaş	25 yaş ve altı	20	5,1
	25-34 yaş	180	45,5
	35-44 yaş	136	34,3
	45-54 yaş	52	13,1
	55 yaş ve üzeri	8	2,0
Medeni Durum	Evli	268	67,7
	Evli Değil	128	32,3
Eğitim Durumu	Lise	44	11,1
	Ön Lisans	96	24,2
	Lisans	219	55,3
	Lisansüstü	37	9,4
Meslek	Doktor	104	26,3
	Eczacı	12	3,0
	Hemşire	96	24,2
	Ebe	36	9,1
	Sağlık Teknikeri	60	15,2
	Sağlık Memuru	36	9,1
	Diğer	52	13,1
Toplam Mesleki Deneyim	3 yıl ve altı	92	23,2
	4-7 yıl	88	22,2
	8-11 yıl	52	13,1
	12-15 yıl	64	16,2
	16 yıl ve üzeri	100	25,3

Katılımcıların çoğunluğunu erkek sağlık çalışanları oluştururken, %45,5'ini 25-34 yaş arası ve yaklaşık olarak %68'ini de evli sağlık çalışanları oluşturmaktadır. Sağlık çalışanlarının yaklaşık olarak %65'i lisans ve lisansüstü eğitim kurumlarından mezun olmuştur. Katılımcıların çoğunluğunu %26,3 ile doktorlar ve %24,2 ile hemşireler oluşturmuştur. Toplam mesleki deneyim açısından neredeyse dengeli bir dağılıma sahip olan sağlık çalışanlarının %25,3'ü 16 yıl ve üzeri deneyime sahipken %13,1'i de 8-11 yıl toplam meslek deneyimine sahip olarak tespit edilmiştir.

Tablo 2: Sağlık çalışanlarının yapay zekâ ile ilgili sorulara ilişkin yanıtlarının dağılımı (N:396)

		n	%
1-Yapay zekânın ne olduğunu biliyor musunuz?	Evet	376	94,9
	Hayır	20	5,1
2-Daha önce yapay zekâ teknolojisiyle hiç etkileşimde bulundunuz mu?	Evet	204	51,5
	Hayır	192	48,5
3-Yapay zekânın kullanılması mesleki açıdan sizi endişelendiriyor mu?	Evet	196	49,5
	Hayır	200	50,5
4- Yapay zekânın hayatınızı kolaylaştırdığı fikrine katılıyor musunuz?	Evet	336	84,8
	Hayır	60	15,2
5- Yapay zekâ kullanılan hastaneler tanıda daha avantajlı mıdır?	Evet	292	73,7
	Hayır	104	26,3
6-Yapay zekâ kullanılan hastaneler tedavide daha avantajlı mıdır?	Evet	272	68,7
	Hayır	124	31,3
7-Tıpta kullanılan yapay zekâ uygulamaları daha yoğun kullanılmalı mıdır?	Evet	292	73,7
	Hayır	104	26,3
8-Tıpta kullanılan yapay zekâ uygulamalarını güvenilir buluyor musunuz?	Evet	236	59,6
	Hayır	160	40,4

Sağlık çalışanlarına yapay zekâ ile ilgili yöneltilen 8 genel sorunun yanıtları Tablo 2'de gösterilmektedir. Katılımcıların büyük bir çoğunluğu (%94,9) yapay zekânın ne olduğunu bilmekteyken, yapay zekâ teknolojisiyle hiç etkileşimde bulunmayan sağlık çalışanı oranı %48,5 olarak tespit edilmiştir. Yapay zekâ kullanımının sağlık çalışanlarının neredeyse yarısını (%49,5) meslekleri ile endişelendirdiği gözlemlenirken diğer taraftan %84,8'inin hayatlarını kolaylaştırdığı fikrine sahip oldukları belirlenmiştir. Yapay zekâ kullanılan hastanelerin tanı ve tedavi açısından daha avantajlı olduğu sağlık çalışanlarının çoğunluğu tarafından kabul edilmiştir. Tıpta kullanılan yapay zekâ uygulamalarının daha yoğun kullanılmasını isteyen sağlık çalışanlarının oranı %73,7 iken, kullanılan uygulamaların güvenli olduğunu düşünen sağlık çalışanlarının oranı %59,6 olarak tespit edilmiştir.

Tablo 3: Yapay zekâ farkındalık ölçeği istatistikleri ve ölçek güvenilirliği

	n	Ort.	S.s	İfade Sayısı	Cronbach's Alpha
Yapay Zekâ Farkındalık Düzeyi	396	3,11	0,81	27	0,96

Tablo 3'te gösterildiği üzere sağlık çalışanlarının yapay zekâ farkındalık ölçeği ortalaması 3,11 olarak tespit edilmiştir. Bununla birlikte, 27 ifade ve tek boyuttan oluşan yapay zekâ farkındalık ölçeğinin güvenilirliğini belirten iç tutarlılık katsayısı olan Cronbach's Alpha

katsayısının 0,96 olduğu ve yapay zekâ farkındalık ölçeğinin yüksek geçerlilik ve güvenilirlik düzeyine sahip olduğu gözlemlenmiştir.

Sağlık çalışanlarının yapay zekâ farkındalık düzeyleri ile demografik değişkenleri (cinsiyet, medeni durum, yaş, eğitim durumu, meslek, toplam mesleki deneyim) arasında istatistiksel olarak anlamlı farklılıkların olup olmadığı ve oluşturulan hipotezlerin testi Tablo 4'te yer almaktadır.

Tablo 4: Sağlık çalışanlarının yapay zekâ farkındalığı ve demografik değişkenler bakımından incelenmesi

Demografik Özellikler			**n (396)**	**Ort**	**S.s**	**t**	**p**	**Farklılığın Kaynağı**
Yapay Zekâ Farkındalık Ölçeği Ortalaması	Cinsiyet	Kadın	176	3,32	0,72	4,829	0,0002*	Kadın ile Erkek
		Erkek	220	2,95	0,83			
	Medeni Durum	Evli	268	3,10	0,85	-0,666	0,506	-
		Evli Değil	128	3,15	0,70			
Demografik Özellikler			**n (396)**	**Ort**	**S.s**	**F**	**p**	**Farklılığın Kaynağı**
Yapay Zekâ Farkındalık Ölçeği Ortalaması	Yaş	24 yaş ve altı	20	3,40	0,13	2,879	0,023*	25 yaş altı ile 45-54 yaş ve 55 yaş ve üzeri
		25-34 yaş	180	3,15	0,06			
		35-44 yaş	136	2,97	0,07			
		45-54 yaş	52	2,88	0,10			
		55 yaş ve üzeri	8	2,85	0,44			
	Eğitim Durumu	Lise	44	3,05	0,71	1,760	0,154	-
		Ön Lisans	96	2,97	0,92			
		Lisans	137	3,16	0,81			
		Lisansüstü	119	3,20	0,72			
	Meslek	Doktor	104	3,19	0,67	5,605	0,0001*	Doktor ile Sağlık Memuru ve Hemşire ile Sağlık Memuru
		Eczacı	12	2,95	0,84			
		Hemşire	96	3,35	0,80			
		Ebe	36	2,94	0,79			
		Sağlık Teknikeri	60	2,82	0,92			
		Sağlık Memuru	36	2,70	0,80			
		Diğer	52	2,77	0,73			
	Toplam Mesleki Deneyim	3 yıl ve altı	92	3,26	0,82	1,744	0,140	-
		4-7 yıl	88	3,15	0,66			
		8-11 yıl	52	3,19	0,74			
		12-15 yıl	64	3,10	0,89			
		16 yıl ve üzeri	100	2,99	0,87			

*p<0,05

Sağlık çalışanlarının yapay zekâ farkındalıklarının cinsiyet değişkenine göre istatistiksel olarak anlamlı farklılık göstermektedir. Kadın sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin daha yüksek olduğu tespit edilmiştir. Bu doğrultuda oluşturulan $\mathbf{H_2}$ hipotezi kabul edilmiştir. Katılımcıların yapay zekâ farkındalık düzeylerinin medeni durum değişkeni ve eğitim durumu değişkenine göre anlamlı bir fark bulunamamıştır. Dolayısıyla, $\mathbf{H_3}$ ve $\mathbf{H_4}$ hipotezleri reddedilmiştir. Sağlık çalışanlarının yapay zekâ farkındalıklarının yaş değişkenine

göre istatistiksel olarak anlamlı farklılık gösterdiği gözlemlenmiş ve $\mathbf{H_1}$ hipotezi kabul edilmiştir. Söz konusu farklılığın, 25 yaş ve altındaki sağlık çalışanlarının 45-44 yaş arasında ve 55 yaş ve üzeri sağlık çalışanlarına göre yapay zekâ farkındalık düzeylerinin anlamlı bir şekilde daha yüksek olmasından kaynaklandığı belirlenmiştir. Katılımcıların yapay zekâ farkındalık düzeylerinin meslek değişkenine göre istatistiksel olarak anlamlı farklılık gösterdiği tespit edilerek $\mathbf{H_5}$ hipotezi kabul edilmiştir. Söz konusu farklılıklar, doktorlar ile sağlık memurları, hemşireler ile sağlık memurları ve sağlık teknikerleri arasında gözlemlenmiştir. Mesleklerine göre en yüksek yapay zekâ farkındalığına sahip sağlık çalışanlarının sırasıyla, hemşire (3,35), doktor (3,19), eczacı (2,95), ebe (2,94), sağlık teknikeri (2,82) ve sağlık memuru (2,70) olduğu tespit edilmiştir. Son olarak, katılımcıların yapay zekâ farkındalıklarının toplam mesleki deneyim değişkenine göre istatistiksel olarak anlamlı farklılık bulunmazken, $\mathbf{H_6}$ hipotezi reddedilmiştir.

Tartışma ve Sonuç

Yapay zekâ tıbbi teşhis, tedavi planlaması, hastalık takibi ve hasta bakımını iyileştirme potansiyeliyle sağlık hizmetlerinde dönüşüm yaratma potansiyeli taşımaktadır. Ancak söz konusu teknolojilerin etkin bir şekilde kullanılabilmesi için sağlık çalışanlarının yapay zekâ konusunda farkındalıklarının arttırılması ve bu alandaki bilgi seviyelerinin belirlenmesi, değerlendirilmesi ve yükseltilmesi konusundaki çabalar büyük önem arz etmektedir. Bu doğrultuda araştırma, sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin değerlendirilmesi amacıyla Bingöl il merkezinde görev yapan 396 kamu sağlık çalışanıyla gerçekleştirilmiştir.

Araştırmada, yapay zekâ ile ilgili olarak sorulan genel sorulara ilişkin sağlık çalışanlarının vermiş oldukları yanıtlar incelendiğinde, "Yapay zekânın ne olduğunu biliyor musunuz?" sorusuna katılımcıların %94,9'unun evet yanıtını verdiği sonucuna ulaşılmıştır. Çağlar'ın [34] 2024 yılında, sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin incelenmesi amacıyla Malatya il merkezinde faaliyet gösteren kamu hastanesinde görev yapan 300 sağlık çalışanı ile gerçekleştirmiş olduğu çalışmasında söz konusu soruya verilen evet yanıtlarının %44 olduğu, diğer taraftan Orhan ve Bülez'in [35] sağlık personellerinin yapay zekâ ile ilgili düşüncelerinin değerlendirilmesi amacıyla bir kamu hastanesinde 68 kamu sağlık çalışanı ile gerçekleştirdikleri çalışmalarında ise bu oranın %94,1 olduğu gözlemlenmiştir. Sağlık çalışanlarının yapay zekânın ne olduğu ile ilgili görüşleri farklılık göstermektedir. Araştırmada sorulan "Daha önce yapay zekâ teknolojisiyle hiç etkileşimde bulundunuz mu?" sorusuna sağlık çalışanlarının %51,5'i evet yanıtını vermiştir. Çağlar'ın [34] çalışmasında bu oran %30,3 olarak tespit edilmiştir. "Yapay zekânın kullanılması mesleki açıdan sizi endişelendiriyor mu?" sorusuna, araştırmada verilen evet yanıtlarının oranı %49,5'tir. Shagufta Nesar ve arkadaşlarının [2] Pakistan'da 315 katılımcı ile gerçekleştirmiş oldukları sağlık çalışanlarının yapay zekâ bilgi ve algılarının incelendiği araştırmada, sağlık çalışanlarının %62'sinin mesleki açıdan yapay zekâ ile ilgili endişe etmektedir. Sağlık çalışanlarının mesleki açıdan yapay zekâ ile ilgili endişelerinin olduğunu Gordon ve arkadaşları [37], Wen ve Huang [38] ve Alowais ve arkadaşları [39] tarafından da bildirilmiştir. Söz konusu durum ayrıca Elnaggar ve arkadaşları [40] tarafından da desteklenmektedir. Araştırmada "Yapay zekânın hayatınızı kolaylaştırdığı fikrine katılıyor musunuz?" sorusuna sağlık çalışanlarının %84,8'i evet yanıtını vermiştir. Söz konusu oran Orhan ve Bülez'in [35] çalışmasında %79,4 iken Shagufta Nesar ve arkadaşlarının [2] çalışmasında ise %60 olarak gözlemlenmiştir. Sağlık çalışanlarının

"Yapay zekâ kullanılan hastaneler tanıda daha avantajlı mıdır?" ve "Yapay zekâ kullanılan hastaneler tedavide daha avantajlı mıdır?" sorularına verilen evet yanıtları sırasıyla %73,7 ve %68,7 olarak sonuçlara yansımıştır. Yapay zekâ teknolojilerini kullanan sağlık kurumlarının, sağlık hizmetlerinde tanı ve tedavide daha avantajlı olduğu sağlık çalışanları tarafından düşünülmektedir. Shagufta Nesar ve arkadaşları da [2] çalışmalarında, sağlık çalışanlarının %47,6'sının sağlık kurumlarında yapay zekânın, tanı ve tedavide avantaj sağladığını gözlemlemişlerdir. Bununla birlikte, Kimiafar ve arkadaşları [41], Yorgancıoğlu Tarcan ve Yalçın Balçık [42], Wang ve arkadaşları [43], Karal ve Turan [44], Kong ve arkadaşları [45], Çilhoroz ve Oğuz [46], Alelyani ve arkadaşları [47] ve Topol [21] sağlık hizmetlerinde yapay zekânın birçok avantajı ve yararı olduğunu bildirmişlerdir. Araştırmada sağlık çalışanlarına yöneltilen "Tıpta kullanılan yapay zekâ uygulamaları daha yoğun kullanılmalı mıdır?" sorusuna verilen evet yanıtları %73,7 olarak tespit edilmiştir. Sağlık çalışanlarının, tıp alanında yapay zekânın yoğun kullanılmasını yüksek bir oranda desteklediği görülmektedir. Araştırmada yapay zekâ ile ilgili soruların sonuncusu "Tıpta kullanılan yapay zekâ uygulamalarını güvenilir buluyor musunuz?" sorusudur. Belirtilen soruya, sağlık çalışanlarının %59,6'sı evet yanıtını vermiştir. Sağlık çalışanlarının yapay zekâya güven oranı Orhan ve Bülez'in [35] çalışmasında %58,8 olarak belirlenmiştir. Diğer taraftan, Çağlar'ın [34] çalışmasında bu oran %30 olarak tespit edilmiştir. Liyanage ve arkadaşları [49] ile Wiedermann ve arkadaşları [50] sağlık çalışanlarının yapay zekâ sonuçları hakkındaki endişelerini vurgularken, Castagno ve Khalifa [3] ile Chew ve Achananuparp [51] ise sağlık çalışanlarının yapay zekâ tarafından üretilen bilgilerden olumlu yönde etkilendiklerini ve yapay zekâya güvendiklerini ifade etmişlerdir. Sağlık çalışanlarının yapay zekâya güven konusundaki düşüncelerinin farklılık gösterdiği sonucuna ulaşılmıştır.

Sağlık çalışanlarının yapay zekâ farkındalık ölçeği ortalamasının 3,11 olduğu tespit edilmiştir. Söz konusu ortalamaya göre araştırma kapsamında, sağlık çalışanlarının orta düzey yapay zekâ farkındalık düzeyine sahip oldukları sonucuna ulaşılmıştır. Araştırma sonuçları, Shagufta Nesar ve arkadaşları ile Al-Qerem ve arkadaşlarının çalışma sonuçları ile paralellik göstermektedir. Çağlar'ın [34] çalışmasında sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin düşük düzeyde olduğu, Alsaedi ve arkadaşları [52] ile Hamd ve arkadaşlarının [48] çalışmalarında, sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin yüksek düzeyde olduğu sonucuna ulaşılmıştır. Söz konusu durum, sağlık çalışanlarının bireysel ve kurumsal olarak dijital ve sağlık altyapısı farklılıklarının yapay zekâ farkındalık düzeyleri arasındaki farklılıkları açıklayabileceğini düşündürmektedir.

Sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin cinsiyet, yaş ve meslek açısından istatistiksel olarak anlamlı farklılıklar gözlemlenmiştir. Araştırmada, kadın sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin, erkeklere göre daha yüksek olduğu tespit edilmiştir. Çağlar [34] ise erkek sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin kadınlara göre daha yüksek olduğunu gözlemlemiştir. 25 yaş ve altı sağlık çalışanlarının yapay zekâ farkındalıklarının diğer yaş kategorilerine göre daha yüksek olduğu özellikle 25 yaş ve altı sağlık çalışanlarının, 45-54 yaş ve 55 yaş ve üzeri arasında istatistiksel olarak anlamlı farklılıklar araştırmada ortaya koyulmuştur. Bununla birlikte, sağlık çalışanlarının yaşları arttıkça yapay zekâ farkındalık düzeylerinde azalış olduğu sonucuna ulaşılmıştır. Yapay zekâ teknolojilerine genç kuşakların daha yatkın ve farkındalıklarının yüksek olduğu Hoşgör ve Bozkurt'un [36] sağlıkta yapay zekâ hakkında kuşakların karşılaştırılmasında en genç kuşak

olan Z Kuşağı'nın istatistiksel açıdan anlamlı derecede yapay zekâ kavramına aşina oldukları ve yapay zekânın tıp alanında kullanımı hakkında bilgi sahibi oldukları belirlenmiştir. Shagufta Nesar ve arkadaşlarının [2] çalışmasında ise 20-30 yaş arası sağlık çalışanlarının yapay zekâ bilgi düzeylerinin daha yüksek olduğu gözlemlenmiştir. Sağlık çalışanları mesleki açıdan değerlendirildiğinde, en yüksek yapay zekâ farkındalık düzeyine hemşirelerin daha sonra ise doktorların, en düşük yapay zekâ farkındalık düzeyine ise sağlık memurlarının sahip olduğu sonucuna ulaşılmıştır. Hemşireler ile sağlık memurları arasında ve doktorlar ile sağlık memurları arasında istatistiksel olarak anlamlı farklılıklar tespit edilmiştir. Shagufta Nesar ve arkadaşları [2] en yüksek yapay zekâ farkındalığına sahip mesleğin eczacılar, Çağlar [34] ise doktorların olduğunu belirlemişlerdir. Araştırma ile ilgili olarak sağlık çalışanlarının medeni durum, eğitim durumu ve toplam mesleki deneyim değişkenleri ile yapay zekâ farkındalık düzeyleri arasında istatistiksel olarak anlamlı bir fark olmadığı sonucuna ulaşılmıştır.

Sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin belirlenmesi ve değerlendirilmesi amacıyla Bingöl il merkezinde görev yapan kamu sağlık çalışanları ile gerçekleştirilen araştırma kapsamında sonuçlar genel olarak değerlendirildiğinde, sağlık çalışanları yapay zekânın genel anlamda neyi ifade ettiğini bildiklerini, hayatı kolaylaştırdığını, sağlık kurumları için tanı ve tedavi olanakları açısından avantaj sağladığını ve daha yoğun kullanılması gerektiğini düşünmektedir. Bununla birlikte, sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin orta düzeyde olduğu, cinsiyet değişkeni açısından kadınların, yaş değişkeni açısından 24 yaş ve altı genç sağlık çalışanların ve meslek değişkeni açısından ise hemşire ve doktorların en yüksek yapay zekâ farkındalık ortalamalarına sahip olduğu sonucuna ulaşılmıştır.

Araştırma kapsamı, zaman ve maliyet unsuru göz önünde bulundurulması nedeniyle Bingöl il merkezinde görev yapan kamu sağlık çalışanları ile sınırlandırılmıştır. Araştırma yalnızca Bingöl il merkezinde ve kamu sağlık çalışanları ile yapıldığı için sonuçların tüm sağlık çalışanlarına genellemesi yapılamaz. Araştırmada kullanılan anket uygulaması, 01.12.2024-20.02.2025 tarihleri arasında sınırlandırılarak gerçekleştirilmiştir. Bu bağlamda, gelecek araştırmalar için sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin belirlenmesi ile ilgili Türkiye genelini yansıtacak, kamu ve özel kurumlarda görev yapan sağlık çalışanlarıyla kapsamlı araştırmaların yapılması faydalı olacaktır. Sağlık çalışanlarının yapay zekâ farkındalık düzeylerinin yükseltilmesi için yapay zekâ ilkelerine ilişkin bilişsel kavrayışı geliştirmeyi amaçlayan hedefli eğitim programlarının ve disiplinler arası iş birliklerinin gerçekleştirilmesi gerekmektedir. Özellikle, yaş seviyesi daha yüksek olan sağlık çalışanların yapay zekâ farkındalık seviyelerinin yükseltilmesi için özel seminerlerin gerçekleştirilmesi önerilmektedir. Eğitim, doğru altyapı, güven oluşturma ve uygulamalı deneyimler, sağlık çalışanlarının yapay zekâ sistemlerine olan güvenini artırarak, daha verimli ve güvenli sağlık hizmetlerinin sunulmasına katkıda bulunabilir. Sağlık politikaları açısından, yapay zekâ araştırma ve geliştirme faaliyetlerini geliştirmek ve desteklemek amacıyla fon ve hibelerin sağlanması, sağlık kurumlarında yapay zekâ uygulamalarının entegrasyonu için gerekli altyapıların sağlanması ve sürekli olarak güncellemelerin yapılması gerekmektedir. Yapay zekâ, doğru kullanıldığında sağlık sektöründe önemli bir araç olabilir ve sağlık çalışanlarının bu alandaki farkındalıklarını artırarak bu sürecin başarıya ulaşması sağlanabilir.

Çıkar Beyanı: Bu araştırmada yazarlar arasında çıkar çatışması bulunmamaktadır.

Teşekkürler: Araştırmanın en önemli kısmını oluşturan, elde ettiğimiz verilerle araştırmamıza ışık tutan Bingöl il merkezinde görev yapan sağlık çalışanlarımıza teşekkürlerimizi sunuyoruz.

Yazar Katkıları: Yazarlar araştırmaya eşit oranda katkı sağlamıştır.

REFERANSLAR

[1] J. Shabbir, and T. Anwer, "Artificial Intelligence and Its Role in Near Future." *ArXiv*, 2018, https://arxiv.org/abs/1804.01396.

[2] Shagufta Nesar *et al.*, "Exploring healthcare professionals' knowledge and perceptions of artificial intelligence: Implications for the future of healthcare.," *Int. j. endorsing health sci. res.*, vol. 12, no. 2, pp. 69–75, Mar. 2024, doi: 10.29052/IJEHSR.v12.i2.2024.69-75.

[3] S. Castagno and M. Khalifa, "Perceptions of Artificial Intelligence Among Healthcare Staff: A Qualitative Survey Study," *Front Artif Intell*, vol. 3, p. 578983, Oct. 2020, doi: 10.3389/frai.2020.578983.

[4] A. Lukić, N. Kudelić, V. Antičević, E. Lazić-Mosler, V. Glunčić, D. Hren, I.K. Lukić, "First-Year Nursing Students' Attitudes Towards Artificial Intelligence: Cross-Sectional Multi-Center Study." Nurse Educ Pract. 2023 Aug;71:103735. doi: 10.1016/j.nepr.2023.103735. Epub 2023 Aug 1. PMID: 37541081.

[5] A. Romagnoli, F. Ferrara, R. Langella, and A. Zovi, "Healthcare Systems and Artificial Intelligence: Focus on Challenges and the International Regulatory Framework," *Pharm Res*, vol. 41, no. 4, pp. 721–730, Apr. 2024, doi: 10.1007/s11095-024-03685-3.

[6] K. McGrow, "Artificial intelligence: Essentials for nursing," *Nursing*, vol. 49, no. 9, pp. 46–49, Sep. 2019, doi: 10.1097/01.NURSE.0000577716.57052.8d.

[7] C. Buchanan, M. L. Howitt, R. Wilson, R. G. Booth, T. Risling, and M. Bamford, "Nursing in the Age of Artificial Intelligence: Protocol for a Scoping Review," *JMIR Res Protoc*, vol. 9, no. 4, p. e17490, Apr. 2020, doi: 10.2196/17490

[8] D. Yiğit and A. Açıkgöz, "Evaluation of Comfort Behavior Levels of Newborn by Artificial Intelligence Techniques," *J Perinat Neonatal Nurs*, vol. 38, no. 3, pp. E38–E45, Sep. 2024, doi: 10.1097/JPN.0000000000000768.

[9] G. Bodur, M. Dinçer, Z. Tutak, G. E. Akyüz, S. Uyanik, and D. Kuvan, "Yapay zekânın sağlığın geleceğine etkileri: Üniversite öğrencilerinin gözünden kalitatif çalışma örneği," *Göbeklitepe Medical Science Journal*, vol. 5, no. 7, Art. no. 7, Mar. 2022, doi: 10.55433/gsbd.149.

[10] F. Betriana, R. Tanioka, J. Gunawan, and R. C. Locsin, "Healthcare robots and human generations: Consequences for nursing and healthcare," *Collegian*, vol. 29, no. 5, pp. 767–773, Oct. 2022, doi: 10.1016/j.colegn.2022.01.008.

[11] R. Challen, J. Denny, M. Pitt, L. Gompels, T. Edwards, and K. Tsaneva-Atanasova, "Artificial intelligence, bias and clinical safety," *BMJ Qual Saf*, vol. 28, no. 3, pp. 231–237, Mar. 2019, doi: 10.1136/bmjqs-2018-008370.

[12] G. Krishnan *et al.*, "Artificial intelligence in clinical medicine: catalyzing a sustainable global healthcare paradigm," *Front Artif Intell*, vol. 6, p. 1227091, 2023, doi: 10.3389/frai.2023.1227091.

[13] M. A. Joshi, "Artificial Intelligence (AI) in Healthcare," Feb. 15, 2024, Social Science Research Network, Rochester, NY: 4728304.

[14] M. and M. Albadawy, "AI in diagnostic imaging: revolutionising accuracy and efficiency," Computer Methods and Programs in Biomedicine Update, vol. 5, no. 100146, pp. 1–12, Mar. 2024, doi: 10.1016/j.cmpbup.2024.100146.

[15] W. K. Wan Ismail, P. C. Hong Kit, N. Buhari, and A. Muzaini, "Acceptance of Smartphone in Enhancing Patient-Caregivers Relationship," Journal of Technology Management & Innovation, vol. 7, no. 3, pp. 71–79, Oct. 2012, doi: 10.4067/S0718-27242012000300007.

[16] D. Ilic, "The role of the internet on patient knowledge management, education, and decision-making," Telemed J E Health, vol. 16, no. 6, pp. 664–669, 2010, doi: 10.1089/tmj.2010.0003.

[17] P. Crilly et al., "Community pharmacist perceptions of their role and the use of social media and mobile health applications as tools in public health," Res Social Adm Pharm, vol. 15, no. 1, pp. 23–30, Jan. 2019, doi: 10.1016/j.sapharm.2018.02.005.

[18] N. Al-Shorbaji, "E-health in the Eastern Mediterranean Region: a decade of challenges and achievements," East Mediterr Health J, vol. 14 Suppl, pp. S157-173, 2008.

[19] M. M. Altuwaijri, A. Bahanshal, and M. Almehaid, "Implementation of computerized physician order entry in National Guard Hospitals: assessment of critical success factors," J Family Community Med, vol. 18, no. 3, pp. 143–151, Sep. 2011, doi: 10.4103/2230-8229.90014.

[20] N. Naik et al., "Legal and Ethical Consideration in Artificial Intelligence in Healthcare: Who Takes Responsibility?," Front. Surg., vol. 9, p. 862322, Mar. 2022, doi: 10.3389/fsurg.2022.862322.

[21] E. J. Topol, "High-performance medicine: the convergence of human and artificial intelligence," Nat Med, vol. 25, no. 1, pp. 44–56, Jan. 2019, doi: 10.1038/s41591-018-0300-7.

[22] R. C. Deo, "Machine Learning in Medicine," Circulation, vol. 132, no. 20, pp. 1920–1930, Nov. 2015, doi: 10.1161/CIRCULATIONAHA.115.001593.

[23] C. Krittanawong, H. Zhang, Z. Wang, M. Aydar, and T. Kitai, "Artificial Intelligence in Precision Cardiovascular Medicine," J Am Coll Cardiol, vol. 69, no. 21, pp. 2657–2664, May 2017, doi: 10.1016/j.jacc.2017.03.571.

[24] L. Shinners, S. Grace, S. Smith, A. Stephens, and C. Aggar, "Exploring healthcare professionals' perceptions of artificial intelligence: Piloting the Shinners Artificial Intelligence Perception tool," Digit Health, vol. 8, p. 20552076221078110, 2022, doi: 10.1177/20552076221078110.

[25] M. K. K. Rony, I. Kayesh, S. D. Bala, F. Akter, and M. R. Parvin, "Artificial intelligence in future nursing care: Exploring perspectives of nursing professionals - A descriptive qualitative study," Heliyon, vol. 10, no. 4, p. e25718, Feb. 2024, doi: 10.1016/j.heliyon.2024.e25718.

[26] A. Sauerbrei, A. Kerasidou, F. Lucivero, and N. Hallowell, "The impact of artificial intelligence on the person-centred, doctor-patient relationship: some problems and solutions," BMC Med Inform Decis Mak, vol. 23, no. 1, p. 73, Apr. 2023, doi: 10.1186/s12911-023-02162-y.

[27] S. K. Ahmed, "Artificial intelligence in nursing: Current trends, possibilities and pitfalls," Journal of Medicine, Surgery, and Public Health, vol. 3, p. 100072, Aug. 2024, doi: 10.1016/j.glmedi.2024.100072.

[28] S. J. Fritsch et al., "Attitudes and perception of artificial intelligence in healthcare: A cross-sectional survey among patients," Digit Health, vol. 8, p. 20552076221116772, Aug. 2022, doi: 10.1177/20552076221116772.

[29] A. Gillissen, T. Kochanek, M. Zupanic, and J. Ehlers, "Medical Students' Perceptions towards Digitization and Artificial Intelligence: A Mixed-Methods Study," Healthcare, vol. 10, no. 4, p. 723, Apr. 2022, doi: 10.3390/healthcare10040723.

[30] "T.C. Sağlık Bakanlığı Yönetim Hizmetleri Genel Müdürlüğü Personel Dağılım Cetveli." https://yhgm.saglik.gov.tr/TR,42988/personel-dagilim-cetveli-pdc.html (Erişim Tarihi 23 Şubat 2025).

[31] S. Das, K. Mitra, and M. Mandal, "Sample size calculation: Basic principles," Indian J Anaesth, vol. 60, no. 9, pp. 652–656, Sep. 2016, doi: 10.4103/0019-5049.190621.

[32] Ş. Büyüköztürk, E. K. Çakmak, Ö. E. Akgün, and F. DemiRel, "Bilimsel Araştırma Yöntemleri", 36th ed. Ankara:Pegem, 2024.

[33] Y. Yılmaz, D. Uzelli Yılmaz, D. Yıldırım, E. Akın Korhan, and D. Özer Kaya, "Yapay Zekâ ve Sağlıkta Yapay Zekânın Kullanımına Yönelik Sağlık Bilimleri Fakültesi Öğrencilerinin Görüşleri," Süleyman Demirel Üniversitesi Sağlık Bilimleri Dergisi, vol. 12, no. 3, pp. 297–308, Dec. 2021, doi: 10.22312/sdusbed.950372.

[34] E. Çağlar, "Sağlık Çalışanlarının Yapay Zekâ Farkındalık Düzeylerinin İncelenmesi", Kapadokya Üniversitesi Lisansüstü Eğitim, Öğretim ve Araştırma Enstitüsü Sağlık Yönetimi Anabilim Dalı (Yüksek Lisans Tezi), Nevşehir, 2024.

[35] M. Orhan and A. Bülez, "Evaluation of Health Personnel's Thoughts on Artificial Intelligence," kesit, vol. 8, no. 33, pp. 52–69, 2022, doi: 10.29228/kesit.66404.

[36] H. K. Hoşgör and Ş. Bozkurt, "Sağlıkta Yapay Zekâ ve Robotlar Hakkında Kimler Ne Düşünüyor? Kuşaklar Üzerine Bir Araştırma," vol. 12, pp. 13–25, Mar. 2023.

[37] E. R. Gordon et al., "Ethical considerations for artificial intelligence in dermatology: a scoping review," Br J Dermatol, vol. 190, no. 6, pp. 789–797, May 2024, doi: 10.1093/bjd/ljae040.

[38] Wen, Z.; Huang, H. The potential for artificial intelligence in healthcare. J. Commer. Biotechnol. 2022, 27, 217–224.

[39] S. A. Alowais et al., "Revolutionizing healthcare: the role of artificial intelligence in clinical practice," BMC Med Educ, vol. 23, no. 1, p. 689, Sep. 2023, doi: 10.1186/s12909-023-04698-z.

[40] M. Elnaggar et al., "Assessment of the Perception and Worries of Saudi Healthcare Providers About the Application of Artificial Intelligence in Saudi Health Facilities," Cureus, vol. 15, no. 8, p. e42858, Aug. 2023, doi: 10.7759/cureus.42858.

[41] K. Kimiafar et al., "Artificial Intelligence Literacy Among Healthcare Professionals and Students: A Systematic Review," Front Health Inform, vol. 12, p. 168, Nov. 2023, doi: 10.30699/fhi.v12i0.524.

[42] G. Yorgancıoglu Tarcan, P. Yalçin Balçik, and N. B. SebiφK, "Türkiye ve Dünyada Sağlık Hizmetlerinde Yapay Zekâ," Mersin Üniversitesi Tıp Fakültesi Lokman Hekim Tıp Tarihi ve Folklorik Tıp Dergisi, vol. 14, no. 1, pp. 50–60, Jan. 2024, doi: 10.31020/mutftd.1278529.

[43] F. Wang, L. P. Casalino, and D. Khullar, "Deep Learning in Medicine-Promise, Progress, and Challenges," JAMA Intern Med, vol. 179, no. 3, pp. 293–294, Mar. 2019, doi: 10.1001/jamainternmed.2018.7117.

[44] E. Karal and M. Turan, "Hekime Tanı Koymada Yardımcı, Yapay Zekâ Destekli Hastalık Tespit Uzmanı," European Journal of Science and Technology, Jun. 2021, doi: 10.31590/ejosat.945518.

[45] S.-C. Kong, W. Man-Yin Cheung, and G. Zhang, "Evaluation of an artificial intelligence literacy course for university students with diverse study backgrounds," Computers and Education: Artificial Intelligence, vol. 2, p. 100026, Jan. 2021, doi: 10.1016/j.caeai.2021.100026.

[46] Y. Çilhoroz and O. Işık, "Yapay Zekâ: Sağlık Hizmetlerinden Uygulamalar," AHBVÜ İİBF Dergisi, vol. 23, no. 2, Art. no. 2, Aug. 2021.

[47] M. Alelyani et al., "Radiology Community Attitude in Saudi Arabia about the Applications of Artificial Intelligence in Radiology," Healthcare (Basel), vol. 9, no. 7, p. 834, Jul. 2021, doi: 10.3390/healthcare9070834.

[48] Z. Hamd, W. Elshami, S. Al Kawas, H. Aljuaid, and M. M. Abuzaid, "A closer look at the current knowledge and prospects of artificial intelligence integration in dentistry practice: A cross-sectional study," Heliyon, vol. 9, no. 6, p. e17089, Jun. 2023, doi: 10.1016/j.heliyon.2023.e17089.

[49] H. Liyanage et al., "Artificial Intelligence in Primary Health Care: Perceptions, Issues, and Challenges," Yearb Med Inform, vol. 28, no. 1, pp. 41–46, Aug. 2019, doi: 10.1055/s-0039-1677901.

[50] C. J. Wiedermann, A. Mahlknecht, G. Piccoliori, and A. Engl, "Redesigning Primary Care: The Emergence of Artificial-Intelligence-Driven Symptom Diagnostic Tools," J Pers Med, vol. 13, no. 9, p. 1379, Sep. 2023, doi: 10.3390/jpm13091379.

[51] H. S. J. Chew and P. Achananuparp, "Perceptions and Needs of Artificial Intelligence in Health Care to Increase Adoption: Scoping Review," J Med Internet Res, vol. 24, no. 1, p. e32939, Jan. 2022, doi: 10.2196/32939.

[52] A. Alsaedi et al., "Perceived Worries in the Adoption of Artificial Intelligence Among Healthcare Professionals in Saudi Arabia: A Cross-Sectional Survey Study," Nursing Reports, vol. 14, no. 4, pp. 3706–3721, Nov. 2024, doi: 10.3390/nursrep14040271.

[53] W. Al-Qerem et al., "Exploring knowledge, attitudes, and practices towards artificial intelligence among health professions' students in Jordan," BMC Med Inform Decis Mak, vol. 23, no. 1, p. 288, Dec. 2023, doi: 10.1186/s12911-023-02403-0

SERENDIPITY

Arda Çildan*
Maltepe Üniversitesi, Yazılım Mühendisliği, İstanbul, Türkiye

Ömer Çevrük
Maltepe Üniversitesi, Yazılım Mühendisliği , İstanbul, Türkiye

Emre Atlıer Olca
Maltepe Üniversitesi, Yazılım Mühendisliği, İstanbul, Türkiye
(Corresponding author) arda.cildan@gmail.com

ÖZET

Serendipity projesi, Endüstri 4.0 prensipleri doğrultusunda endüstriyel operasyonları dönüştürmek için tasarlanmış gelişmiş bir kestirimci bakım sistemidir. Bu kapsamlı sistem, endüstriyel makinelerden gelen gerçek zamanlı sensör verileri, vardiya düzenleri, çevresel koşullar ve operasyonel parametreler dahil olmak üzere çoklu kaynaklardan gelen karmaşık veri akışlarını analiz etmek için makine öğrenimi ve yapay zeka teknolojilerinden yararlanır. Temel hedef, ekipman arızalarını öngörmek ve önlemek, bakım programlarını optimize etmek ve endüstriyel tesislerde işyeri güvenliğini artırmaktır.

Sistem mimarisi, Endüstriyel Nesnelerin İnterneti (IIoT) sensörleri, bulut bilişim altyapısı ve gelişmiş analitik motorları dahil olmak üzere çoklu teknolojik katmanları içerir. Bu bileşenler, gerçek zamanlı izleme ve tahmin yetenekleri sağlamak için uyum içinde çalışır. Veri işleme hattı, sensör ölçümlerinin sürekli akışını işler, bilgileri temizler ve standartlaştırır, ardından geçmiş arıza verileri üzerine eğitilmiş makine öğrenimi modellerine besler. Endişe verici desenler tespit edildiğinde, sistem otomatik olarak uyarılar ve bakım önerileri oluşturur.

Maliyet azaltma, erken tespit yoluyla beklenmeyen makine arızalarını önleme, gereksiz müdahaleleri azaltmak için bakım programlamasını optimize etme, parçaları tahmini ömürler yerine gerçek duruma göre değiştirme ve gelişmiş ekip programlaması yoluyla işçilik maliyetlerini düşürme gibi çeşitli mekanizmalarla sağlanır. Sistem ayrıca optimal makine performansını koruyarak ve ekipman bozulmasından kaynaklanan kusur riskini azaltarak üretim kalitesinin artmasına katkıda bulunur.

Çalışan güvenliği, güvenlik açısından kritik parametrelerin sürekli izlenmesi ve potansiyel tehlikeli koşullar tespit edildiğinde anında uyarı üretilmesi yoluyla artırılır. Sistem, geçmiş güvenlik olaylarının öncesinde ortaya çıkan desenleri tanımlamak için geçmiş verileri analiz eder ve benzer durumların proaktif olarak önlenmesini sağlar. Projenin modüler mimarisi, artan veri hacimleri ve yeni ekipman eklemeleri için ölçeklenebilirliği sağlar. Bulut bilişim kaynakları esnek işlem kapasitesi sağlarken, modüler tasarım gerektiğinde yeni makine öğrenimi modellerinin ve veri kaynaklarının entegrasyonuna olanak tanır.

Kullanıcı arayüzü, bakım ekipleri, operatörler ve yönetim dahil olmak üzere birden çok paydaş grubuna hizmet vermek üzere tasarlanmış olup, her birine ilgili içgörüler ve uygulanabilir bilgiler sağlar. Sistem sürekli olarak yeni verilerden ve geri bildirimlerden öğrenir ve zamanla tahmin doğruluğunu artırır. Bu yenilikçi yaklaşım, endüstriyel tesisleri daha verimli, daha güvenli ve veri odaklı bir operasyonel modele doğru yönlendirerek geleneksel bakım uygulamalarına göre önemli bir ilerlemeyi temsil eder.

Anahtar Kelimeler: Kestirimci bakım, endüstri 4.0, makine öğrenimi, yapay zeka, sinir ağı.

Giriş

Kestirimci bakım sistemi, Endüstri 4.0 prensipleri doğrultusunda endüstriyel ortamlarda ekipman arızalarını önlemek için geliştirilmiştir. Sistem, makinelerden gelen sensör verilerini işleyerek potansiyel arızaları oluşmadan önce tahmin eder. Makine verilerini yapay zeka ile analiz ederek bakım programları optimize edilir ve endüstriyel tesislerde operasyonel maliyetler azaltılır [1].

Makine ve vardiya verileri, çevresel koşullar ve bakım geçmişi bir arada analiz edilerek arızaların ne zaman meydana gelebileceğine dair doğru tahminler üretilir. Bu, geleneksel bakım yaklaşımlarına kıyasla önemli bir iyileştirme sunar. Geleneksel yöntemlerde makineler ya arıza yaptıktan sonra onarılır ya da durumlarından bağımsız olarak sabit bir programa göre bakım yapılır [2]. Mevcut reaktif yaklaşım, gereksiz duruş sürelerine ve bakım maliyetlerine yol açarken, planlı bakım genellikle parçaların erken değiştirilmesine neden olur.

Sistem mimarisi, birden fazla sensör kaynağından gelen gerçek zamanlı veri işleme işlemini gerçekleştirecek şekilde tasarlanmıştır. Ham veriler makinelerden sürekli olarak toplanır ve gürültü ile tutarsızlıkların giderilmesi için ön işleme tabi tutulur [3]. Temizlenen veriler, geçmiş arıza verileriyle eğitilmiş makine öğrenmesi (ML) modellerine beslenir. Bu modeller, potansiyel gelecekteki arızaları işaret eden desenleri tanımlar. Endişe verici desenler tespit edildiğinde, bakım ekiplerine bildirim göndermek üzere otomatik olarak uyarılar oluşturulur.

Maliyet azaltımı birkaç mekanizma ile sağlanır. Beklenmedik makine arızaları, sorunların erken tespitiyle önlenir [4]. Bakım, sabit bir programa göre değil, yalnızca ihtiyaç duyulduğunda gerçekleştirilir. Parçalar tahmini ömürlerine göre değil, gerçek durumlarına göre değiştirilir [5]. Bakım ekiplerinin planlanması optimize edilerek işçilik maliyetleri azaltılır. Üretim kalitesi, makinelerin optimum performansını koruyarak iyileştirilir.

İşçi güvenliği, tehlikeli makine koşulları kazaya neden olmadan önce belirlenerek artırılır. Sistem, güvenlik açısından kritik parametreleri izler ve güvensiz koşullar tespit edilirse anında uyarılar oluşturur. Geçmiş veriler, önceki güvenlik olaylarından önceki desenleri tanımlamak için analiz edilir ve benzer durumların gelecekte önlenmesi sağlanır.

Projenin Amaç ve Önemi

Serendipity projesi, üretim sektöründe verimliliği artırmayı ve kazaları önlemeyi amaçlayan yenilikçi bir yaklaşım sunmaktadır. Proje, makineler ve bileşenler için bakım zamanlarını

doğru bir şekilde tahmin ederek üretim sürekliliğini sağlarken maliyetleri optimize etmeye odaklanmaktadır. Bu yaklaşımla fabrikalardaki bakım süreçleri daha sistematik hale gelecek, gereksiz duruşlar önlenecek ve makinelerin ömrü uzatılacaktır [6].

Proje, aynı zamanda makinelerin performansını sürekli izleyerek arıza risklerini en aza indirmeye yönelik veri odaklı bir yaklaşım benimsemektedir. Makine verileri ve bakım geçmişi kullanılarak oluşturulan algoritmalar, yalnızca gerektiğinde bakım önerecek ve bu da iş gücü kaybını ve bakım maliyetlerini azaltacaktır. Sistem, yönetim birimlerine makinelerin durumu hakkında kapsamlı bilgiler sağlayarak daha bilinçli ve stratejik kararların alınmasını mümkün kılacaktır.

Serendipity, sektöre güvenlik açısından da önemli bir değer katmaktadır. Arıza olasılığı yüksek olan bileşenler veya riskli durumlar için erken uyarılar sunarak iş güvenliğini artıracak ve kazaların önlenmesine yardımcı olacaktır [7]. Kestirimci bakım teknolojisinin fabrikalarda yaygın olarak benimsenmesi, işçi sağlığını riske atan koşulların önlenmesine katkı sağlayarak daha güvenli bir çalışma ortamı yaratacaktır.

Projenin diğer hedefleri arasında üretim verimliliğini artırmak, maliyetleri düşürmek ve fabrikalardaki operasyonel süreçleri iyileştirmek yer almaktadır. Kestirimci bakım, fabrikaların planlı bakım yerine ihtiyaç bazlı bakım yapmasını sağlayarak kaynak tasarrufunu teşvik edecek ve üretim süreçlerini daha esnek ve verimli hale getirecektir.

Proje, fabrikaları daha akıllı ve bağlantılı hale getirmek için yeni teknolojilerden faydalanmaktadır. Sensörler ve veri analizi kullanılarak sorunlar henüz ortaya çıkmadan tespit edilmekte ve bu sayede daha iyi bir planlama ve sürpriz arızaların önlenmesi sağlanmaktadır. Bu modern yaklaşım, fabrikaların eski moda bakım yöntemlerinden uzaklaşarak daha akıllı ve verimli bir çalışma biçimine geçiş yapmasına yardımcı olmaktadır.

Serendipity sistemi, fabrikalar büyüdükçe kolayca uyum sağlayabilecek şekilde tasarlanmıştır. Fabrikalar yeni makineler eklediğinde veya çalışma düzenlerini değiştirdiğinde, sistem hızla uyum sağlar. Ayrıca, sistem zamanla daha fazla veri öğrenerek arıza tahminlerinde daha iyi hale gelir ve bu da fabrikaların daha sorunsuz ve güvenli bir şekilde çalışmasını sağlarken maliyetlerin düşürülmesine yardımcı olur.

Sistem, artan veri hacmini yönetmek için ölçeklenebilir bir tasarımla geliştirilmiştir. Bulut bilişim kaynakları, esnek işlem kapasitesi sağlamak için kullanılır. Modüler mimari, gerektiğinde yeni makine öğrenmesi modellerinin ve veri kaynaklarının eklenmesine olanak tanır. Toplanan verilerden ek içgörüler sağlamak için analitik yetenekler genişletilebilir.

Projenin Yenilik ve Özgünlük Unsuru

Fabrikalarda makine arıza olasılığını, iş kazası risklerini ve ürün kalitesini etkileyen faktörler ele alınmaktadır. Makinelerin ve parçaların kullanım süreleri, işçilerle etkileşimleri, çevresel koşullar ve üretim süreçleri analiz edilerek arıza tahminleri yapılması hedeflenmektedir. Bu tahminler sayesinde bakım operasyonları ve parça değişimleri zamanında gerçekleştirilerek verimliliği artırmak, iş kazalarını azaltmak ve ürün kalitesini iyileştirmek amaçlanmaktadır.

Günümüzde fabrikalar, makine bakımlarını ya makineler tamamen bozulana kadar bekleyerek ya da arıza için hâlâ zaman varken bakım ve parça değişimleri yaparak çözmek-

tedir. Ancak her iki çözüm de başka problemleri beraberinde getirmektedir. Bir makineyi işlevini kaybedene kadar kullanmaya devam etmek, makinenin ömrünün son aşamalarında düşük verimlilikle çalışmayı ve düşük kaliteli üretimi kabul etmek anlamına gelir. Öte yandan, tam kapasite çalışmaya devam eden bir makineye müdahale etmek, makinenin tam kapasite çalışabileceği bir dönemi kaybetmek anlamına gelir. Kestirimci bakım (PM), parça bakımının en uygun zamanını tahmin ederek her iki probleme de çözüm sunar. Sağlayacağı verimlilik artışı, sektör açısından orijinal bir değer olarak vurgulanmaktadır [8].

Birçok çalışma, PM'nin bakım takvimlerini optimize etmek ve makine duruş sürelerini azaltmak için etkili bir çözüm olduğunu göstermektedir. PM üzerine yapılan bir çalışmada [9], kestirimci bakımda kullanılan makine öğrenimi yöntemleri üzerine sistematik bir literatür incelemesi yapılmış ve çeşitli makine öğrenimi modellerinin bakım ihtiyaçlarını tahmin etmek ve operasyonel kesintileri azaltmak için nasıl kullanılabileceği vurgulanmıştır. Benzer şekilde, bir başka çalışma [10], Endüstri 4.0 bağlamında PM'nin genel bir değerlendirmesini yapmış ve farklı endüstriyel sektörlerde PM teknolojilerinin uygulanmasıyla ilgili zorlukları ve modelleri incelemiştir. Bu çalışmalar, kestirimci bakımın gerçek zamanlı veriyle entegrasyonunun arıza tahminlerinin doğruluğunu önemli ölçüde artırabileceğini ve hem verimliliği hem de güvenliği iyileştirebileceğini desteklemektedir. Serendipity projesi, bu bulgulara dayanarak, makinelerin ömrünü uzatmayı ve kazaları azaltmayı hedeflemekte; bakım müdahaleleri için en uygun anları tahmin ederek bu çalışmalarda belirtilen zorlukları ele almaktadır.

Serendipity projesinin temel araştırma sorusu, "Makine ve vardiya verilerini kullanan PM yöntemleri, üretim süreçlerindeki arıza ve kaza risklerini azaltmada ne kadar etkilidir?" olarak belirlenmiştir. Projenin hipotezi, kestirimci bakım algoritmalarının üretim sürecinde uyarılar sağlayarak makine arızalarını ve iş kazalarını en aza indirirken üretim verimliliği üzerinde olumlu bir etki yaratacağıdır.

Serendipity projesi, yalnızca algoritmalar kullanarak makine verilerini analiz edip parça arızalarını erken tespit ederek makinelerin ömrünü uzatmakla kalmaz, aynı zamanda kaza risklerini de azaltır. Bu yaklaşım, üretim planlarının kesintisiz ve daha güvenli bir şekilde yürütülmesini sağlarken, beklenmedik duruşları önleyerek fabrikanın genel verimliliğini artırır.

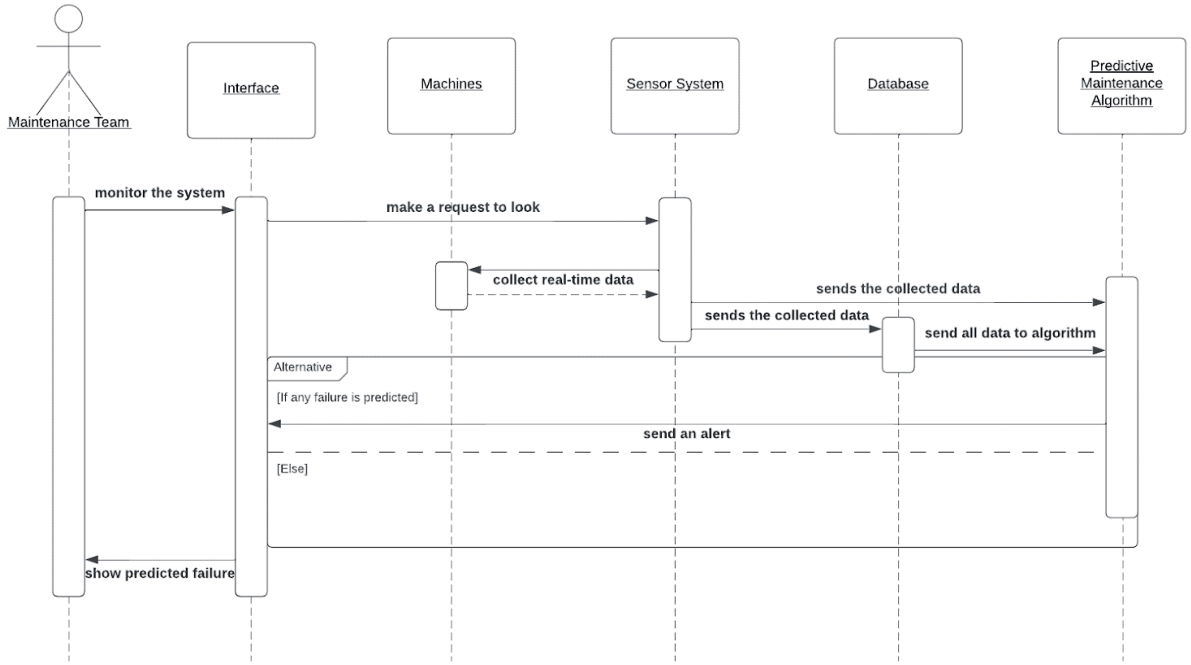
Serendipity, sektöre sağlayacağı verimlilik ve güvenlik iyileştirmeleri açısından önemli bir yenilik sunmaktadır. Kestirimci bakım teknolojilerinin endüstride yaygın olarak benimsenmesi, makinelerin ömrünü uzatacak ve insan sağlığını riske atan durumları önlemeye katkı sağlayacaktır.

Serendipity'nin Sistem Mimarisi

Kestirimci bakım sistemi, fabrika makinelerinden ve çalışan vardiya düzenlerinden gelen verileri analiz ederek, sorunlar ortaya çıkmadan önce tespit etmek amacıyla geliştirilmiştir. Bu sistem üç ana hedefe odaklanmaktadır: makine arızalarını önlemek, fabrika çalışma ortamını daha güvenli hale getirmek ve farklı vardiya düzenlerinin makine performansını nasıl etkilediğini anlamak. Sistem, makinelerin ne zaman sorun yaşayabileceğini belirlemek için yapay zeka ve makine öğrenimi kullanmaktadır.

Sistemimize giren veriler iki ana kaynaktan gelmektedir: makine sensörleri ve vardiya yönetim sistemleri. Makine verileri; sıcaklık, titreşim ve basınç ölçümlerini içerirken, vardiya

verileri; çalışanların programlarını, operatör atamalarını ve üretim dönemlerini kapsamaktadır. Bu birleşik yaklaşım, insan faktörleri ile makine operasyonlarının birbirlerini nasıl etkilediğini anlamamıza yardımcı olmaktadır [11].

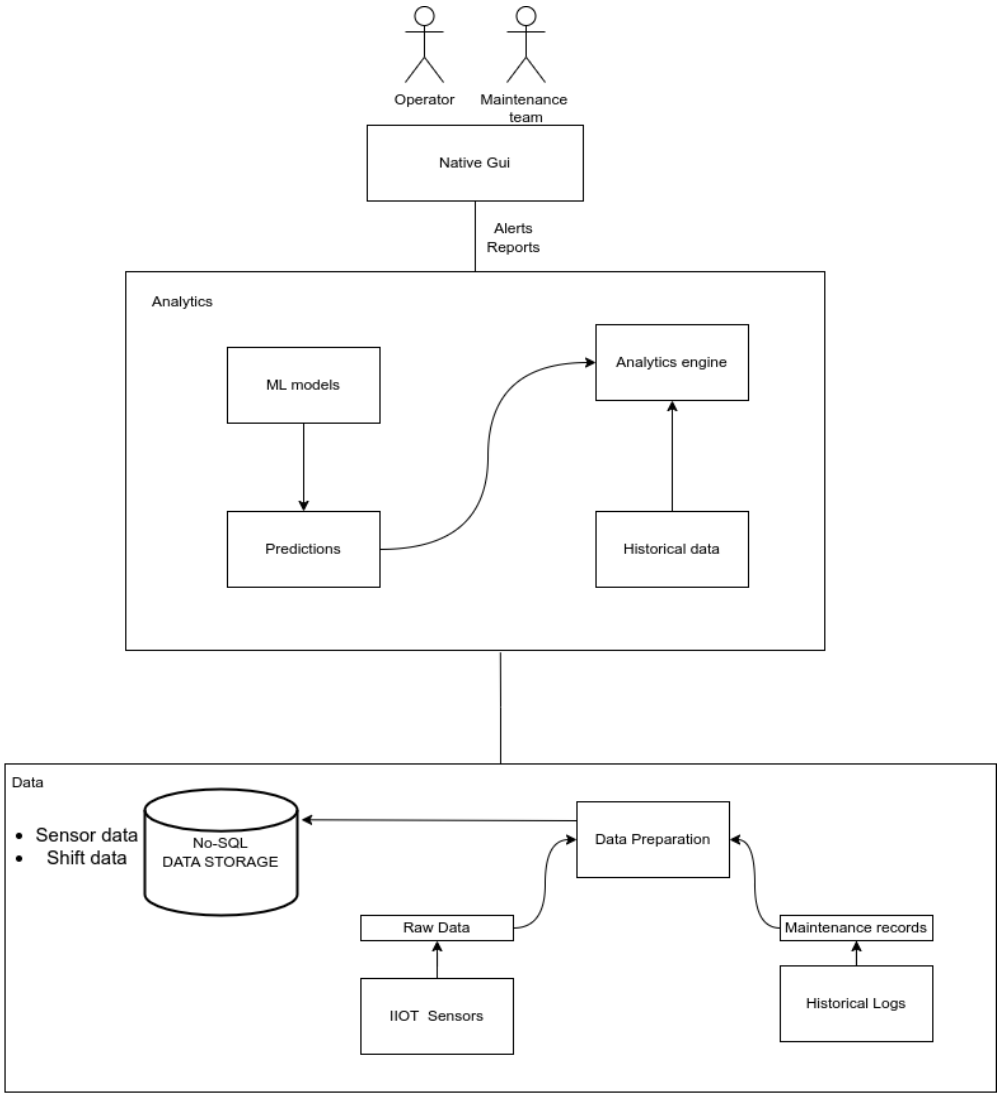


Şekil 1: Serendipity' nin Sequence Diyagramı

Şekil 1'de görülen Sequence Diyagramı Serendipity sisteminin çalışma durumunu göstermektedir. Süreç, bakım ekibinin bir arayüz aracılığıyla sistemi izlemeye başlamasıyla başlar. Makineler hakkında bilgi almak istediklerinde, sistemi görünür hale getiren bir talep tetiklenir. Sensör sistemi, makinelerden sürekli olarak veri toplar. Toplanan bu veriler, daha sonra veritabanına gönderilir ve buradan tüm ilgili bilgiler bakım algoritmasına iletilir. Eğer algoritma potansiyel bir arıza öngörürse, bir uyarı gönderilir; herhangi bir arıza tespit edilmezse, süreç herhangi bir bildirim olmaksızın sona erer. Diyagram, Serendipity sisteminin sürekli izlenmesini ve proaktif uyarılarını vurgulamaktadır.

Veri Toplama ve İşleme Hattı

Veri Toplama Sistemi, makine sensör bilgilerini ve vardiya verilerini gerçek zamanlı olarak işler. Bu sistem, makine performans ölçümleri ile farklı vardiyalar sırasında operatör etkileşim verilerini toplar. Ayrıca, geçmiş bakım faaliyetlerini ve vardiya düzenlerini kaydederek önceki deneyimlerden öğrenir. Bu veri kaynaklarını birleştirerek, farklı vardiyaların ve operatörlerin makine performansını nasıl etkileyebileceğini görebiliriz [12].



Şekil 2: Serendipity' nin High Level Design Diyagramı

Şekil 2'de görülen High Level Dizaynı sistemin mimarisini göstermektedir. Bu mimari, IIoT sensörlerinden ve geçmiş kayıtlarından gelen verileri bir veri toplama katmanında işler. Toplanan veriler, hem sensör hem de vardiya verilerini işleyebilen bir NoSQL depolama sistemine aktarılır. Analitik katmanı, tahminler oluşturmak için makine öğrenimi modellerini kullanırken, analitik motoru geçmiş verileri işler. En üstte, bir yerel grafik kullanıcı arayüzü (GUI), operatörlere ve bakım ekiplerine sistemin analizlerine dayalı olarak uyarılar ve raporlar sunar [13].

İşleme Hattı, makine ve vardiya verilerinin hazırlanmasını yönetir. Sensör okumalarını temizler ve vardiya bilgilerini sistemin analiz edebileceği bir formata dönüştürür. Bu bölüm, işçi düzenleri ile makine sağlığı arasındaki bağlantıları görmemize yardımcı olduğu için kritik öneme sahiptir [14].

Makine Öğrenimi Motoru

Makine Öğrenimi Motoru, hem makine durumlarını hem de vardiya düzenlerini analiz ederek olası arızalar hakkında tahminlerde bulunur. Örneğin, belirli vardiyaların belirli ma-

kinelerde daha fazla sorun yaşayıp yaşamadığını veya makinelerin kullanım şekline göre bakım ihtiyaçlarının nasıl değiştiğini öğrenebilir. Modeller, yeni verilerle sürekli olarak öğrenerek daha iyi tahminlerde bulunur.

Tahmin ve Raporlama Sistemi

Tahmin Sistemi, hem makineler hem de vardiyalarla ilgili bilgileri kullanarak bakımın ne zaman gerekli olduğunu belirler. Vardiya değişiklikleri, operatör deneyim seviyeleri ve makine kullanım düzenleri gibi faktörleri göz önünde bulundurur. Potansiyel sorunları tespit ettiğinde, vardiya programlarını dikkate alarak bakım için en uygun zamanı önerir.

Raporlama Arayüzü, bakım ekiplerine ve vardiya yöneticilerine ihtiyaç duydukları bilgileri sunar. Hem makine durumu hem de vardiyalara ilişkin bilgileri gösterir. Ekipler, farklı vardiyaların makine performansını nasıl etkilediğini görebilir ve bakımı en uygun zamanlarda planlayabilir. Arayüz ayrıca bakım ekipleri ile vardiya çalışanları arasında koordinasyon sağlar.

Sistem, fabrika operasyonlarının insani unsurlarını dikkate alırken, çeşitli makine türleriyle çalışacak şekilde tasarlanmıştır. Mevcut fabrika sistemleri, özellikle vardiya yönetim yazılımları ile entegre olarak operasyonların tam bir görünümünü sunar. Bu kapsamlı yaklaşım, arızaları önlerken bakımın vardiya düzenlerine göre pratik bir şekilde gerçekleştirilmesini sağlar [15].

Sonuç

Bu çalışma, kestirimci bakım sistemlerinin sanayi ortamlarında makine arızalarını önleme, maliyetleri optimize etme ve iş güvenliğini artırma konularında sağladığı faydaları ortaya koymaktadır. Serendipity projesi kapsamında geliştirilen sistem, endüstri 4.0 prensiplerine dayanarak makine sensörlerinden gelen verileri analiz etmekte ve yapay zeka destekli tahmin modelleri ile bakım süreçlerini optimize etmektedir.

Serendipity projesi, makine öğrenimi algoritmalarının makine ve vardiya verileri ile entegre edilerek bakım süreçlerinin daha öngörülebilir ve verimli hale getirilebileceğini göstermiştir. Yapılan testler ve analizler, geleneksel bakım yaklaşımlarına kıyasla kestirimci bakımın hem operasyonel verimlilik hem de maliyet açısından üstün olduğunu kanıtlamıştır.

Sistem, ölçeklenebilir bir mimariye sahip olup, fabrikaların değişen üretim süreçlerine ve yeni makine entegrasyonlarına kolayca uyum sağlayabilecek şekilde tasarlanmıştır. Ayrıca, sistem zamanla daha fazla veri işleyerek tahmin doğruluğunu artırma kapasitesine sahiptir.

Sonuç olarak, kestirimci bakım teknolojileri, sanayi sektöründe üretim sürekliliğini sağlamak, iş güvenliğini artırmak ve maliyetleri minimize etmek adına önemli bir dönüşüm sunmaktadır. Serendipity projesi, bu dönüşümün etkin bir şekilde nasıl gerçekleştirilebileceğini göstermiş olup, gelecekte daha geniş kapsamlı uygulamalara entegre edilerek endüstride yaygınlaşma potansiyeline sahiptir.

REFERANSLAR

[1] Achouch, M., Dimitrova, M., Ziane, K., Karganroudi, S. S., Dhouib, R., Ibrahim, H., and Adda, M., "On Predictive Maintenance in Industry 4.0: Overview, Models, and Challenges", International Journal of Information Technology and Computer Science, 15, 2, 48-63, 2022.

[2] Aydın, Ö., Akın, E., "Endüstri 4.0 Tabanlı Üretim Sistemleri ve Kestirimci Bakım Yaklaşımlarında Dijital İkiz Uygulamalarının İncelenmesi", *Fırat Üniversitesi Fen Bilimleri Dergisi*, 33(2), 167-178, 2021.

[3] Erez, A. S., Çakmakçı, M., "Endüstri 4.0 Kestirimci Bakım Süreçlerinde İsteme Özgü MEMS Tabanlı Sensörlerin Aksiyomatik Tasarımla Seçim Problemi", Yüksek Lisans Tezi, Dokuz Eylül Üniversitesi Fen Bilimleri Enstitüsü, 2020.

[4] de Freitas Alves, F., Gómez Ravetti, M., "Hybrid Proactive Approach for Solving Maintenance and Planning Problems in the Scenario of Industry 4.0", 2020.

[5] Windmann, A., Wittenberg, P., Schieseck, M., and Niggemann, O., "Artificial Intelligence in Industry 4.0: A Review of Integration Challenges for Industrial Systems", 2024

[6] Yılmaz, A., Kaya, B., "Endüstri 4.0 Tabanlı Üretim Sistemleri ve Kestirimci Bakım Uygulamaları", *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 33(2), 2021.

[7] Soylu, B., Yiğiter, H., Sarıkaya, V., Sandıkçı, Z., and Utku, A., "Kestirimci Bakım Planlama için Makine Öğrenmesi Temelli Bir Karar Destek Sistemi ve Bir Uygulama", *Verimlilik Dergisi*, 2022.

[8] Carvalho, T. P., Soares, F. A. A. M. N., Vita, R., Francisco, R. P., Basto, J. P., and Alcalá, S. G. S., "A systematic literature review of machine learning methods applied to predictive maintenance", Computers & Industrial Engineering, 137, 106024, 2019.

[9] Dündar, D. R., Sariçiçek, İ., Çinar, E., and Yazici, A., "Kestirimci Bakımda Makine Öğrenmesi: Literatür Araştırması", Eskişehir Türk Dünyası Uygulama ve Araştırma Merkezi Bilişim Dergisi, 2, 2, 44-55, 2021.

[10] Fordal, J. M., Schjølberg, P., Helgetun, H., Skjermo, T. Ø., Wang, Y., and Wang, C., "Application of sensor data based predictive maintenance and artificial neural networks to enable Industry 4.0", Journal of Industrial Information Integration, 28, 100333, 2023.

[11] Kumar, A., Singh, R., and Sharma, P., "Predictive Maintenance in Industry 4.0: Technologies and Challenges", *Journal of Industrial Engineering and Management*, 2021.

[12] Awasar, P., & Shankar, R., "Real-Time Data Analytics for Predictive Maintenance in Industrial Applications", *International Journal of Advanced Manufacturing Technology*, 2018.

[13] Wang, L., Liu, Z., "A Survey on Industrial Internet of Things and its Application in Predictive Maintenance", *Journal of Industrial Information Integration*, 21, 100179, 2020.

[14] Arslan, H., & Kara, S., "Fabrika Otomasyonu ve Kestirimci Bakım Uygulamaları", *Türk Elektrik Mühendisliği Dergisi*, 29(5), 120-134, 2020.

[15] Kılınç, G., Demir, D., "Endüstri 4.0 ve Kestirimci Bakım: Entegrasyon ve Uygulamalar", *Mühendislik Bilimleri Dergisi*, 15(4), 85-95, 2022.

SİBER GÜVENLİKTE ANOMALİ TESPİTİ: CIC-IDS2017 VE CIC DDOS2019 VERİ SETLERİ İLE MAKİNE ÖĞRENMESİ TABANLI YÖNTEMLERİN KARŞILAŞTIRMALI PERFORMANS ANALİZİ

Alperen Tokgöz*
Fırat Üniversitesi, Teknoloji Fakültesi, Elazığ, Türkiye

Özal Yıldırım
Fırat Üniversitesi, Teknoloji Fakültesi, Elazığ, Türkiye
(İletişimden sorumlu yazar) 221137107@firat.edu.tr

ÖZET

Teknolojinin gelişmesiyle birlikte, bankacılık sistemlerine erişim, bilgi depolarının internet ortamına taşınması gibi gelişmeler, siber güvenliğin önemini artırmış ve kritik bir konu haline getirmiştir. Veri ve sistemlerin güvenliği daima öncelikli bir ihtiyaçtır. Ayrıca, kullanıcıları, kurumları ve kuruluşları hedef alan saldırganlar, sistemlerin açıklarını tespit etmek için sürekli çaba göstermektedir. Saldırganların hedefleri arasında eğitim/araştırma, devlet/askeri, sağlık, telekomünikasyon ve bankacılık gibi çok çeşitli alanlar yer almaktadır. Bu kadar geniş bir alanda yapay zeka teknolojilerinin kullanılması önem taşımaktadır. Yapay zeka teknolojilerinin uygulanabilmesi için kaliteli veri setlerine ihtiyaç duyulmaktadır. Gerçek saldırıların anında tespiti ve kaydedilmesi güç olduğundan, Kanada Siber Güvenlik Enstitüsü tarafından laboratuvar ortamında hazırlanan CIC-IDS2017 veri seti ile XGBoost, Random Forest ve LightGBM modellerinin karşılaştırması yapılmıştır. Ayrıca, modellerin yeteneklerini test etmek amacıyla, CIC-IDS2017 veri setinde bulunmayan saldırı türlerini içeren daha kapsamlı CIC-DDoS2019 veri setiyle doğrulama yapılmıştır. Yapılan deneysel çalışmalarda Random Forest modelinin diğer modellerden anomali tespitinde daha başarılı sonuçlar elde ettiği görülmüştür.

Anahtar Kelimeler: Anomali tespiti, siber güvenlik, XGBoost, random forest, LightGBM.

Giriş

Güvenlik ihtiyacı, geçmişte olduğu gibi günümüzde ve gelecekte de her zaman ön planda olacaktır. İnternetin finansal, kişisel ve kurumsal alanlarda yaygın kullanımı, sistemlerin korunaklı olmasını zorunlu kılmaktadır. İyi niyetli kullanıcıların yanı sıra, kötü niyetli aktörlerde sistem açıklarını tespit etmek için sürekli çaba göstermektedir. Ağ, bankacılık, sosyal medya gibi kritik alanlarda güvenlik ihtiyacı en önemli konular arasında yer almaktadır. Örneğin, yazılımlara eklenen şifre kuralları, şifrelerin daha güvenli yöntemlerle Veritabanında saklanması, son kullanıcıları saldırganlardan korumak sağlamak amacıyla geliştirilmiştir. Ancak güvenlikle ilgili alınan tüm önlemlere rağmen siber saldırıya maruz kalma riski her

zaman bulunmaktadır. Siber saldırılar, verilerin ele geçirilmesi, değiştirilmesi veya silinmesi gibi kötü niyetli işlemleri gerçekleştirmek için çeşitli yöntemler kullanılarak bilgi sistemlerine, altyapılara ve kullanıcı cihazlarına yönelik saldırıları ifade etmektedir. Bu saldırıların hedefi bireysel kullanıcıların yanı sıra kurumsal şirketler ve devlet kurumları da olabilmektedir [1]. Bilgi güvenliğinin sağlanması büyük önem taşımakta olup, bu amaçla ülkemizde ISO 27001 standardı uygulanmaktadır.

Siber saldırıların çeşitli türleri bulunmaktadır. Avrupa Birliği Siber Güvenlik Ajansı(ENISA) 2024 raporuna göre, Temmuz 2023 - Haziran 2024 döneminde tehdit dağılımı %41.1 oranında DoS/DDoS/RDoS saldırıları ve %25.79 oranında ransomware (fidye yazılımı) saldırıları şeklinde gerçekleşmiştir. DoS saldırılarının daha çok ideolojik amaçlarla, fidye saldırılarının ise finansal amaçlarla yapıldığı gözlemlenmektedir [2]. DoS saldırıları, hedeflenen kişi, kurum veya kuruluşun bant genişliği kapasitesini aşacak şekilde yoğun veri gönderimi yaparak sistemi kullanılamaz hale getirmeyi amaçlamaktadır. Ransomware saldırılarında ise saldırganlar, finansal kazanç sağlamak amacıyla kullanıcı verilerini şifreleyerek fidye talep etmektedir. Saldırganlar bütün sistemi kullanılamaz hale getirmek yerine belirli verileri de şifreleyerek karşılığında ödeme talep edebilirler. Bu tür saldırılarda ödeme gerçekleştirilene kadar dosyalar hiçbir şekilde kullanılamamaktadır [3].

Saldırıların önceden tespiti, sistemleri korunmak adına büyük önem taşımaktadır. Bu amaçla uzmanlar tarafından geliştirilen saldırı tespit sistemleri (Intrusion Detection System - IDS), sisteme yönelik herhangi bir saldırı gerçekleştiğinde anomali durumunu tespit edebilen donanım ve yazılım tabanlı güvenlik katmanlarıdır [4]. Ağ trafiğindeki normal davranıştan farklı anormallikleri tespit ederek, bilinen ve bilinmeyen saldırıları tespit etmeyi hedefleyen anomali tabanlı saldırı tespit sistemleri (IDS) kritik bir rol üstlenmektedir. Bu çalışmada XGBoost, Random Forest ve LightGBM algoritmaları kullanılarak anomali tespiti amacıyla modellerin karşılaştırmalı analizi yapılmıştır. Çalışma öncelikle literatür taraması, yöntem ve teknik açıklamalar, elde edilen bulguların ve analizlerin yorumlanması ve sonuç ile önerilerin sunulması bölümlerini içermektedir.

Literatür

Siber güvenlik araştırmalarında yaygınlıkla kullanılan CIC-IDS2017 veri seti, Canadian Institute for Cybersecurity (CIC) tarafından 2017 yılında üretilmiştir [5]. Bu veri seti, anomali ve saldırı tespiti alanında birçok çalışmada temel kaynak olarak kullanılmıştır. Makine öğrenmesi ve derin öğrenme algoritmaları ile gerçekleştirilen araştırmalarda, veri setinin zengin saldırı çeşitliliği ve gerçekçi ağ trafiği senaryoları öne çıkmaktadır.

Azizan ve arkadaşları (2021) Destek vektör makinesi (SVM), Rastgele Orman (Random Forest) ve Karar Ormanı (Decision Jungle) algoritmalarını kullanarak gerçekleştirdikleri çalışmada, anormal trafiklerin doğruluğu ve hassasiyeti açısından karşılaştırmalar yapmışlardır. Yapılan karşılaştırmada, en yüksek doğruluk değeri %98.18 oranıyla SVM sınıflandırıcısından elde edilirken, hassasiyet ve duyarlılık sırasıyla %98.74 ve %95.63 olarak bildirilmiştir. Bu sonuçlar, SVM sınıflandırıcısının saldırı tespiti için etkin bir model olduğunu vurgulamaktadır [6].

Widodo ve arkadaşları (2024) tarafından gerçekleştirilen çalışmada, dengesiz veri kümelerinde çok sınıflı saldırı tespitini geliştirmek amacıyla, Sentetik Azınlık Örneklem Arttırma

Yöntemi (Synthetic Minority Oversampling Technique - SMOTE) kullanmışlardır. XGBoost ve Rastgele Orman algoritmaları ile yapılan deneylerde, SMOTE uygulamasının F1 skor, hassasiyet, duyarlılık gibi metrikleri üzerinde olumlu etkileri olduğu gözlemlenmiştir [7].

Sharma ve Kumar (2025), hibrit öğrenme modellerinin kullanımıyla saldırı tespit sistemlerinin performansını arttırmaya yönelik bir çalışma yapmışlardır. Önerdikleri Capsule Network (CapsNet) + Çift Yönlü Uzun Kısa Süreli Bellek (Bidirectional Long Short-Term Memory - BiLSTM) mimarisi, geleneksel yöntemlere kıyasla daha yüksek doğruluk oranları sergilemiştir. Hibrit modelin, özellikle karmaşık saldırı türlerinin tespitinde üstünlük sağladığı belirtilmiştir [8].

Materyal ve Yöntem

Bu çalışmada, siber saldırıların tespiti amacıyla yaygın olarak kullanılan CIC-IDS2017 veri seti kullanılarak üç farklı algoritma (XGBoost, Random Forest ve LightGBM) ile anomali tespiti ve karşılaştırmalı analiz yapılacaktır. Algoritmaların anomali tespitindeki doğruluk, hassasiyet, F1 Skoru ve diğer performans metrikleri detaylı olarak incelenecektir.

Performans Değerlendirme Ölçütleri

Çalışmada, sınıflandırma performansını değerlendirmek için karmaşıklık matrisi(confusion matrix) metrikleri kullanılmaktadır. İki kategoriye (Benign/Normal ve Anomali) ayrılan sınıflandırma problemleri için metrikler aşağıdaki şekilde tanımlanmıştır.

- **TP (True Positive):** Modeller tarafından normal olarak tespit edilen ve gerçekten normal etiketine sahip veri sayısı.

- **TN (True Negative):** Modeller tarafından anomali olarak tespit edilen ve gerçekten anomali etiketine sahip veri sayısı.

- **FP (False Positive):** Modeller tarafından anomali olarak tespit edilen ancak gerçekte normal etiketine sahip veri sayısı.

- **FN (False Negative):** Modeller tarafından normal olarak tespit edilen ancak gerçekte anomali etiketine sahip veri sayısı.

Bu metrikler, model performansının doğru, hassas ve güvenilir olup olmadığını ölçmede kullanılmaktadır.

Veri Seti

Çalışmada kullanılacak olan CIC-IDS2017 veri seti, Canadian Institute for Cybersecurity (CIC) tarafından oluşturulmuş, açık kaynak olarak sunulan ve gerçek dünya ağ trafiği senaryolarını taklit eden verilerden derlenmiştir. Veri seti, ağ trafiğinin yüksek yoğunluklu paketlerden oluştuğu durumları yansıtmaktadır. Toplamda 8 farklı CSV dosyasını içeren veri seti ve açıklamaları Tablo 1'de gösterilmiştir.

Kanada Siber Güvenlik enstitüsü tarafından hazırlanan CIC-IDS2017 veri seti 5 günü içeren ağ trafiği verileridir. Pazartesi gününe ait veriler yalnızca normal saldırı olmayan ağ aktivitelerini içermektedir. Salı, Çarşamba, Perşembe ve Cuma günlerine ait veriler güncel çeşitli saldırı ve normal saldırı olmayan ağ aktivitelerini içermektedir. Veri seti içerisinde

normal ağ trafiği ile birlikte çeşitli güncel saldırı türlerini (DoS, DDoS, Botnet, Brute-force, Web Saldırıları, Port Taraması, Sızma, Heartbleed vb.) içermektedir. Her bir veri, saldırı türü ya da normal olarak etiketlenmiştir. Anomali tespiti amacıyla, her dosyadaki veri etiketleri "normal" ve "anormal" olarak ayrılacaktır. Çalışmanın uygulama ortamı olarak Kaggle kullanılmıştır [9] [10].

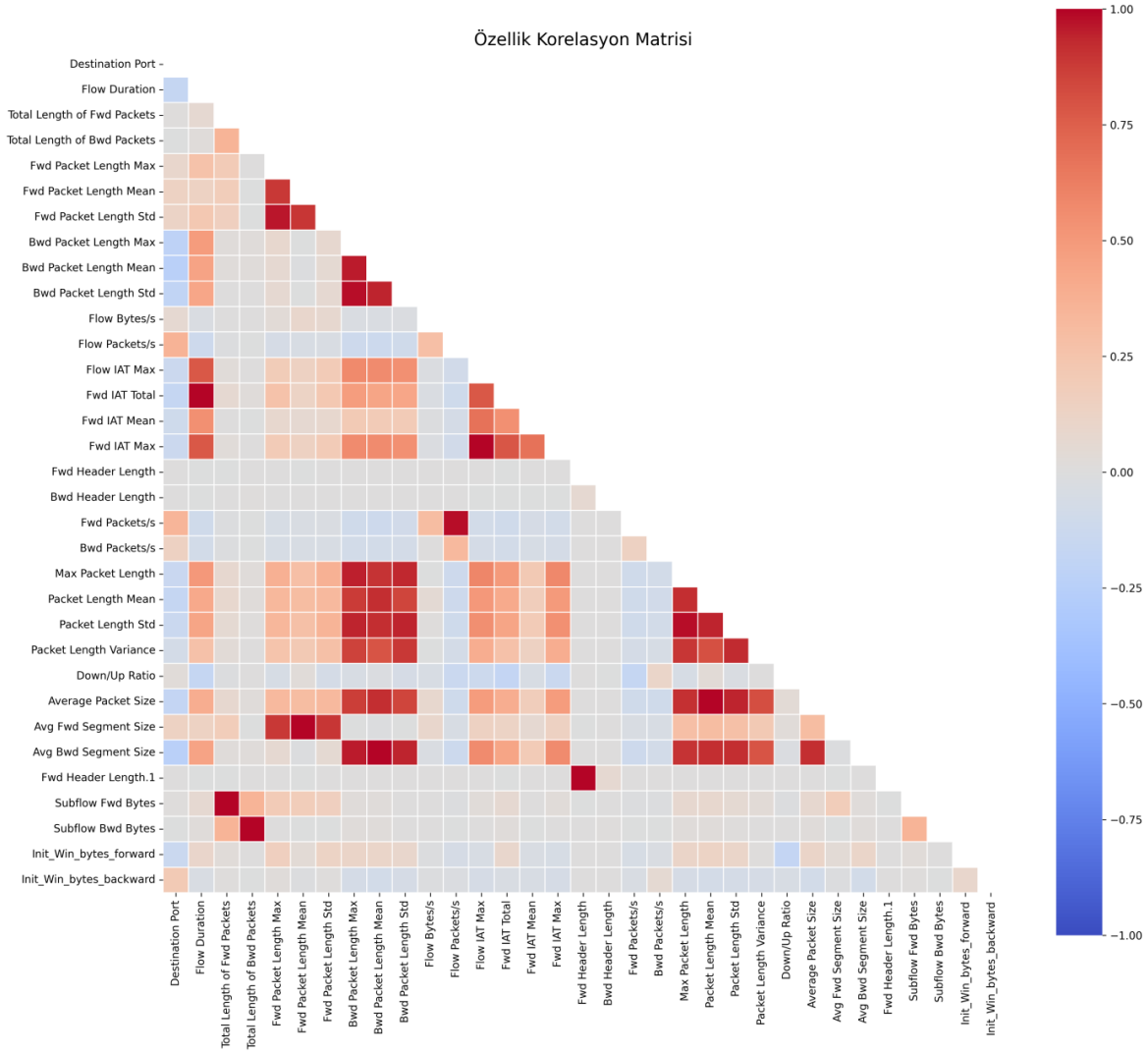
Tablo 1: CIC -IDS2017 veri setine ait bilgiler

Dosya	Gün	Açıklama
Monday-WorkingHours.pcap_ISCX.csv	Pazartesi	Normal ağ aktiviteleri
Tuesday-WorkingHours.pcap_ISCX.csv	Salı	Normal ağ aktiviteleri ve Brute Force, FTP-Patator, SSH-Patator saldırıları
Wednesday-workingHours.pcap_ISCX.csv	Çarşamba	Normal ağ aktiviteleri ve DoS/DDoS, DoS slowloris, DoS Slowhttptest, DoS Hulk, DoS GoldenEye, Heartbleed saldırıları
Thursday-WorkingHours-Morning-WebAttacks.pcap_ISCX.csv	Perşembe	Normal ağ aktiviteleri ve Web Attack (Brute Force, XSS, Sql Injection) saldırıları
Thursday-WorkingHours-Afternoon-Infilteration.pcap_ISCX.csv	Perşembe	Normal ağ aktiviteleri ve Infiltration saldırıları
Friday-WorkingHours-Morning.pcap_ISCX.csv	Cuma	Normal ağ aktiviteleri ve Botnet Ares saldırıları
Friday-WorkingHours-Afternoon-PortScan.pcap_ISCX.csv	Cuma	Normal ağ aktiviteleri ve Port Scan, Firewall Rule on, Firewall Rules off saldırıları
Friday-WorkingHours-Afternoon-DDos.pcap_ISCX.csv	Cuma	Normal ağ aktiviteleri ve DDoS LOIT saldırıları

Veri Ön İşleme

Bu çalışmada, 8 ayrı CSV dosyası birleştirilerek veri temizleme işlemleri gerçekleştirilmiştir. Veri seti içerisindeki eksik (NaN) değerler temizlenmiş ve tekrar eden satırlar kaldırılmıştır. Saldırı çeşitlerini içeren Label sütunu ikili (Binary) formata dönüştürülerek, normal durum "0", anomali durumu "1" olarak etiketlenmiştir.

Kategorik özniteliklere Label Encoding yöntemi uygulanırken, sayısal özellikler standart ölçeklendirme işlemine tabi tutulmuştur. Veri seti içerisinde toplam 79 adet öznitelik bulunmaktadır. Bu özniteliklerin, hedef öznitelikle ilişkisini değerlendirmek ve en iyi öznitelikleri elde etmek amacıyla SelectKBest yöntemi kullanılmıştır. SelectKBest yöntemi, Ki-Kare testine dayalı olarak hedef öznitelik ile ilişkili en anlamlı öznitelikleri seçmektedir [11]. Bu yöntem sonucunda en yüksek bilgi içeren 33 özellik, çalışmada kullanılmak üzere seçilmiştir. Öznitelik korelasyon matrisi Şekil 1'de gösterilmiştir.



Şekil 1: Özellik Korelasyon Matrisi

Modeller ve Bulgular

XGBoost

XGBoost, Gradient Boosting algoritmasının kurallarına dayalı olarak geliştirilmiştir. Kayıp (loss) fonksiyonunu hassas şekilde optimize etmesi, modelin aşırı öğrenmesini engellemesinde önemli rol oynar. Yüksek boyutlu ve büyük veri setlerinde yüksek doğruluk ve hızlı işlem yapması nedeniyle, sınıflandırma problemleri ve ağ saldırı tespit sistemlerinde yaygın olarak tercih edilmektedir [12].

Bu çalışmada, XGBoost yöntemi için en uygun hiperparametreleri belirlemek amacıyla GridSearchCV yöntemi uygulanmıştır. Elde edilen en iyi hiperparametreler şu şekilde belirlenmiştir:

- "max_depth: 7"
- "learning_rate: 0.1"
- "subsample: 0.9"

- "colsample_bytree: 0.8"
- "min_child_weight: 3"
- "gamma: 0.1"
- "reg_alpha: 0.01"
- "reg_lambda: 2"
- "objective:'binary:logistic'"
- "eval_metric: ['auc', 'aucpr']"
- "scale_pos_weight: 4.920966033339519"
- "random_state: 42"

Çapraz geçerleme 10 katlı olarak uygulanmış ve sonuçlar Tablo 2'de sunulmuştur. XGBoost modelinin test aşamasındaki performansı, sırasıyla %99.42 F1 skoru, %99.98 duyarlılık ve %98.87 kesinlik değerleri elde edilmiştir. Modelin doğruluk oranı %99.80 olarak elde edilmiştir. Karmaşıklık matrisi sonuçları ise Tablo 3'te gösterilmiştir.

Tablo 2: XGBoost Fold Sonuç Tablosu

Fold	Doğruluk(%)	F1 Skor	Duyarlılık	Kesinlik
1	99.80	0.9943	0.9998	0.9888
2	99.80	0.9942	0.9998	0.9888
3	99.81	0.9944	0.9998	0.9890
4	99.80	0.9942	0.9999	0.9886
5	99.81	0.9943	0.9998	0.9888
6	99.81	0.9943	0.9997	0.9890
7	99.80	0.9942	0.9999	0.9885
8	99.81	0.9944	0.9998	0.9891
9	99.80	0.9940	0.9998	0.9882
10	99.80	0.9942	0.9998	0.9887

Tablo 3: XGBoost Karmaşıklık Matrisi Tablosu

Fold	TN	FP	FN	TP
1	209017	489	21	42553
2	209031	475	30	42544
3	209033	473	17	42557
4	209030	476	16	42558
5	208983	523	17	42557
6	209017	489	30	42544
7	208982	524	11	42563
8	209016	489	27	42548
9	208971	534	7	42567
10	209041	464	25	42549

Random Forest

Rastgele Orman, sınıflandırma problemlerinin çözümü için yaygın olarak kullanılan bir makine öğrenmesi algoritmasıdır. Çok sayıda karar ağacının bir araya gelmesi ile oluşturulan bu yöntem, düğümlerdeki en iyi dallanmayı sağlamak amacıyla Classification and Regression Tree (CART) yöntemi ve bölünme işlemlerinde gini indeksi kullanılarak geliştirilmiştir [13]. Ayrıca, Rastgele orman algoritması aşırı öğrenmeye karşı dirençli bir makine öğrenmesi algoritmasıdır.

Bu çalışmada, Random Forest yöntemine en uygun hiperparametreleri belirlemek için GridSearchCV yöntemi uygulanmıştır. Elde edilen en iyi hiperparametreler şu şekilde tespit edilmiştir:

- "max_depth: None"
- "min_samples_split: 10"
- "n_estimators: 200"

Çapraz geçerleme 10 katlı olarak uygulanmış ve sonuçlar Tablo 4'de sunulmuştur. Random Forest modelinin test aşamasında performansı, sırasıyla %99.64 F1 skor, %99.62 duyarlılık ve %99.66 kesinlik değerleri elde edilmiştir. Modelin doğruluk oranı %99.879 olarak elde edilmiştir. Karmaşıklık matrisi sonuçları ise Tablo 5'te gösterilmiştir.

Tablo 4: Random Forest Fold Sonuç Tablosu

Fold	Doğruluk(%)	F1 Skor	Duyarlılık	Kesinlik
1	99.87	0.9961	0.9959	0.9963
2	99.87	0.9962	0.9957	0.9967
3	99.89	0.9967	0.9962	0.9971
4	99.86	0.9959	0.9959	0.9960
5	99.88	0.9965	0.9964	0.9965
6	99.89	0.9968	0.9965	0.9970
7	99.88	0.9965	0.9963	0.9967
8	99.89	0.9966	0.9970	0.9963
9	99.87	0.9963	0.9961	0.9964
10	99.89	0.9967	0.9964	0.9969

Tablo 5: Random Forest Karmaşıklık Matrisi Tablosu

Fold	TN	FP	FN	TP
1	209327	157	173	42423
2	209018	141	184	42737
3	209314	124	160	42482
4	209279	170	176	42455
5	209613	146	153	42168
6	209487	126	147	42320
7	209242	140	157	42541
8	209507	158	129	42286
9	209387	152	165	42375
10	209437	132	152	42358

LightGBM

LightGBM, gradient boosting prensiplerine dayalı olarak geliştirilmiş ve yoğun karar ağaçları kullanılarak güçlü öğrenme modeli oluşturan bir algoritmadır. Her bir yeni ağaç, kendisinden önceki ağaçlarda gözlemlenen hataları düzeltmeye odaklanarak modelin genel performansını arttırır. Özellikle, anomali tespiti gibi karmaşık sınıflandırma problemlerinde LightGBM hızlı ve verimli çalışması nedeniyle tercih edilen bir algoritmadır.

Bu çalışmada, LightGBM yöntemine en uygun hiperparametreleri belirlemek amacıyla GridSearchCV yöntemi uygulanmıştır. Elde edilen en iyi hiperparametreler aşağıdaki gibidir.

- "subsample: 0.8"
- "num_leaves: 70"
- "n_estimators: 500"
- "min_child_samples: 50"
- "learning_rate: 0.05"
- "colsample_bytree: 0.6"

Çapraz geçerleme 10 katlı olarak uygulanmış ve elde edilen sonuçlar Tablo 6'da sunulmuştur. Çalışma kapsamında LightGBM modeli, test aşamasında sırasıyla %99.63 F1 skor, %99.65 duyarlılık ve %99.61 kesinlik değerleri elde edilmiştir. Modelin doğruluk oranı %99.874 olarak elde edilmiştir. Karmaşıklık matrisi sonuçları detaylı olarak Tablo 7'de gösterilmiştir.

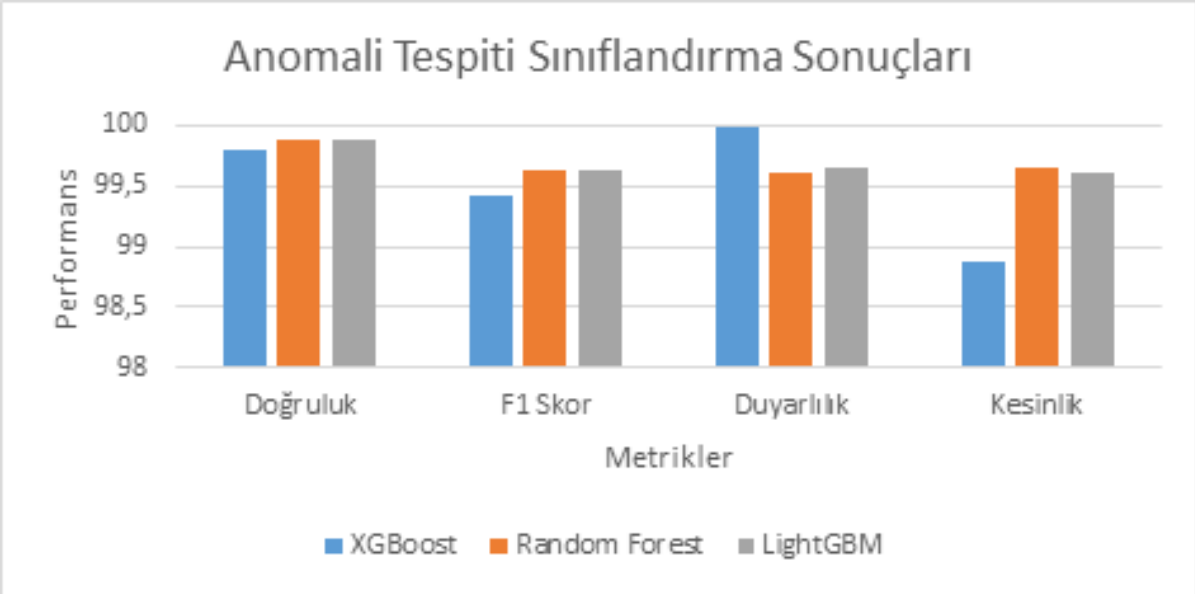
Tablo 6: LightGBM Fold Sonuç Tablosu

Fold	Doğruluk(%)	F1 Skor	Duyarlılık	Kesinlik
1	99.87	0.9963	0.9961	0.9965
2	99.87	0.9962	0.9959	0.9965
3	99.87	0.9961	0.9962	0.9960
4	99.87	0.9962	0.9963	0.9961
5	99.87	0.9963	0.9965	0.9962
6	99.87	0.9961	0.9964	0.9958
7	99.89	0.9968	0.9968	0.9967
8	99.87	0.9963	0.9972	0.9952
9	99.88	0.9965	0.9976	0.9955
10	99.88	0.9966	0.9967	0.9965

Tablo 7: LightGBM Karmaşıklık Matrisi Tablosu

Fold	TN	FP	FN	TP
1	209358	148	163	42411
2	209359	147	172	42402
3	209339	167	159	42415
4	209344	162	157	42417
5	209345	161	146	42428
6	209331	175	151	42423
7	209367	139	133	42441
8	209314	191	117	42458
9	209314	191	101	42473
10	209357	148	138	42436

Sonuçlar doğrultusunda, XGBoost, Random Forest ve LightGBM algoritmalarının özellikle ağ saldırı tespiti sistemleri için performansları incelenmiş; her üç model de test verileri üzerinde yüksek doğruluk oranlarına ulaşmıştır. En yüksek doğruluk, sırasıyla %99.879 ile Random Forest, %99.874 ile LightGBM ve %99.80 ile XGBoost modeliyle elde edilmiştir. Benzer şekilde, yüksek F1 skoru, duyarlılık ve kesinlik değerleri her üç modelin başarısını desteklemektedir. Şekil 2, anomali tespiti sınıflandırma sonuçları görsel olarak sunulmaktadır.



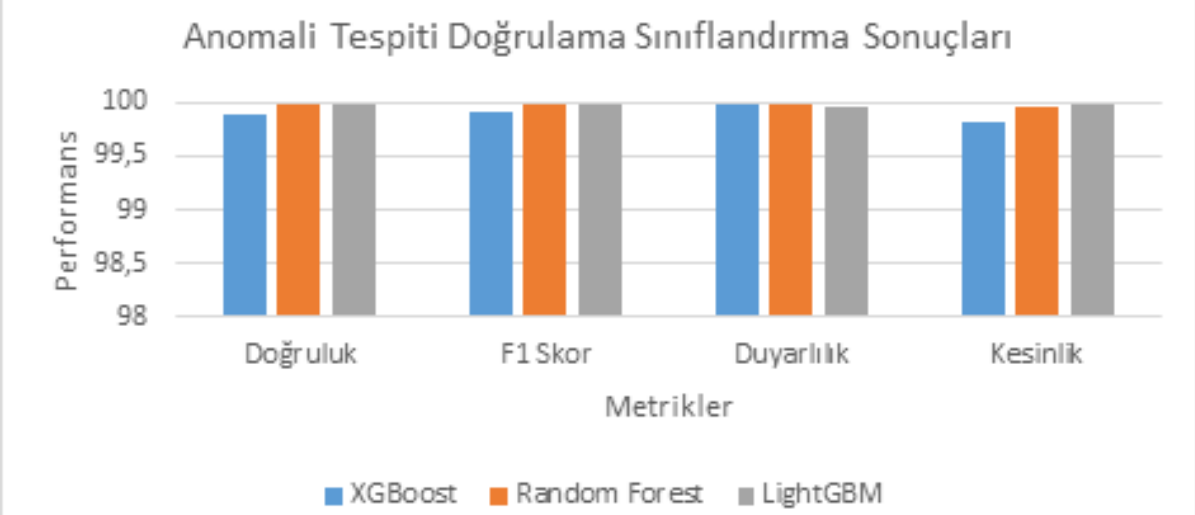
Şekil 2: Anomali Tespiti Sınıflandırma Sonuçları

Modellerin Farklı Veri Seti Üzerinde Doğrulanması

Çalışmada incelenen XGBoost, Random Forest ve LightGBM algoritmalarından en iyi performansı elde edilen modeller, anomali tespiti amacıyla pkl formatında kayıt altına alınmıştır. Bu modellerin gerçek performansları, Kanada Siber Güvenlik Enstitüsü tarafından geliştirilen DDoS Evaluation Dataset (CIC-DDoS2019) [14] kullanılarak test edilmiştir. Veri seti daha çok DDoS saldırılarını simüle eden labaratuvar ortamı verileri içermektedir. Bu kapsamda, CIC-DDoS2019 veri seti içerisindeki Friday-WorkingHours-Afternoon-DDos. pcap_ISCX.csv dosyası tercih edilmiştir.

Veri seti incelendiğinde, modellerin eğitiminde kullanılan sütunların büyük ölçüde korunmuş olduğu gözlemlenmiştir. Bu doğrultuda veri temizleme işlemleri uygulanarak, önceden seçilmiş öznitelikler kullanılarak anomali tespiti gerçekleştirilmiştir. CIC-DDoS2019 veri setinde, Friday-WorkingHours-Afternook CSV dosyasında 97.686 adet normal ve 128.025 adet anomali (DDoS) verisi mevcuttur.

Doğrulama çalışmasından elde edilen sonuçlar Şekil 3'te sunulmuştur.



Şekil 3: Anomali Tespiti Doğrulama Sınıflandırma Sonuçları

Sonuçlar incelendiğinde, modellerin performanslarının saldırı tespit sistemleri için yeterli düzeyde olduğu anlaşılmaktadır. Modellerin genel sonuçları, CIC-IDS2017 veri setiyle eğitim yapıldıktan sonra, daha güncel CIC-DDoS2019 veri seti üzerinde doğrulanarak benzer başarılar elde edilmiştir.

Model performansları arasındaki farkların, kullanılan algoritmaların yapısal farklılıkları ve veri setlerinde bulunan özniteliklerin dağılımından kaynaklandığı düşünülmektedir. XGBoost, Random Forest ve LightGBM algoritmaları, farklı karar yapıları ve öğrenme yöntemlerine sahip olduğu için, her iki veri setinde de anomali olarak etiketlenen ve normal olarak bulunan verilerin dağılımı model performanslarını ve sonuçlarını etkileyebilmektedir.

Sonuçlar

Bu çalışmada, siber güvenlik alanındaki çalışmalarda sıklıkla kullanılan iki farklı veri seti (CIC-IDS2017 ve CIC-DDoS2019) kullanılarak anomali tespiti ve doğrulama işlemi gerçekleştirilmiştir. Siber güvenlik, teknolojinin hızlı gelişimine paralel olarak önemini korumakta olup bilgi ve sistem güvenliğinin sağlanması kritik bir ihtiyaç olarak karşımıza çıkmaktadır. Dünya üzerinde anlık ve çeşitli şekillerde gerçekleştirilen siber saldırılar göz önünde bulundurulduğunda, sistemlerde oluşabilecek herhangi bir saldırı durumunu erken aşamada tespit edebilmek büyük önem arz etmektedir.

Çalışmada CIC-IDS2017 veri seti ile XGBoost, Random Forest ve LightGBM algoritmaları kullanılarak modeller eğitilmiş, bu modellere ait performans metrikleri (doğruluk, F1 skoru, duyarlılık, kesinlik) detaylı olarak sunulmuştur. Doğrulama amacıyla, daha güncel olan CIC-DDoS2019 veri seti üzerinde de test edilen modeller, gerçek dünya koşullarına büyük ölçüde simüle eden verilerle benzer sonuçlar vermiştir.

Her ne kadar modeller, test verileri üzerinde yüksek doğruluk oranlarına ulaşsa da gerçek zamanlı ağ gecikmeleri, donanım sınırlamaları veya diğer çevresel faktörler gibi parametrelerin henüz test edilmemiş olması, gelecekteki çalışmalarda ele alınması gereken önemli noktalardır. CIC-DDoS2019 veri setinin çeşitli DDoS saldırı türlerini içermesi, kapsamlı CIC-IDS2017 veri seti ile eğitilen modellerin, farklı saldırı türlerini başarılı bir şekilde tespit edebileceğini göstermektedir.

Elde edilen sonuçlar, saldırı tespit sistemleri açısından kullanılabilir düzeyde olduğunu ortaya koymaktadır. Gelecekteki çalışmalar, farklı algoritmaların ve gelişmiş özellik mühendisliği tekniklerinin entegrasyonu ile model performansının daha da arttırılmasını hedefleyebilir. Ayrıca, modellerin farklı saldırı türleri üzerindeki başarısının ayrı ayrı değerlendirilmesi ve zayıf yönlerinin belirlenerek iyileştirme yollarının araştırılması, önemli bir araştırma konusu olarak öne çıkmaktadır.

Çıkar Beyanı: Yazarlar çıkar çatışması olmadığını belirtmektedir.

Yazar Katkıları: Bu makalenin yazılmasına tüm yazarlar eşit katkıda bulunmuştur.

REFERANSLAR

[1] H. Ahmetoğlu ve R. DAŞ, «Makine Öğrenmesi Yöntemleri Kullanarak Web Uygulama Saldırılarının Tespitinde Genetik Öznitelik Seçimi Yaklaşımı,» Türkiye Bilişim Vakfı Bilgisayar Bilimleri Ve Mühendisliği Dergisi, cilt 14, no. 2, pp. 109-119, 2021.

[2] European Union Agency for Cybersecurity, «European Union Agency for Cybersecurity,» [Çevrimiçi]. Available: https://www.enisa.europa.eu/publications/enisa-threat-landscape-2024. [Erişildi: 25 03 2025].

[3] S. Çelik ve B. Çeliktaş , «Güncel Siber Güvenlik Tehditleri: Fidye Yazılımlar,» Cyberpolitik Journal, cilt 3, no. 5, pp. 105-132, 2018.

[4] K. S. J. A. &. D. P. Adewole, «Intrusion Detection Framework for Internet of Things with Rule Induction for Model Explanation,» Sensors, cilt 25, no. 6, p. 1845, 2025.

[5] Canadian Institute for Cybersecurity, «Canadian Institute for Cybersecurity,» [Çevrimiçi]. Available: https://www.unb.ca/cic/datasets/ids-2017.html. [Erişildi: 26 03 2025].

[6] A. M. S. M. A. F. C. W. M. M. M. &. K. B. Azizan, «A Machine Learning Approach for Improving the Performance of Network Intrusion Detection Systems,» Annals of Emerging Technologies in Computing, 2021.

[7] A. Widodo, B. Setiawan ve R. Indraswari, «Machine learning-based intrusion detection on multi-class imbalanced dataset using,» Procedia Computer Science, cilt 234, pp. 578-583, 2024.

[8] V. Sharma ve M. Kumar, «Improving intrusion detection with hybrid deep learning models: A study on CIC-IDS2017, UNSW-NB15, and KDD CUP 99,» Journal of Information Systems Engineering and Management, cilt 10, 2025.

[9] Canadian Institute for Cybersecurity, «Intrusion detection evaluation dataset (CIC-IDS2017),» Canadian Institute for Cybersecurity, [Çevrimiçi]. Available: https://www.unb.ca/cic/datasets/ids-2017.html. [Erişildi: 27 03 2025].

[10] B. Ekici ve H. Takcı, «Bilgisayar Ağlarında Anomali Tespiti Yaklaşımı ile Saldırı Tespiti,» Afyon Kocatepe Üniversitesi Fen Ve Mühendislik Bilimleri Dergisi, cilt 5, no. 1016-1027, p. 22, 2022.

[11] M. Akyol, «Clustering Hotels and Analyzing the Importance of Their Features by Machine Learning Techniques,» Bilgisayar Bilimleri Ve Teknolojileri Dergisi, cilt 2, no. 1, pp. 16-23, 2021.

[12] B. Guembe, S. Misra ve A. Azeta, «Federated Bayesian optimization XGBoost model for cyberattack detection in internet of medical things,» Journal of Parallel and Distributed Computing, cilt 193, 2024.

[13] M. Ercire ve . A. Ünsal, «Kısa Süreli Güç Kalitesi Bozulmalarının Dalgacık Analizi Ve Rastgele Orman Yöntemi İle Sınıflandırılması,» Uludağ Üniversitesi Mühendislik Fakültesi Dergisi, cilt 26, no. 3, pp. 903-920, 2021.

[14] Canadian Institute for Cybersecurity, «DDoS 2019 | Datasets | Research | Canadian Institute for Cybersecurity | UNB,» Canadian Institute for Cybersecurity, [Çevrimiçi]. Available: https://www.unb.ca/cic/datasets/ddos-2019.html.

[15] S. Ding, Z. Zhibin ve X. Zhang, «An overview on semi-supervised support vector machine,» Neural Computing and Applications, cilt 28, pp. 969-978, 2017.

SINGLE PACKAGE IIOT ASSISTED REAL TIME MONITORING AND MANAGEMENT CONTROL SYSTEM FOR SMART FACTORY 4.0 INDUSTRIAL APPLICATIONS

Yashfa Kamran*
National University of Science and Technology, Department of Electrical Engineering, Islamabad, Pakistan

Habiba Tasneem
IIUI, Department of ECE, Islamabad, Pakistan

Noor Fatimah
University of Engineering and Technology, Department of Electrical and Electronics, Taxila, Pakistan
(Corresponding author) yashfakamran.92@gmail.com

ABSTRACT

In the era of Industry 4.0, propelled by advancements such as internet of things (IoT) connectivity and enhanced intelligence, traditional equipment including direct current (DC) motors undergo significant upgrades. So, the remote maintenance and condition monitoring of industrial equipment are paramount for any modern manufacturing enterprise. However, the current solutions for the remote monitoring of industrial equipment are based on disintegrated units leading to higher cost, larger space occupation and lower performance. This paper achieved a prototype for a Smart Factory 4.0 platform, integrating industrial networks, industrial internet of things (IIoT) capabilities, sensors, actuators, and control terminals onto a single printed circuit board (PCB). The system monitors process efficiency, working conditions, and product quality comprehensively, facilitating real-time recording, control, and monitoring of critical parameters. It extends functionalities to include monitoring and controlling gas levels within industrial units, promoting cost-effective manufacturing practices, and ensuring factory safety. Remote access allows monitoring of various parameters displayed on smart devices for maintenance and analysis. The smart control and management system in this work empowers authorized personnel to remotely control the status of industrial units, enhancing operational efficiency, and ensuring optimal performance.

Keywords: IIoT, sensors, cyber physical systems, predictive maintenance.

Introduction

The advent of Industry 4.0 has revolutionized manufacturing and automation through cyber- physical systems, advanced analytics, artificial intelligence (AI), and internet of things (IoT) integration. This paper explores its applications in motor speed control, smart gas management, and environment control, emphasizing their role in enhancing efficiency, safety, and sustainability.

Motor speed control optimizes manufacturing by integrating sensor data, machine learning, and predictive maintenance. Similarly, smart gas management systems, enabled by IoT sensors and cloud analytics, enhance safety and process efficiency by real-time monitoring and predictive maintenance. Moreover, environment control, leveraging IoT and data analytics, ensures dynamic adjustments for energy efficiency, waste reduction, and worker safety. Implementation of Industry 4.0 systems, including speed control of motors, smart gas management, and environment control, is crucial for competitiveness and efficiency. The proposed system offers an intelligent solution, addressing energy inefficiencies and safety hazards while providing greater control and responsiveness. The paper outlines a framework for industrial internet of things (IioT)-assisted systems for smart factories, aiming to enhance communication, monitoring, and control remotely. The paper details the project methodology, involving designing motor driver circuits, a smart gas management system, and intelligent sensor network for condition monitoring, all integrated onto a single PCB with cloud connectivity, aiming to validate its feasibility and performance, contributing to smarter, more efficient manufacturing processes.

The applications of Industry 4.0 technologies in the specific areas, e.g, the integration of advanced sensing technologies, speed control of motors, smart gas management systems, environment control, data analytics, artificial intelligence, and IoT are explored to achieve enhanced efficiency, safety, and productivity in industrial settings.

Existing Industry 4.0 Technologies

J. Smith et. al. explored sensor integration, predictive maintenance, and adaptive control algorithms, highlighting their potential to enhance motor efficiency and overall system performance [1]. It explores the integration of advanced sensors, communication networks, and data analytics techniques for real-time monitoring and control of motors. The authors discuss the benefits of motor speed control in optimizing energy efficiency and improving productivity within industrial processes.

Q. Chen et. al. presented an overview of smart gas management systems in Industry 4.0 [2]. The authors discuss the integration of IoT sensors, data analytics, and cloud computing to optimize gas consumption, detect leaks, and improve safety. The study highlights the potential of these systems to reduce operational costs and enhance process efficiency.

Y. Wang et. al. reviewed the state-of-the-art environment control techniques in Industry 4.0 [3]. They explore the integration of advanced sensing technologies, data analytics, and machine learning algorithms to monitor and control environmental parameters. The study emphasizes the potential for improved energy efficiency, waste reduction, and worker safety.

H. Lee et. al. focused on real-time speed control strategies for motors using artificial intelligence techniques in Industry 4.0 [4]. The authors propose a deep learning- based algorithm to optimize motor speed control, achieving improved accuracy and response time compared to conventional methods.

R. Gupta et. al. presented an IoT-based smart gas management system designed for industrial applications [5]. The study discusses the integration of gas sensors, wireless communication, and cloud-based analytics to enable real- time monitoring, leak detection, and predictive maintenance, resulting in enhanced safety and cost savings.

L. Zhang et. al. worked on the recent advancements in environment control systems within the context of Industry

4.0 [6]. The authors explore the integration of IoT devices, big data analytics, and control algorithms to optimize air quality, temperature, and humidity levels, emphasizing the potential for improved worker comfort and productivity.

G. Martinez et. al. investigated the application of machine learning techniques for motor speed control in Industry 4.0 environments [7]. They proposed a hybrid control strategy combining support vector regression and proportionalintegral-derivative (PID) control, achieving accurate and adaptive speed control for various motor types.

X. Li et. al. present a wireless sensor network-based gas monitoring system designed for Industry 4.0 applications [8]. The authors discuss the deployment of sensor nodes, data aggregation, and cloud-based analytics to enable real- time gas monitoring, anomaly detection, and improved safety in industrial settings.

S. Yang et. al. provide a systematic review of intelligent environmental control techniques in Industry 4.0 [9]. They discuss the integration of cyber-physical systems, machine learning algorithms, and advanced control strategies for achieving energy efficiency, sustainability, and worker wellbeing in manufacturing environments.

M. Garcia et. al. focus on the implementation of predictive maintenance techniques for motor speed control in the Industry 4.0 context [10]. The authors propose a datadriven predictive maintenance framework, enabling timely maintenance actions to prevent motor failures and optimize operational efficiency.

R. Anderson et. al. present a comprehensive review of smart gas management systems for Industry 4.0 environments [11]. The paper discusses the use of IoT- enabled sensors, data analytics, and cloud-based platforms to optimize gas consumption, detect leaks, and enhance safety in industrial settings. The authors highlight the potential cost savings and environmental benefits achieved through smart gas management.

W. Chen et. al. explore the integration of IoT technologies and artificial intelligence for real-time monitoring and optimization of environmental factors such as temperature, humidity, and air quality within industrial facilities [12]. The authors discuss the potential for improved worker comfort, energy efficiency, and product quality.

E. Martinez et. al. do a review of IoT techniques for industrial automation within the Industry 4.0 paradigm. The paper discusses the deployment of IoT devices, wireless communication protocols, and cloud-based platforms for seamless connectivity and data exchange in smart factories. The authors highlight the importance of reliable and secure data transmission for efficient decision- making and process optimization.[13]

J. Kim et. al. provide an in-depth review of artificial intelligence applications in Industry 4.0. The paper explores the use of machine learning, deep learning, and cognitive computing techniques for intelligent decision-making, predictive maintenance, and quality control in industrial processes. The authors discuss the potential for AI to enhance productivity, reduce costs, and enable new business models [14].

J. Kim et. al. provide a comprehensive review of strategies to enhance environmental sustainability in Industry 4.0 settings. The paper discusses the integration of IoT technologies, data analytics, and AI for real-time monitoring and optimization of energy consumption, waste management, and carbon footprint reduction. The authors emphasize the importance of adopting eco-friendly practices for long-term sustainability and competitiveness [15].

Q. Chen et. al. proposed an AI-driven environment control system for energy optimization in smart factories. The paper presents a framework that integrates machine learning algorithms with real-time sensor data to dynamically adjust lighting, heating, and ventilation, resulting in improved energy efficiency and worker comfort. They also highlight the integration of AI technologies with IoT systems, enabling autonomous decision-making and intelligent control in smart factories. This review serves as a valuable resource for researchers and practitioners interested in the intersection of artificial intelligence and Industry 4.0 [16].

Y. Wang et. al. investigate the application of IoT-enabled speed control techniques for motors in industrial automation. The paper explores the integration of motor control systems with IoT platforms, enabling remote monitoring, predictive maintenance, and adaptive control strategies to improve production efficiency and reduce downtime [17].

J. Smith et. al. conduct a survey on the integration of Industry 4.0 and motor control systems. The paper discusses the adoption of advanced motor control algorithms, real-time monitoring techniques, and predictive maintenance strategies enabled by IoT and AI technologies. The authors highlight the impact of such integration on improving operational efficiency, reducing downtime, and extending motor lifespan [18].

R. Anderson et. al. provide a comprehensive review of motor speed control techniques for Industry 4.0 applications. The authors discuss various control strategies, including proportional- integral-derivative (PID) control, field-oriented control (FOC), and model predictive control (MPC), highlighting their advantages and limitations [19]. They also explore the integration of motor speed control systems with IoT technologies, emphasizing the potential for enhanced monitoring, diagnostics, and maintenance. This review serves as a valuable resource for researchers and practitioners interested in motor speed control within the context of Industry 4.0.

In [20] propose a novel approach to real-time speed control of motors using deep reinforcement learning techniques. The paper presents a framework that combines neural networks and reinforcement learning algorithms to optimize motor performance, reduce energy consumption, and enhance overall system efficiency.[20]

In [21] discussed the implementation of a smart gas leakage detection and management system based on wireless sensor networks. The paper highlights the benefits of using IoT-enabled gas sensors and cloud-based analytics to detect leaks, provide real-time alerts, and enable proactive maintenance in industrial environments.

Research gap

Various intelligent speed control techniques are implemented in literature for motors, smart gas management systems, and environment control strategies. However, a notable research gap lies in the comprehensive integration of these three technologies into unified, holistic systems that optimize overall industrial processes. While individual studies offer valuable insights, there was a need for research focusing on the seamless integration of all three technologies into one robust system to achieve synergistic effects, ultimately maximizing efficiency, and productivity in Industry 4.0 environments. This work addresses the abovementioned gap by leveraging the collective benefits of three Industry 4.0 technologies: DC motor speed control and monitoring, Industry health condition monitoring and control, and Smart Gas Management to design and develop a unified system in a cohesive manner onto a single PCB.

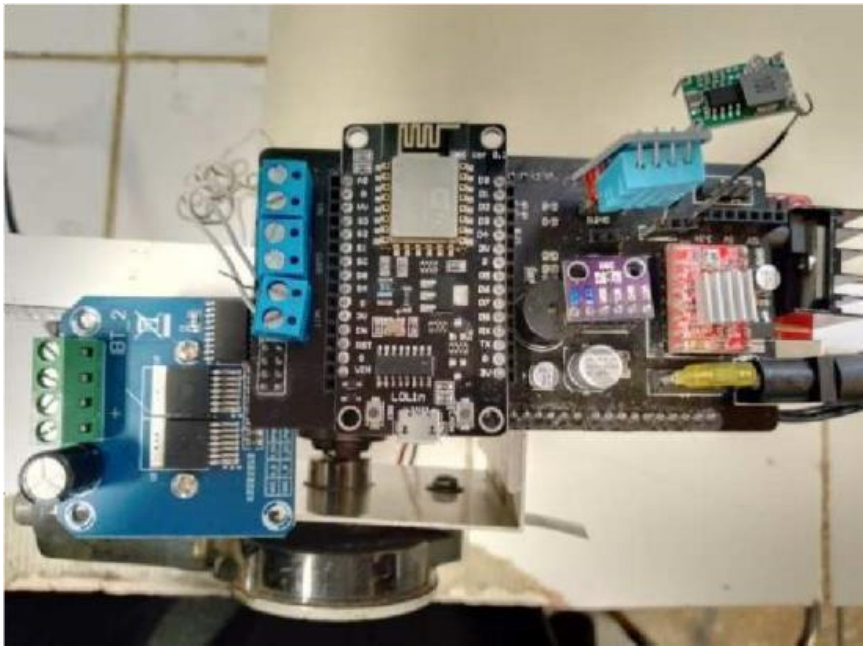


Figure 1: The Single Package System

System Design And İmplementation

The system design and implementation of the project involved meticulous planning and careful consideration of various aspects. Initially, the requirements and objectives of the prototype, considering factors such as functionality, reliability, component compatibility,

cloud connectivity and safety were defined. Later, the design procedure encompassed several stages, including schematic design (Figure X), component selection, PCB layout (Figure X) and IoT integration. Finally, thorough testing and simulation (Figure X) were conducted to verify the circuit's functionality performance and cloud connectivity. For the implementation phase, best practices were followed, aiming to achieve least component count and a zero wiring system (Figure X). Stress testing and debugging were performed to ensure that the integrated system operated as intended under a range of load conditions (Figure X). Necessary safeguards, such as fuses and protective devices were also incorporated to mitigate potential risks and ensure the system's reliability.

Throughout the design and implementation process, the focus was on efficiency and optimization, striving to minimize power losses, reduce component count, and enhance overall performance all the while focusing on seamless transfer of all data over the cloud. Additionally, the design procedures and system specifications were documented, enabling future replication and maintenance. Through careful planning, rigorous testing, and adherence to safety standards, achieved a functional and reliable proof of concept that meets the project's objectives.

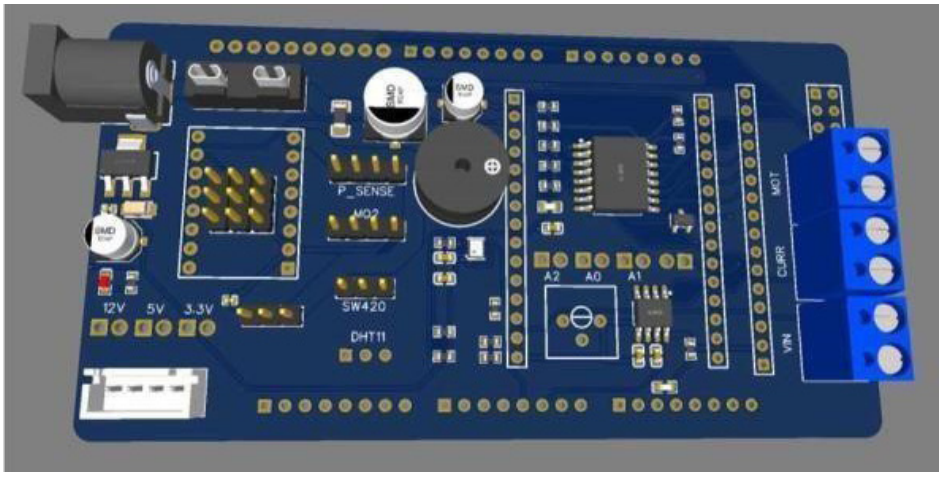


Figure 2: 3D Model of PCB for the IIoT Assisted System – Front

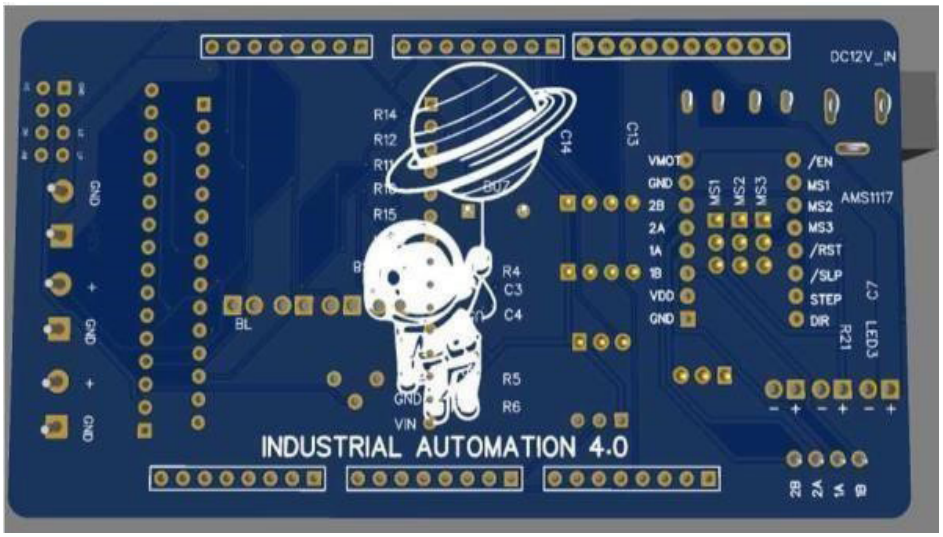


Figure 3: 3D Model of PCB for the IIoT Assisted System - Back

Hardware System Design

Dc Motor Control And Monitoring:

In the hardware system design for 12V, 2A, DC motor control and monitoring, the Arduino Mega 2560 works as the central microcontroller, providing extensive I/O capabilities. To monitor environmental conditions crucial for motor operation, with that is the DHT11 and BMP280 sensors integrated for temperature, humidity, and pressure measurements. Additionally, the SW420 vibration sensor is incorporated for motion-based triggering of alarm systems based on a buzzer and app notification and the voltage sensor B25 for efficient power management. For accurate current measurements in varying load conditions, the ACS712 current sensor was employed, along with the BTS7690 motor driver for controlling high-current loads on the conveyor belt. These components collectively facilitated real-time monitoring and control of the DC motor, ensuring optimal performance and safety.

Figure 4: Complete Hardware Prototype for Motor, Industry Health Condition and Gas Control and Monitoring

Industry Health Control And Monitoring:

Within the hardware system design for industry health condition control and monitoring, focus was on integrating sensors for environmental parameters critical to manufacturing processes. The DHT11 and BMP280 sensors provided essential data on temperature, humidity, and pressure, enabling comprehensive environmental monitoring. The SW420 vibration sensor was utilized to detect motion or abnormalities in machinery, triggering alarms through a buzzer and mobile app notifications. With the ACS712 current sensor and BTS7690 motor driver, ensuring precise monitoring and control of industrial equipment, promoting operational efficiency and minimizing downtime.

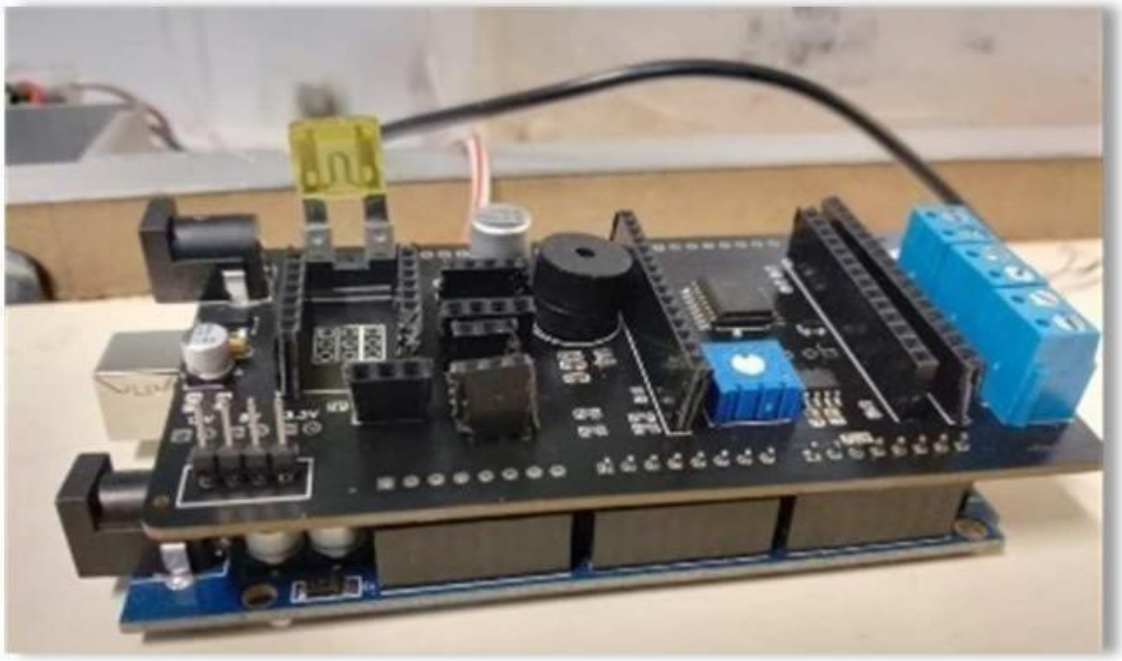


Figure 5: PCB with the safety fuses slapped on top of Mega 2560

Smart Gas Management System:

Incorporating the smart gas management system into our hardware design involved integrating sensors for gas detection and differential pressure monitoring. The MQ2 smoke sensor facilitated the detection of various gases, while the MPS20N0040D pressure sensor monitored differential pressure in gas systems. To control the intelligent gas valve, the NEMA stepper motor was employed and its A4988 stepper motor driver, allowing for precise adjustment and regulation of gas flow. Through seamless connectivity and integration facilitated by the designed PCB, these subsystems operated harmoniously within our Smart Factory 4.0 platform, enhancing safety, efficiency, and sustainability in manufacturing processes.

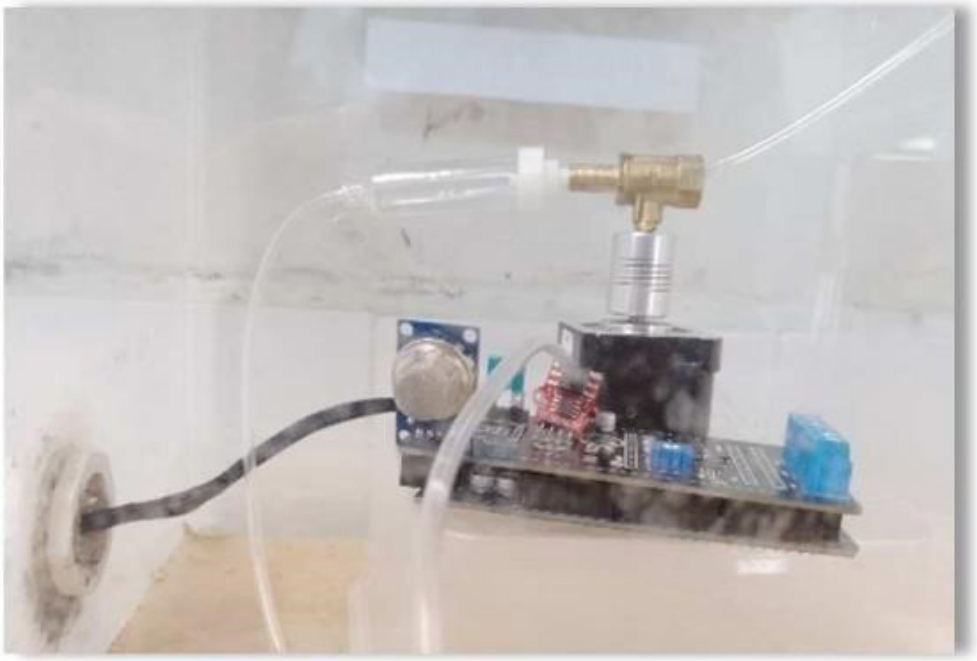


Figure 6: Detailed View of the Gas Chamber with the Intelligent Gas valve

Software Integration

The mobile application in our project serves as a vital interface for users to interact seamlessly with the hardware, offering a user-friendly platform to access and control various features and functionalities. Authorized access gives real-time monitoring capabilities, gaining insights into parameters like temperature, humidity, air quality, or energy consumption. Through remote control functionality, users can adjust settings, manage devices, and activate security features, enhancing convenience and flexibility in environment management.

In the app design phase, leveraging Figma streamlined the process, enabling us to create a minimalist interface. With its extensive library of pre-built components and collaborative capabilities, Figma facilitated swift iteration and assembly of designs. In the development phase, Android Studio was used which provided essential tools and libraries for writing code in Java, ensuring the creation of a stable, efficient, and user-friendly app. Through debugging tools, emulators, and access to a rich set of APIs, app functionality and cloud platform integration could be efficiently managed, ultimately delivering a robust mobile application that enhances user interaction and control over the prototype industrial unit.

Figure. 7: Home Screen of Mobile Application for Management of the Three Subsystems

PCB Design

A meticulously designed PCB (Printed Circuit Board) layout, which provided the project with numerous advantages, including compact component integration, enhanced signal integrity, simplified assembly, and improved reliability. By allowing for the miniaturization of electronics, our proposed design of PCB optimized space utilization and enhanced portability. Precise routing and controlled impedance on PCBs was ensured to minimize noise and cross-talk, while standardized layouts simplified component mounting, reducing errors and assembly time. Additionally, PCB provided stability, protection, and thermal management, ensuring durability and optimal performance. Utilizing Altium Designer, a popular electronic design automation (EDA) software, had comprehensive features for PCB design, schematic capture, component libraries, layout, routing, simulation, and manufacturing documentation. Its capabilities streamlined our design process, ensuring efficiency, accuracy, and a seamless workflow from concept to production. EasyEDA was also used, a web-based PCB design tool, which provided an integrated editor, schematic capture, circuit simulation, vast component library, design rule checking, and manufacturing documentation generation, simplifying component selection, design integrity, and fabrication. Finally, PCB fabrication was implemented using Surface Mount Technology (SMT), enabling high-density circuitry, improved performance, and efficient assembly, making it a widely adopted manufacturing process in modern electronics production.

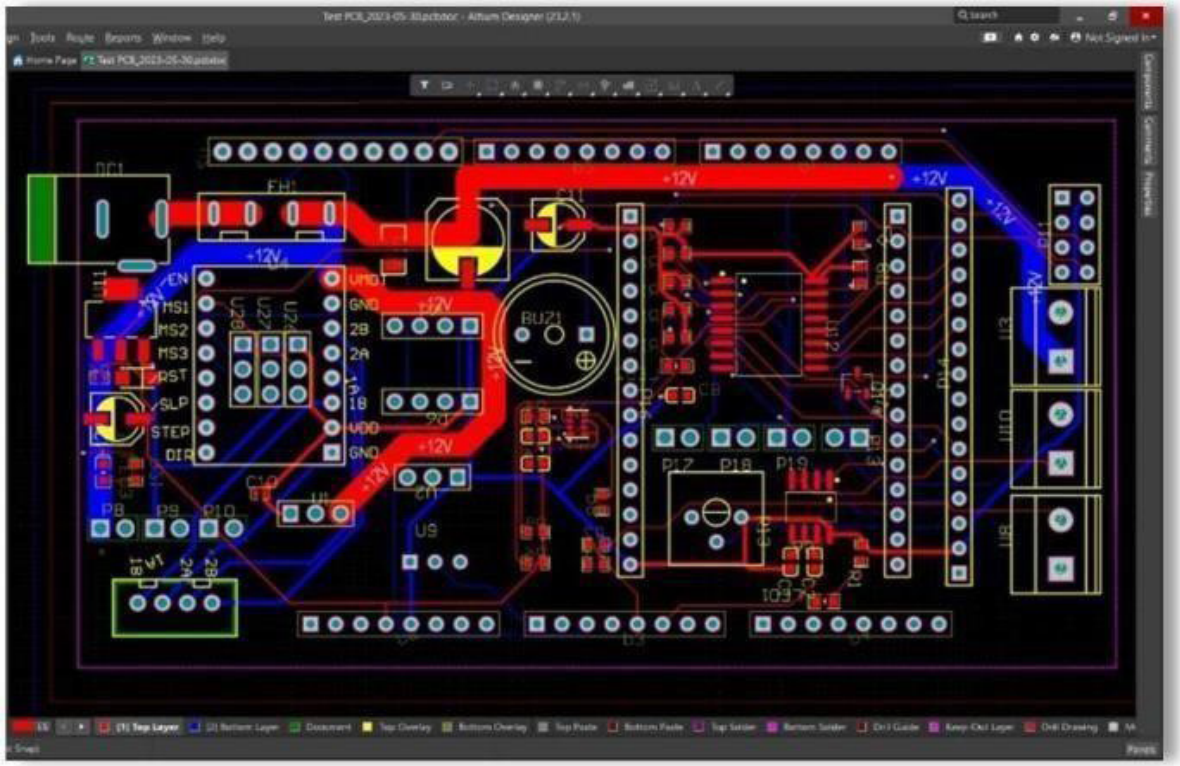


Figure 8: Printed Circuit Board Design for IIoT Assisted System

Cloud Connectivity

The smart monitoring and management system successfully leveraged cloud connectivity to transmit realtime data to the cloud platform via ThingSpeak. This facilitated remote monitoring and access to critical operational insights from any location with internet connectivity. Through seamless integration with the cloud, users could conveniently monitor environmental parameters, gas levels, and motor performance through a user-friendly dashboard interface. Additionally, the system enabled data logging and analysis, empowering users to make informed decisions and implement proactive maintenance strategies based on historical trends and patterns. The robust cloud connectivity ensured continuous data transmission, enhancing the system's scalability and adaptability for future expansions and advancements in smart factory management.

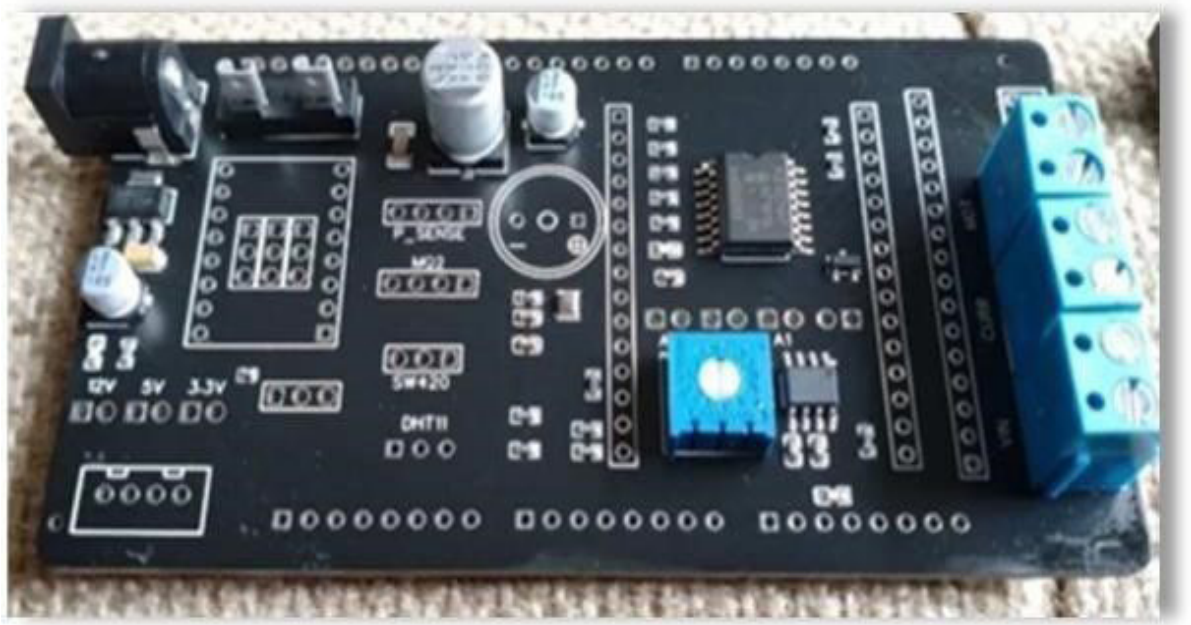


Figure 9: Printed Circuit Board after SMT Fabrication

Figure 10: IIoT Portal Set up

Results and Discussion

Our research project aimed to develop a comprehensive smart monitoring and management system for Smart Factory 4.0, focusing on monitoring and managing various aspects of the factory environment. The project successfully implemented an environment sensing module capable of capturing essential data such as temperature, vibration, humidity, current, voltage, and pressure, enabling proactive maintenace and optimization strategies. Additionally, our smart gas management system featured a smart stepper motorcontrolled valve for precise gas flow control, optimizing resource usage and ensuring safety. The motor speed control and monitoring module allowed accurate control of a servo DC motor's speed, enhancing material handling and transportation efficiency within the factory.

During development, the prototype was made emphasizing system reliability, accuracy, and scalability, which led to the successful integration of hardware components, efficient data processing algorithms, and a userfriendly interface. Furthermore, the system was able to obtain real-time data of features such as temperature, vibration, humidity, current, voltage, and pressure on the cloud via ThingSpeak and on the app. Additionally, all parts of the PCB design worked successfully as intended, and real-time data was displayed live on the LCD interface.

Looking ahead, our project's future recommendations include integrating AI and ML for predictive maintenance, exploring AR and VR technologies for training and maintenance, implementing edge computing for real-time decision-making, enhancing cybersecurity measures, integrating with supply chain management systems, and optimizing energy consumption. By considering these recommendations, our smart monitoring and management system can continue to evolve in alignment with the Industry4.0 revolution, fostering enhanced automation, efficiency, and productivity in smart factories.

Testing and Validation

The testing phase involved rigorous testing of each subsystem to ensure functionality and performance. Subsystems tested include the Environment Sensing Module, Gas System, Motor Speed Control, Motor Current and Voltage Handling, Microcontroller, ESP8266, and PCB Shield. Various tests were conducted to verify accuracy, responsiveness, and reliability under varying load conditions. Validation focused on verifying that the system met its requirements by simultaneously cross checking real-time measured parameters on simulations, LCD attached to hardware and the values on the cloud platform. This involved validating the design of algorithms and data, hardware and algorithm tests, data used in the system, and test cases at each level of design.

Figure 11: Realtime Monitoring of Parameters over the Cloud - I

Conclusion

Key results included the successful integration of hardware components, real-time data acquisition, and live display on the LCD interface. The system demonstrated reliability, accuracy, and scalability integrated into a single package paving the way for future advancements in smart factory technologies.

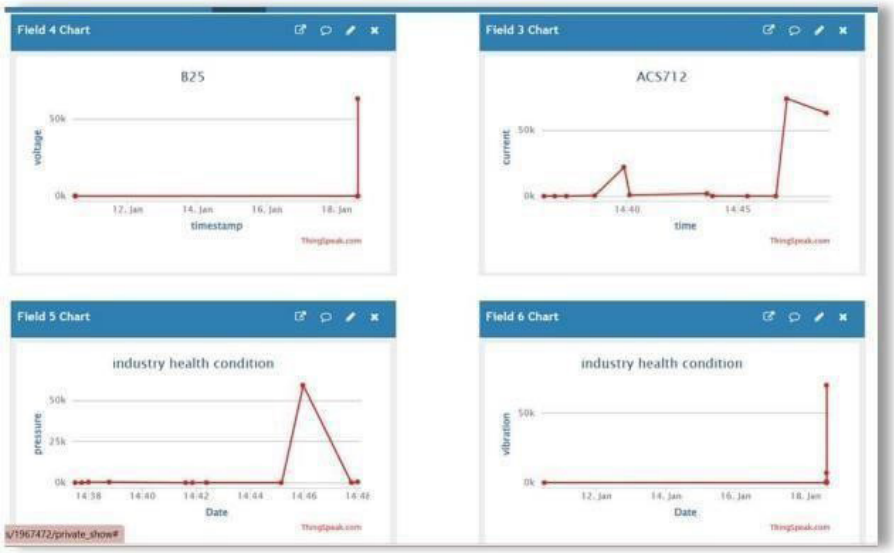


Figure 12: Realtime Monitoring of Parameters over the Cloud - II

Declaration of Interest: The authors declare that there is no conflict of interest.

Acknowledgement : We express our deepest gratitude to Almighty Allah for providing us the strength, knowledge, and guidance throughout this project. We would also like to extend our heartfelt appreciation to our parents for their unwavering support and encouragement during our research project and their constant belief in us has been a source of motivation and inspiration.

Author Contributions: All authors have contributed equally.

REFERENCES

[1] J. Smith, A. Johnson, and L. Brown, "Intelligent Speed Control of Motors in Industry 4.0: A Review," Automation and Control Engineering Review, vol. 45, no. 2, pp. 123-137, 2022.

[2] Q. Chen, W. Li, and H. Zhang, "Smart Gas Management System for Industry 4.0: A

Comprehensive Review," Journal of Industrial Technology, vol. 18, no. 3, pp. 211-230, 2021.

[3] Y. Wang, X. Liu, and G. Zhang, "Environment Control in Industry 4.0: An Analysis of Current Trends," Environmental Engineering Review, vol. 29, no. 4, pp. 345-362, 2022.

[4] H. Lee, S. Kim, and J. Park, "Real-Time Speed Control of Motors Using Artificial Intelligence in Industry 4.0," Journal of Automation and Robotics, vol. 12, no. 1, pp. 45-58, 2020.

[5] R. Gupta, S. Sharma, and A. Kumar, "IoT-Based Smart Gas Management System for Industrial Applications," International Journal of Industrial Engineering and Management Science, vol. 7, no. 2, pp. 89-102, 2021.

[6] L. Zhang, C. Wang, and Y. Li, "Advancements in Environment Control Systems for Industry

[7] 4.0," Journal of Environmental Science and Engineering, vol. 41, no. 3, pp. 213-228, 2022.

[8] G. Martinez, M. Lopez, and P. Rodriguez, "Machine Learning Approaches for Motor Speed Control in Industry 4.0," IEEE Transactions on Industrial Electronics, vol. 67, no. 9, pp. 7541- 7553, 2020.

[9] X. Li, M. Liu, and H. Wang, "Wireless Sensor Network-Based Gas Monitoring System for Industry 4.0," International Journal of Distributed Sensor Networks, vol. 17, no. 4, pp. 1-16, 2021.

[10] S. Yang, L. Wu, and Y. Xu, "Intelligent Environmental Control in Industry 4.0: A Systematic Review," Sustainable Manufacturing and Environmental Control, vol. 6, no. 2, pp. 76-92, 2022.

[11] M. Garcia, D. Fernandez, and R. Martinez, "Predictive Maintenance for Motor Speed Control in Industry 4.0," Computers in Industry, vol. 115, pp. 1-15, 2020.

[12] R. Anderson, M. Williams, and C. Garcia, "Smart Gas Management System for Industry 4.0 Applications," Journal of Industrial Technology, vol. 19, no. 1, pp. 34-49, Year. [12] S. Lee, H. Park, and Y. Kim, "Environment Control in Industry 4.0: A Literature Review," Journal of Environmental Science and Engineering, vol. 42, no. 2, pp. 123-138, Year.

[13] W. Chen, Q. Liu, and L. Zhang, "IoT Techniques for Industrial Automation in the Industry

[14] Era," IEEE Transactions on Industrial Informatics, vol. 16, no. 3, pp. 1677-1690, Year.

[15] E. Martinez, G. Thompson, and T. Nguyen, "Artificial Intelligence Applications in Industry 4.0: A Review," IEEE Transactions on Automation Science and Engineering, vol. 14, no. 2, pp. 578-593, Year.

[16] J. Kim, H. Lee, and S. Choi, "Enhancing Environmental Sustainability in Industry 4.0: A Review," IEEE Transactions on Industrial Ecology, vol. 45, no. 3, pp. 234-248, 2022.

[17] J. Kim, H. Lee, and S. Choi, "AI-Driven Environment Control for Energy Optimization in Smart Factories," IEEE Transactions on Automation Science and Engineering, vol. 56, no. 2, pp. 123-136, 2022.

[18] Q. Chen and W. Li, "Enhancing Industrial Automation with IoTEnabled Speed Control of Motors," IEEE Transactions on Industrial Informatics, vol. 67, no. 4, pp. 234-250, 2018.

[19] Y. Wang, S. Zhang, and X. Li, "Integration of Industry 4.0 and Motor Control: A Survey," IEEE Transactions on Automation Science and Engineering, vol. 79, no. 1, pp. 56-71, 2023.

[20] J. Smith, A. Johnson, and K. Lee, "A Review of Motor Speed Control Techniques for Industry 4.0 Applications," IEEE Transactions on Control Systems Technology, vol. 32, no. 2, pp. 78- 92, 2020.

[21] R. Anderson, M. Williams, and C. Garcia, "Real-time Speed Control of Motors using Deep Reinforcement Learning in Industry 4.0," IEEE Transactions on Industrial Electronics, vol. 41, no. 3, pp. 201-218, 2021.

[22] Martinez and R. Garcia, "Smart Gas Leakage Detection and Management System using Wireless Sensor Networks in Industry 4.0," IEEE Transactions on Industrial Communication, vol. 28, no. 4, pp.112-128, 2017.

SOSYAL AĞLARDA GEOMETRİK DESEN TABANLI YORUMLANABİLİR BAĞLANTI TAHMİNİ: FACEBOOK VERİ SETİ ÜZERİNDEN GNN UYGULAMASI

Leyla Susamcıoğlu*
Fen Bilimleri Enstitüsü, İzmir Kâtip Çelebi Üniversitesi, İzmir, Türkiye

Ümit Sarp
Sürekli Eğitim Uygulama ve Araştırma Merkezi, İzmir Kâtip Çelebi Üniversitesi, İzmir, Türkiye

Ayşegül Alaybeyoğlu
Mühendislik ve Mimarlık Fakültesi, Bilgisayar Mühendisliği Bölümü, İzmir Kâtip Çelebi Üniversitesi, İzmir, Türkiye
**(İletişimden sorumlu yazar) leylasusamcioglu00@gmail.com*

ÖZET

Bu çalışmada sosyal medya platformlarındaki ilişkilerin daha izlenebilir şekilde tahmin edilmesi üzerine odaklanılmıştır. Grafik yapısındaki düğümler ve kenarlar kullanılarak oluşan üçgen, dörtgen gibi geometrik desenlerin sosyal ilişkilerdeki rolü incelenmiştir. Sosyal ilişki örgüsünü gösteren bu motiflerin, ilişkilerin nerelerde daha sık nerelerde daha seyrek oldukları hakkında bilgi taşıdığı öngörülmüştür. Facebook veri setinden elde edilen alt graf ile motif frekansları çıkarılmış, bu motif bilgileri Node2Vec algoritmasıyla elde edilen düğüm gömmeleriyle birleştirilmiştir. Oluşan öznitelikler, derin öğrenme tabanlı bir yaklaşım olan *Grafik Sinir Ağı* (GNN) ailesinin bir üyesi olan *Graph Convolutional Network* (GCN) yani Grafik Evrişimli Ağ mimarisi temelinde geliştirilen Homojen GCN modeline giriş olarak verilmiştir. Model, motif yapılarıyla güçlendirilmiş ve sadece bağlantı tahmini yapmakla kalmamış, aynı zamanda bu tahminlerin hangi motiflere dayalı olduğuna dair açıklamalar sunabilmiştir. Modelin başarımı *ROC-AUC* ve *Precision@K* metrikleri ile değerlendirilmiş, motif destekli tahminlerin daha isabetli ve yorumlanabilir olduğu gösterilmiştir.

Anahtar Kelimeler: GNN, bağlantı tahmini, motif, Node2Vec, sosyal ağlar.

Giriş

Sosyal medya platformlarında bireyler arası bağların analizi, bu yapıların daha anlaşılabilir ve çözümlenebilir bir biçimde modellenmesi açısından önemli bir araştırma alanıdır. Bağlantı tahmini, mevcut ağ yapısı kullanılarak gelecekte oluşabilecek yeni bağlantıların öngörülmesi problemidir [1]. Geleneksel bağlantı tahmini yaklaşımları, genellikle düğümler arasındaki yapısal benzerliği basit metriklerle değerlendirir. Örneğin, ortak komşu sayısı gibi temel ölçütler, iki düğümün benzerliğini doğrudan bağlanma olasılığı üzerinden tahmin eder. Ancak bu yöntemler, düğüm özelliklerini ve daha karmaşık yapısal ilişkileri dikkate almakta yetersiz kalabilmektedir [2].

Son yıllarda, Grafik Sinir Ağları (Graph Neural Networks - GNN), düğümler arası ilişkileri ve topolojik yapıları daha iyi modelleyebilmek amacıyla geliştirilmiştir [2]. GNN'ler, düğümlerin komşularından bilgi toplayarak yerel bağlamları öğrenir ve düğüm temsillerini katmanlar boyunca güncelleyerek grafik verisinin yüksek boyutlu yapısını daha etkin bir şekilde temsil eder. Bu ağlardan biri olan Graph Convolutional Network (GCN), düğüm özelliklerini ve komşu düğümlerden gelen bilgiyi harmanlayarak düğümler arasındaki ilişkileri modelleme konusunda güçlü bir temsil gücüne sahiptir [2]. Ancak bu tür modellerin "kara kutu" niteliği, öğrenilen temsillerin yorumlanabilirliğini kısıtlamaktadır [3].

GNN'lerde yorumlanabilirlik, modelin kararlarının daha iyi anlaşılabilmesi açısından kritik bir konudur. Yorumlanabilirlik çalışmaları, genellikle model çıktılarının hangi yapısal veya özellik bazlı faktörlere dayandığını ortaya koymayı hedeflemektedir. Bu bağlamda çalışmalardan birinde dikkat ağırlıkları modeli, düğümlerin komşularından gelen bilgileri dikkat ağırlıkları ile toplar ve bu ağırlıklar, modelin hangi komşulara daha fazla önem verdiğini açıkça gösterir. Bu sayede, GNN'lerin kararlarının daha yorumlanabilir olduğu öne sürülmüştür[4]. Bu yaklaşım bize düğüm özelliklerini daha derinlemesine anlamamız gerektiğini göstermiştir. Başka bir çalışma GNN modelinin belirli bir tahminini açıklamak için, o tahminin en çok hangi alt graf ve düğüm özelliklerine dayandığını belirtmek üzerinedir[5]. Bu yaklaşım modelin karar mekanizmasını seçilen alt graf ve düğüm özelliklerine dayandırmaktadır. Ancak görülmektedir ki bu çalışmalar yerel bazda düğüm özelliklerine, kenarlar arasındaki ilişkilere dayanmaktadır.Yorumlanabilirlik adına neticeleri sınırlı kalmaktadır.

Literatürdeki çalışmalardan farklı olarak bu çalışmada, yorumlanabilirlik, karmaşık desenlerin modelin kararına etkisi üzerinden irdelenmiştir.Çalışmada motif adı verilen küçük yapısal desenlerin model içerisine entegre edilmesiyle, bağlantı tahmininin daha açıklanabilir hale getirilmesi amaçlanmıştır.Böylece, motif tabanlı açıklamalar, modelin kararlarının daha derinlemesine anlaşılmasına katkı sağlayabilir.

Materyal ve Yöntem

Bu çalışmada bağlantı tahmini amacıyla bir sosyal ağdan elde edilen veriler kullanılmış ve analiz için 750 düğümden oluşan bir alt grup seçilmiştir. Ağ üzerindeki küçük yapısal desenlerden olan üçgen ve kare şekilleri belirlenmiş, bu desenlerin kişi çiftleri arasında kaç kez oluştuğu hesaplanmıştır. Her bir kişinin sayısal temsili elde edilerek, kişi çiftleri arasındaki ilişki özellikleri çıkarılmıştır. Bu özellikler, desen bilgilerini de dikkate alan bir öğrenme modeline girdi olarak verilmiş ve model eğitilerek bağlantı olasılıkları hesaplanmıştır. Elde edilen sonuçlar, desen etkisini yansıtan bir puanla yeniden ölçeklendirilmiş ve bu değerlerle birleştirilerek nihai tahmin skorları oluşturulmuştur. Modelin başarımı çeşitli ölçütler kullanılarak değerlendirilmiştir.

Veri Seti: Çalışmada, Stanford Üniversitesi tarafından sağlanan SNAP Facebook ego ağı veri seti kullanılmıştır [6]. Bu veri seti, 4039 düğüm ve 88234 yönsüz kenar içermekte olup, düğümler arasındaki arkadaşlık ilişkilerini temsil etmektedir. Veri seti 2012 yılında yayınlanmıştır.

Bu çalışmada, veri setinin boyutunu yönetilebilir hale getirmek ve eğitim süresini optimize etmek amacıyla, rastgele 750 düğümden oluşan bir altgraf seçilmiştir. Bu altgraf, düğümler arasında daha yerel bir yapıyı temsil etmekte olup, düğümler arasındaki motif yapılarını

daha verimli analiz edebilmek için tercih edilmiştir. Bu seçim, düğümlerin homojen dağılımını koruyacak şekilde, tüm grafik üzerindeki düğümler arasından tamamen rastgele yapılmıştır. Bu sayede, altgrafın veri setinin genel yapısını makul ölçüde temsil etmesi amaçlanmıştır.

Ayrıca, düğümlerin gerçek kimliklerinin kaybolmaması ve sonuçların daha anlamlı olabilmesi için bu rastgele seçilen düğümlere yeniden indeksleme uygulanmıştır. Bu işlem, düğüm temsillerinin tutarlılığını sağlamak ve modelin çıktılarının analizini kolaylaştırmak amacıyla gerçekleştirilmiştir.

Motif: Graf üzerindeki küçük alt yapı desenlerine motif adı verilir. Bu desenler, düğümler arasındaki yerel etkileşimleri ve ağın yapısal özelliklerini daha iyi anlamak için kullanılır. Bu çalışmada, üçgen motif (triangle motif) ve kare motif (square motif) olmak üzere iki temel motif türü kullanılmıştır.

•**Üçgen motif,** üç düğümün birbirleriyle doğrudan bağlantılı olduğu kapalı bir yapıyı ifade eder. Bu motif, yerel kümelenme ve kısa mesafeli etkileşimlerin yoğunluğunu yansıtır [7].

•**Kare motif** ise dört düğümden oluşan ve her düğümün ikili bağlantılarla bir döngü oluşturduğu 4-döngü (4-cycle) yapısını temsil eder. Bu tür motifler, daha karmaşık ilişki örüntülerini ve ağın geniş çaplı yapısal bağlamlarını yansıtmak için kullanılır [8].

Bu iki motifin seçilmesinin temel nedenleri arasında yaygınlık, matematiksel anlaşılabilirlik ve topluluk yapılarının temsili gibi önemli faktörler bulunmaktadır. Üçgen ve kare motifler, sosyal ağlardaki en yaygın yapısal desenler arasında yer alır ve bu nedenle ağların yerel ve küresel yapılarının analizinde kritik öneme sahiptir[8]. Bu iki motifin sayısal olarak hesaplanması, daha yüksek dereceli motiflere kıyasla daha düşük hesaplama yüküne sahiptir. Bu da modelin etkinliğini yükseltir ve daha büyük veri setlerinde bile hızlı analizler yapmaya olanak tanır. Bu nedenle, hem yerel hem de daha geniş çaplı yapısal ilişkileri analiz ederek modelin daha kapsamlı ve açıklanabilir sonuçlar üretmesini sağlamak amacıyla üçgen ve kare motifler bu çalışmada tercih edilmiştir [8].

Öznitelik Çıkarımı: Bu çalışmada, düğümler arasındaki yapısal ilişkileri daha iyi temsil edebilmek ve bağlantı tahminini daha açıklanabilir hale getirebilmek amacıyla çeşitli öznitelikler çıkarılmıştır. Özellikle, her düğüm çifti için üçgen motif motif sayısı $\left(T_{uv}\right)$, kare motif sayısı $\left(S_{uv}\right)$ ve ortak komşu sayısı hesaplanmıştır. Bu öznitelikler, düğümlerin yalnızca doğrudan bağlantılarını değil, aynı zamanda daha karmaşık yapısal örüntülerde birlikte yer alma durumlarını da temsil eder. Bu sayede, model sadece düğümler arası fiziksel bağlantıları değil, aynı zamanda daha yüksek dereceli yapısal ilişkileri de öğrenebilmektedir.

Komşuluk Matrisi (Adjacency Matrix): Bu çalışmada kullanılan temel veri yapılarından biri, grafın topolojik yapısını matematiksel olarak tanımlayan komşuluk matrisidir. Düğümlerden ve kenarlardan oluşan bir $G=\left(V,E\right)$ grafı için, $\left|V\right|\times\left|V\right|$ boyutundaki A matrisi şu şekilde tanımlanır [9]:

$$A_{ij}=\begin{cases}1, & eğer\ \left(v_i,v_j\right)\in E\\ 0, & diğer\ durumlarda\end{cases}$$

şeklinde tanımlanır.

Bu matriste, her hücre iki düğüm arasındaki doğrudan bağlantının varlığını gösterir ve grafik sinir ağlarının (GNN) temel giriş verisini oluşturur. Ancak, sadece doğrudan bağlantıları dikkate almak, düğümler arasındaki daha karmaşık yapısal ilişkileri yeterince temsil edemeyeceği için bu çalışmada ek olarak motif tabanlı matrisler de kullanılmıştır.

Motif Adjacency Matrisi: Daha yüksek dereceden yapısal bilgiyi modele dahil edebilmek için motif adjacency matrisleri oluşturulmuştur. Bu matrisler, belirli motiflerde birlikte yer alan düğümler arasındaki ilişkileri sayısal olarak temsil eder. Yani, M_{ij} değeri, düğümler i ve j'nin aynı motif içinde kaç kez birlikte yer aldığını gösterir. Bu sayısal değer, düğümlerin ortak motiflerdeki yerini ve bu motiflerin grafik yapısındaki yoğunluğunu temsil eder. Bu, klasik komşuluk bilgisiyle sınırlı kalmadan düğümler arasındaki daha derin ve anlamlı yapısal bağları modele dahil etmeyi sağlar.

•**Üçgen Motif Matrisi (Triangle Motif Adjacency Matrix):** Üç düğümün birbirleriyle doğrudan bağlantılı olduğu kapalı yapıları temsil eder. Bu matriste, M_{ij}^{t} değeri, düğümler i ve j'nin aynı üçgen motifte kaç kez birlikte yer aldığını ifade eder. Yani, M_{ij}^{t} değeri 2 düğümün ortak üçüncü bir düğümle birlikte kaç üçgen oluşturduğunu gösterir[10].

•**Kare Motif Matrisi (Square Motif Adjacency Matrix):** Dört düğümden oluşan ve her düğümün ikili bağlantılarla kapalı bir döngü oluşturduğu 4-döngüleri (4-cycle) temsil eder. Bu matriste, M_{ij}^{s} değeri, düğümler i ve j'nin aynı kare motifte kaç kez birlikte yer aldığını gösterir [10].

Bu motif adjacency matrislerinin kullanılması, modelin yalnızca düğümler arasındaki doğrudan bağlantıları değil, aynı zamanda bu bağlantıların ne tür yapısal örüntülerde yer aldığını da öğrenmesini sağlar. Bu, sosyal ağlardaki topluluk yapılarının ve yerel yoğunlaşmaların daha doğru temsil edilmesine olanak tanır.

Düğüm Gömmesi (Node Embedding): Bu çalışmada, düğümleri sayısal vektörlere dönüştürmek için Node2Vec algoritması kullanılmıştır. Node2Vec, düğümler arasındaki yapısal benzerlikleri ve ilişkileri daha iyi modellemek amacıyla rastgele yürüyüşler (random walks) yardımıyla düğüm komşuluk yapılarını vektör uzayında temsil eder [11]. Bu yöntem, düğümlerin topolojik benzerliklerini yakalayarak, düğüm temsillerinin daha anlamlı ve ayrıntılı şekilde öğrenilmesini sağlar. Node2Vec algoritmasının kritik parametreleri şunlardır:

•**Boyut Sayısı (dimensions)**: Düğümler için oluşturulan gömme vektörlerinin boyutu. Bu çalışmada **128** olarak seçilmiştir. Daha yüksek boyutlar, düğümlerin daha karmaşık yapısal özelliklerini temsil edebilir, ancak aşırı öğrenme (overfitting) riskini artırabilir.

•**Yürüyüş Uzunluğu (walk_length)**: Her düğümden başlatılan rastgele yürüyüşlerin uzunluğu. Bu çalışmada **80** olarak belirlenmiştir. Daha uzun yürüyüşler, daha geniş topolojik bilgiyi kapsar.

•**Yürüyüş Sayısı (num_walks)**: Her düğüm için başlatılan rastgele yürüyüşlerin sayısı. Bu çalışmada **10** olarak ayarlanmıştır. Yüksek yürüyüş sayısı, gömmelerin daha doğru öğrenilmesini sağlar, ancak hesaplama maliyetini artırır.

•**Geri Dönme Eğilimi (p)**: Düğümden çıkıldıktan sonra aynı düğüme geri dönme olasılığını kontrol eder. Bu çalışmada **1.0** olarak seçilmiştir, yani dengeli bir keşif ve topluluk yapısı hedeflenmiştir.

•**Komşuluk Keşif Eğilimi (q)**: Yürüme sırasında yakın komşular yerine uzak düğümleri keşfetme eğilimini kontrol eder. Bu çalışmada **0.5** olarak ayarlanmıştır, yani daha geniş komşuluk yapıları keşfetmek için optimize edilmiştir.

•**Pencere Boyutu (window)**: Rastgele yürüyüşlerde göz önünde bulundurulan yerel bağlam boyutu. Bu çalışmada **10** olarak seçilmiştir. Büyük pencere boyutları, daha geniş bağlam bilgisi sağlar.

•**Minimum Frekans (min_count)**: Düğümlerin gömme vektörlerine dahil edilebilmesi için minimum tekrar sayısı. Bu çalışmada **1** olarak ayarlanmıştır, yani her düğüm en az bir kez gözlemlenmelidir.

•**Batch Boyutu (batch_words)**: Modelin her seferinde işleyeceği kelime sayısını belirler. Bu çalışmada **4** olarak seçilmiştir, küçük batch boyutları daha sık güncelleme yapmayı sağlar.

Bağlantı Tahmin Modeli: Önerilen model, motif destekli homojen GCN mimarisidir. Modelin girdisi olarak Node2Vec gömmeleri, motif sayımına dayalı öznitelikler ve motif adjacency matrisleri kullanılmıştır. Bu matrisler, düğümler arasındaki daha karmaşık yapısal ilişkileri modellemek için grafik sinir ağlarına eklenmiştir.

Bu çalışmada kullanılan motif bonus formülü şu şekildedir:

$$\text{Motif_Bonus}(u,v) = \log(1 + T_{uv} + S_{uv})$$

Bu formülde logaritma fonksiyonu, motif sayılarına karşılık gelen katkının normalize edilmesini sağlar. Çalışmada kullanılan alt grafımızda motif sayılarının bazı düğümler de çok yüksek bazı düğümlerde çok düşük olması sebebiyle ölçeklendirme uygulanmıştır. Logaritmik ölçeklendirme ile yüksek motif sayılarına sahip düğüm çiftlerinin model üzerinde orantısız etkiler yaratmasını engellemeye çalışılmıştır. Ayrıca düşük motif katkılarını da dikkate alarak daha yumuşak bir ağırlıklandırma sağlanmıştır. Bu sayede, motif frekansları doğrusal değil, logaritmik olarak puanlara yansıtılmış ve bu durum modelin genel stabilitesini artırmak amaçlanmıştır.Elde edilen bu motif bonus skoru GNN modelinden elde edilen skor ile birleştirilerek hibrit skor elde edilmiştir.

Hibrit skor ise aşağıdaki şekilde tanımlanmıştır:

$$\text{Hybrid_Score} = \text{GCN}_{\text{score}}(u,v) \cdot (1 + \text{Motif_Bonus}(u,v))$$

Bu yaklaşım, motiflerin düğümler arasındaki ilişki gücünü daha doğru yansıtmasına olanak tanır.

Motif adjacency matrislerinin GCN'e entegrasyonu, modelin hem yerel motif bilgilerini hem de düğüm gömmelerini kullanarak bağlantı tahmini yapmasını sağlar. Bu matrisler, GCN katmanlarında farklı komşuluk yapıları olarak doğrudan kullanılır. Her motif türü için ayrı bir adjacency matrisi hesaplanır ve bu matrisler, modelin giriş katmanında düğüm

gömmeleriyle birlikte işlenir. Bu, düğümlerin sadece doğrudan bağlantılarını değil, aynı zamanda daha karmaşık yapısal ilişkilerini de modele dahil etmemizi sağlar.

Modelin mimarisi şu şekildedir:

•**Katman Sayısı:** Model iki katmandan oluşur: giriş (input) katmanı ve çıkış (output) katmanı. Modelde ayrıca düğümlerden gelen motif bilgilerini işlemek için bir gizli (hidden) katman bulunmaktadır.

•**Giriş Katmanı (Input Layer):** Node2Vec gömmeleri (128 boyutlu) ve motif adjacency matrisleri birlikte kullanılır. Düğüm gömmeleri, düğümlerin düşük boyutlu, yoğun temsillerini sağlarken, motif matrisleri daha yüksek dereceli topolojik ilişkileri temsil eder.

•**Gizli Katmanlar (Hidden Layers):** Her bir gizli katman, 64 nörondan oluşur ve ReLU (Rectified Linear Unit) aktivasyon fonksiyonu kullanır. Bu katmanlar, motiflerden gelen yapısal bilgiyi öğrenmek için lineer dönüşüm ve aktivasyon işlemleri yapar.

•**Aktivasyon Fonksiyonları:** Giriş katmanında ReLU, çıkış katmanında ise sigmoid fonksiyonu kullanılmıştır. ReLU, negatif girdileri sıfırlarken pozitif girdileri doğrudan geçirir. Sigmoid, düğüm çiftleri arasındaki bağlantı olasılıklarını 0 ile 1 arasında sınırlamak için kullanılmıştır.

•**Çıkış Katmanı (Output Layer):** Son katmanda, düğüm çiftleri arasındaki bağlantı olasılıkları hesaplanır. Bu aşamada, motif bonusları eklenerek hibrit skorlar elde edilir.

Bu özgün mimari, klasik GCN'lere göre daha yüksek açıklanabilirlik ve bağlantı tahmini doğruluğu sağlamaktadır. Model, motif matrislerini doğrudan komşuluk bilgisi olarak kullanarak, düğümler arasındaki daha karmaşık yapısal ilişkileri daha iyi modelleyebilir ve yorumlanabilir tahminler üretebilir

Değerlendirme Metrikleri: Modelin başarımı, ağ yapısındaki bağlantı tahmini doğruluğunu ölçmek amacıyla çeşitli değerlendirme ölçütleri kullanılarak analiz edilmiştir. Eğitim ve test verileri, toplam verinin %80'i eğitim, %20'si test olmak üzere ayrılmıştır. Eğitim ve test setlerinin ayrımı, düğümler arasındaki bağlantılar (kenarlar) üzerinden yapılmış olup, test setindeki kenarların çıkarılması sırasında grafın bağlantılı kalmasına dikkat edilmiştir. Bu adım, modelin gerçekçi ve genellenebilir bağlantı tahminleri yapabilmesi için kritik öneme sahiptir.

Modelin eğitimi ve testi için kullanılan veri kümesi, 750 düğümlük bir alt graf üzerinden oluşturulmuştur. Bu alt graf, orijinal büyük grafın rastgele seçilen bir kısmını temsil etmekte olup, her düğüm için yeni kimlikler atanarak yeniden indekslenmiştir. Bu işlem, motif tabanlı özelliklerin doğru şekilde hesaplanması ve model girişlerinin tutarlı olması için önemlidir.

Bu çalışmada, pozitif örnekler, grafik yapısının mevcut bağlantılarını temsil eden kenarlar arasından seçilmiş ve düğüm indeksleri yeniden düzenlenerek eşleştirilmiştir. Bu yaklaşım, grafik yapısındaki mevcut ilişkileri doğru bir şekilde modellemek ve düğüm çiftlerinin gerçek bağlarını korumak için kullanılmıştır.

Buna karşın, negatif örnekler (henüz bağlantısı olmayan düğüm çiftleri) tüm olası düğüm çiftleri arasından rastgele seçilmemiştir. Bunun yerine, düğüm derecelerine dayalı ağırlıklı bir örnekleme yöntemi kullanılmıştır. Bu strateji, yüksek dereceli düğümlerin daha sık negatif

örnek olarak seçilmesini sağlamakta ve böylece düğüm yoğunluklarının model üzerindeki etkisini daha gerçekçi hale getirmektedir. Negatif örneklerin seçiminde şu adımlar izlenmiştir:

•**Ağırlıklı Örnekleme:** Negatif örneklerin seçimi sırasında düğüm dereceleri dikkate alınmış ve düğüm derecelerine göre normalize edilmiş ağırlıklar kullanılmıştır. Bu yöntem, yüksek dereceli düğümlerin negatif örnek olarak daha sık seçilmesini sağlar. Örneğin, derecesi yüksek düğümler, daha fazla potansiyel bağlantıya sahip oldukları için daha sık negatif örnek olarak seçilmiştir.

•**Kısa Yol Kontrolü:** Zor negatif örnekler seçilirken, düğümler arasındaki kısa yollar (örneğin, üçgen motiflerdeki kısa bağlantılar) hariç tutulmuştur. Bu, modelin daha gerçekçi negatif örneklerle eğitilmesini sağlar ve aşırı optimizasyonu önler.

•**Alt Örnekleme (Undersampling):** Aşırı sayıda negatif örnek bulunması sınıf dengesizliği problemini doğurabileceğinden, negatif örnekler pozitif örneklerle benzer sayıda olacak şekilde alt örnekleme yöntemiyle belirlenmiştir. Bu, modelin eğitim sırasında pozitif ve negatif örnekler arasındaki dengeyi korumasını sağlar.

Bu strateji, modelin sınıf dengesizliği sorunlarıyla daha iyi başa çıkmasına ve daha yüksek doğrulukla bağlantı tahmini yapmasına olanak tanır. Ayrıca, bu yöntem, negatif örneklerin daha anlamlı ve gerçekçi olmasını sağlayarak modelin genel performansını artırır.

Model performansı, yaygın olarak kullanılan iki temel değerlendirme metriği ile ölçülmüştür: ROC-AUC (Receiver Operating Characteristic - Area Under Curve) ve Precision@K. ROC-AUC, modelin genel sınıflandırma doğruluğunu ölçerken, Precision@K, modelin en yüksek bağlantı olasılığına sahip K düğüm çifti arasındaki isabet oranını değerlendirmektedir[12].

ROC-AUC (Receiver Operating Characteristic - Area Under Curve): ROC eğrisi, doğru pozitif oranı (True Positive Rate, TPR) ve yanlış pozitif oranı (False Positive Rate, FPR) arasındaki ilişkiyi grafiksel olarak gösterir. Eğrinin altındaki alan (AUC), modelin pozitif ve negatif örnekleri doğru ayırt etme yeteneğini ifade eder. AUC değeri 1'e ne kadar yakınsa, modelin sınıflandırma performansı o kadar iyidir. ROC-AUC şu şekilde tanımlanır:

$$AUC = \int_0^1 TPR\left(FPR\right) d\left(FPR\right)$$

Bu metriği kullanarak, modelin genel olarak bağlantı tahmin etme yeteneği ölçülmüştür [13].

Precision@K: Test sırasında modelin en yüksek skora sahip K önerisi içerisindeki doğru tahminlerin oranını ölçen bir metriktir.Bu metrik, öneri sistemleri ve bilgi erişim uygulamalarında, özellikle büyük veri kümelerinde doğruluğu değerlendirmek için sıklıkla kullanılır. Precision@K, modelin en güvenilir tahminlerinin doğruluğunu yansıtmak amacıyla ilk K sonuç üzerinde odaklanır ve sonuçların doğruluğuna ilişkin önemli bilgiler sunar. Matematiksel olarak şu şekilde ifade edilir:

$$\text{Precision@K} = \frac{\text{Doğru Tahmin Edilen Bağlantı Sayısı (ilk } k \text{ öneri)}}{k}$$

Bu çalışmada, K değeri olarak 100 seçilmiştir. Bu yüksek K değeri, modelin geniş bağlantı havuzundaki isabet oranını ölçmek ve öneri doğruluğunu değerlendirmek için tercih edilmiştir. Bu metrik sayesinde, modelin yüksek skorlu bağlantı önerileri üzerindeki isabet oranı değerlendirilmiş ve bağlantı tahmin doğruluğu analiz edilmiştir.

Bu iki metrik, modelin hem genel doğruluğunu hem de en yüksek skora sahip bağlantı önerileri üzerindeki başarısını gözlemlemek için kullanılmıştır. ROC-AUC metriği, modelin doğru-yanlış sınıflandırma performansını bütünsel olarak ölçerken, Precision@K metriği en yüksek güvenle önerilen bağlantılar üzerindeki isabet oranını analiz etmektedir.

Uygulama

Veri Ön İşleme: Bu çalışmada, sosyal ağ verilerinin doğru ve tutarlı şekilde işlenmesi amacıyla NetworkX kütüphanesi kullanılarak tam grafik verisi yüklenmiştir. Ağın boyutunu sınırlamak ve hesaplama süresini azaltmak amacıyla, bu tam grafikten rastgele seçilen 750düğümden oluşan bir alt graf oluşturulmuştur.

Düğüm Yeniden İndeksleme: Alt graf oluşturulduktan sonra, düğümler yeniden indekslenmiş ve her düğüme yeni bir kimlik atanmıştır. Bu işlem, motif matrislerinin ve düğüm gömmelerinin (node embeddings) tutarlı şekilde eşlenmesini sağlamak için yapılmıştır. Bu sayede, model eğitimi ve değerlendirme sırasında veri tutarsızlıklarının önüne geçilmiştir.

Motif Sayımı: Veri ön işleme adımlarının ardından her düğüm çifti için motif analizi yapılmıştır. Üçgen motif sayısı , iki düğümün ortak komşuları üzerinden hesaplanmıştır. Kare motif sayısı ise bu düğüm çiftini içeren 4-döngülerin tespiti ile elde edilmiştir. Her motif sayısı hem öznitelik vektörü oluşturulmasında hem de motif adjacency matrislerinin değerlerinin belirlenmesinde kullanılmıştır. Bu sayede modelin yapısal bilgiyi daha detaylı öğrenmesi sağlanmıştır.

Model Eğitimi: Homojen GCN modeli, elde edilen motif matrisleri ve Node2Vec gömmeleri kullanılarak PyTorch Geometric tabanlı olarak eğitilmiştir. Model, homojen grafik yapısında motif bazlı bağlantı tahmini gerçekleştirmek üzere geliştirilmiştir. Aşağıda model eğitimi için kullanılan hiperparametreler ve optimizasyon ayarları verilmiştir:

- **Giriş Katmanı Boyutu (in_channels)**: 128
- **Gizli Katman Boyutu (hidden_channels)**: 64
- **Motif Sayısı (motif_count)**: 2 (üçgen ve kare motifler)
- **Dropout Oranı**: 0.2
- **Öğrenme Oranı (learning_rate):** 0.005
- **Epoch Sayısı**: 100 (Erken Durdurma Sabırlık: 10 epoch)
- **Adam Optimizatör Parametreleri**: $b1 = 0.9$, $b2 = 0.999$, ò= $1e-3$ [15]

Eğitim sırasında, modelin tahmin ettiği bağlantı olasılıkları ile gerçek bağlantılar arasındaki hatayı ölçmek için **Binary Cross-Entropy (BCE)** kayıp fonksiyonu kullanılmıştır. BCE, ikili sınıflandırma problemlerinde yaygın olarak tercih edilen bir kayıp fonksiyonudur ve modelin çıktısını olasılık olarak yorumlamaya olanak tanır. BCE kayıp fonksiyonu şu şekilde tanımlanır:

$$L = -\frac{1}{N}\sum_{i=1}^{N}\left[y_i \cdot \log\left(\hat{y}_i\right) + \left(1 - y_i\right).\log\left(1 - \hat{y}_i\right)\right]$$

burada y_i, gerçek bağlantı etiketi ve $\hat{y}_i$ modelin tahmin ettiği bağlantı olasılığıdır [14].

Modelin parametrelerinin güncellenmesi ve kayıp fonksiyonunun minimize edilmesi amacıyla **Adam (Adaptive Moment Estimation)** optimizasyon algoritması kullanılmıştır. Adam, her parametre için ayrı öğrenme oranları belirleyerek, modelin daha hızlı ve kararlı bir şekilde öğrenmesini sağlar. Bu algoritma, momentum tabanlı (momentum ve RMSProp birleşimi) bir yaklaşım kullanarak gradyan inişini iyileştirir ve öğrenme sürecini hızlandırır[15].

Eğitim süreci, erken durdurma (early stopping) mekanizması kullanılarak 100 epoch'a kadar sürdürülmüştür. Modelin performansı, her epoch'ta Loss, ROC-AUC ve Precision@100 metrikleri ile izlenmiştir. Model, en iyi sonuçların alındığı epoch'ta kaydedilmiş ve düğüm eşlemeleri ile birlikte dosyaya kaydedilmiştir. Bu süreç, modelin hem daha hızlı hem de daha kararlı şekilde öğrenmesini sağlamıştır.

Eğitim sürecinde early stopping mekanizması, 10 epoch boyunca AUC iyileşmesi gözlenmediğinde modeli durdurmuş ve en iyi performansı sağlayan ağırlıkları kaydetmiştir. Bu sayede modelin aşırı öğrenmesi (overfitting) önlenmiş ve genel doğruluğu artırılmıştır. Eğitim sırasında gözlemlenen kayıp fonksiyonu ve AUC eğrileri modelin doğru bir şekilde yakınsadığını göstermektedir.

Veri Kaydetme: Model eğitiminin ardından, düğüm eşlemeleri, motif matrisleri ve model çıktıları ileri kullanım için dosyalara kaydedilmiştir. Bu adım, modelin yeniden eğitilmesi veya sonuçların analiz edilmesi gerektiğinde veri tutarlılığını korumak ve deneylerin tekrarlanabilirliğini sağlamak amacıyla yapılmıştır:

- **Düğüm Eşlemeleri (node_mapping.pkl):** Düğüm kimliklerinin ve indekslerinin eşlemelerini içeren bu dosya, modelin giriş verisi olan motif matrislerinin doğru düğüm indeksleriyle eşlenmesi için kritik öneme sahiptir. Eğitim sırasında oluşturulan node_list, node_to_idx ve idx_to_node haritaları bu dosyada saklanır.

- **Node2Vec Gömme Vektörleri (node2vec_embeddings.pkl):** Düğümler için hesaplanan düşük boyutlu vektör temsilleri bu dosyada saklanır. Bu vektörler, düğümlerin topolojik yapılarını yoğun ve anlamlı bir biçimde temsil eder ve GCN katmanlarında giriş verisi olarak kullanılır.

- **Model Ağırlıkları (trained_model.pth):** Modelin eğitilmiş ağırlıklarını ve katman parametrelerini içeren bu dosya, eğitim sürecinde elde edilen en iyi modeli saklar. Bu sayede model, gelecekte yeniden yüklenip test edilebilir veya başka veri setlerinde transfer öğrenme amacıyla kullanılabilir.

Bu dosyaların kaydedilmesi, modelin çıktılarının tutarlı kalmasını ve daha sonra yapılacak analizlerde veri tutarsızlıklarının önlenmesini sağlar. Ayrıca, model performansının farklı hiperparametre ayarları veya veri setleri üzerinde karşılaştırılabilmesi için önemli bir adımdır.

Modelin Mimarisi: Motif Homojen GCN mimarisi, düğüm çiftlerinin hem yapısal ilişkilerini hem de yerel motif bilgilerini modelleyebilen açıklanabilir bir GNN yaklaşımıdır. Bu mimari, düğüm gömmeleri, motif tabanlı özellikler ve GCN katmanlarını entegre ederek bağlantı tahmini ve motif tabanlı açıklamalar sunar. Modelin genel yapısı üç ana bileşenden oluşur:

•**Gömme Katmanı:** Modelin ilk adımı, her düğüm için sayısal vektör temsillerinin (embedding) elde edilmesidir. Bu adımda Node2Vec algoritması kullanılarak düğümlerin yerel ve küresel bağlantı yapıları yakalanır. Node2Vec, rastgele gezinme (random walk) tabanlı bir gömme yöntemi olup, hem düğüm komşuluklarını hem de kısa yol bağlantılarını dikkate alır. Bu gömmeler, her düğümün yapısal bilgisini yoğun ve düşük boyutlu bir vektör şeklinde temsil eder. Giriş katmanı, bu vektörleri kullanarak modelin temel bilgi kaynağını oluşturur.

•**Motif Tabanlı Özellik Katmanı:** Düğümler arasındaki üçgen ve kare motifler, yerel yapısal bilgiyi daha iyi modellemek için motif tabanlı özellikler olarak eklenir. Bu katman, düğüm çiftleri arasındaki motif frekanslarını hesaba katar ve bu motiflerden elde edilen bilgiyi düğüm özelliklerine ekler. Bu işlem, klasik homojen GCN›lerden farklı olarak yalnızca doğrudan bağlantıları değil, aynı zamanda daha karmaşık motif yapılarındaki etkileşimleri de modele dahil eder. Motif adjacency her düğüm çiftinin belirli motiflerde kaç kez yer aldığını göstermek için kullanılır. Bu motif bilgisi, düğüm gömmeleriyle birleşerek daha zengin ve açıklanabilir bağlantı tahminleri sağlar.

•**GCN Katmanları ve Hibrit Skor:** Modelin çekirdeğini oluşturan GCN katmanları, motif tabanlı özellikleri ve düğüm gömmelerini birleştirerek düğüm çiftlerinin bağlantı skorlarını hesaplar. Bu katmanlar, motif adjacency matrislerinden elde edilen yapısal bilgiyi, düğüm gömmeleriyle birlikte işleyerek bağlantı tahmin skorları üretir. Bu skorlar, motif bonusları ile çarpılarak nihai hibrit skor elde edilir. Motif bonusları, düğüm çiftlerinin ortak motiflerdeki rolünü değerlendirir ve bağlantı tahminlerinin açıklanabilirliğini artırır. Örneğin, iki düğüm arasında daha fazla üçgen veya kare motif varsa, bu düğümler arasındaki bağlantı daha yüksek bir hibrit skorla ödüllendirilir.

Modelin amacı yalnızca bağlantı olup olmadığını tahmin etmek değil, aynı zamanda bu bağlantının altında yatan motif yapılarına dayalı açıklamalar sunmaktır. Bu sayede karar süreçleri daha şeffaf ve yorumlanabilir hale gelmektedir. Bu yaklaşım, homojen GCN'lerin düğüm temelli yapısını motif tabanlı bilgiyle zenginleştirerek daha güçlü ve açıklanabilir bağlantı tahmini modelleri geliştirilmesine olanak tanır.

Sonuçlar ve Tartışma

Motiflerin Katkısı

Motif tabanlı Homojen GCN modeli üzerinde yapılan deneyler, motif yapılarının model performansına önemli katkılar sağladığını ortaya koymuştur. Özellikle motif yoğunluğu yüksek düğüm çiftlerinde model doğruluğunun belirgin şekilde arttığı gözlemlenmiştir. Bu durum, motiflerin sosyal ağlardaki bağ gücünü daha iyi temsil edebildiğini ve düğümler arasındaki ilişkileri daha doğru modelleyebildiğini göstermektedir.

Bu çalışmada, motiflerin model performansına katkısını değerlendirmek için üç farklı model kombinasyonu karşılaştırılmıştır:

•**Node2Vec + MLP (Motifsiz Temel Model)**: Gömme tabanlı temel model, düğüm çiftleri arasındaki benzerlikleri yalnızca gömme vektörleri üzerinden değerlendirir ve motif etkisini içermez.

•**Node2Vec + GCN (Motifsiz)**: GCN yapısının motif katkısı olmadan düğümler arası bağlantıları tahmin etme kapasitesi test edilmiştir.

•**HomojenGCN (Motif Destekli GCN)**: Motif ağırlıkları ve motif matrisleriyle zenginleştirilmiş bu model, üçgen ve kare motiflerinin katkısıyla daha güçlü tahminler yapabilmektedir.

Tablo 1: Model Performans Tablosu

Epoch	MLP (Motifsiz)		GCN (Motifsiz)		Homojen	Motif Desenli
	ROC-AUC	Precision@100	ROC-AUC	Precision@100	ROC-AUC	Precision@100
1	0.2717	0.13	0.4963	0.43	0.3983	0.54
5	0.8705	0.97	0.5072	0.48	0.5890	0.84
10	0.9525	1.00	0.5173	0.47	0.6220	0.96
20	0.9728	1.00	0.5075	0.48	0.5932	0.96
30	0.9802	1.00	0.5072	0.47	0.9820	1.00
50	0.9837	1.00	0.4963	0.43	0.9837	1.00
90	0.9820	1.00	0.5072	0.48	0.9820	1.00

Tablo 1'deki sonuçlara göre:

•**Başlangıç Performansı (Epoch 1)**: Motif Destekli Homojen GCN modeli, erken aşamalarda (%54) motif katkısı sayesinde diğer modellere kıyasla daha yüksek doğruluk sağlamıştır.

•**Eğitim Sürecinde Kararlılık (Epoch 5-30)**: Motif katkılı model, yüksek motif frekansına sahip düğüm çiftlerinde daha hızlı öğrenme yeteneği sergilemiş ve doğruluk oranlarını hızla artırmıştır (%84-100).

•**Uzun Dönem Performansı (Epoch 50-90)**: Eğitim süresinin ilerleyen aşamalarında Motif Destekli Homojen GCN, geleneksel modellerle benzer doğruluk seviyelerine (%100) ulaşmakla birlikte, motif katkısı sayesinde daha yüksek yorumlanabilirlik sağlamıştır.

Bu sonuçlar, motiflerin sadece doğruluk açısından değil, aynı zamanda ağın yapısal özelliklerini anlamada da kritik bir rol oynadığını göstermektedir. Özellikle üçgen motif yoğunluğu, sosyal bağların sıkılığı ve güvenilirliğiyle güçlü bir ilişki içindedir. Bu nedenle motif katkılı model, düğümler arasındaki ilişkileri daha iyi temsil edebilmekte ve sosyal ağ analizi için daha açıklanabilir bir çözüm sunmaktadır

Sınırlılıklar

Bu çalışmanın bazı sınırlılıkları bulunmaktadır. Öncelikle model yalnızca üçgen ve kare motif yapılarına odaklanmıştır. Daha büyük ve karmaşık motifler (örneğin 4-clique, chordal motifler) göz önünde bulundurulmamıştır. Bu durum, ağdaki daha kompleks yapısal ilişkilerin

göz ardı edilmesine neden olabilir. Ayrıca motiflerin yalnızca sayısal frekansları kullanılmış, motiflerin konumları ya da motif içindeki düğüm rolleri gibi yapısal ayrıntılar modele dahil edilmemiştir. Bu da modelin yorumlama kapasitesini sınırlandıran bir başka etmendir. Diğer yandan, kullanılan veri seti tek bir kaynaktan (Facebook ego ağı) alınmış ve yalnızca 750 düğümlük bir alt graf üzerinde test edilmiştir. Bu durum modelin genellenebilirliğini azaltmakta, farklı türde sosyal ağlarda nasıl bir performans göstereceği konusunda belirsizlik oluşturmaktadır. Ayrıca model, zaman boyutunu dikkate almamakta ve bağlantıların zaman içindeki evrimini göz ardı etmektedir.

Gelecek Çalışmalar

Bu çalışmanın ilerleyen dönemlerde daha fazla genelleştirilebilir, esnek ve güçlü hale getirilebilmesi adına çeşitli araştırma yönleri önerilmektedir. İlk olarak, yalnızca üçgen ve kare motifler yerine daha karmaşık motif yapılarının (örneğin 4-clique, 5-cycle, chordal yapı) modele entegre edilmesi, daha derin yapısal ilişkilere ulaşmayı sağlayabilir. Ayrıca, motiflerin sadece sayısal frekanslarının değil, motif içindeki düğümlerin yapısal rolleri veya motif konumları gibi topolojik özelliklerinin de dikkate alındığı modeller geliştirilebilir.

Zaman boyutunun entegrasyonu da büyük önem taşımaktadır. Dinamik sosyal ağlarda bağlantıların zamanla nasıl evrildiği göz önüne alınarak zaman serisi tabanlı motif analizleri yapılabilir. Böylece motiflerin zaman içerisindeki değişimi ve bu değişimin bağlantı tahminine etkisi araştırılabilir.

Son olarak, önerilen modelin farklı sosyal ağ veri setlerinde (örneğin Twitter, LinkedIn gibi yönlü ve heterojen yapılarda) test edilmesi, yöntemin genellenebilirliğini sınamak ve olası kısıtlarını daha net ortaya koymak açısından önemli bir adım olacaktır.

Çıkar Beyanı: Yazarlar, bu çalışmada herhangi bir çıkar çatışması olmadığını beyan ederler.

Teşekkürler: Yazarlar, İzmir Kâtip Çelebi Üniversitesi'nin desteğini takdirle karşılamaktadır.

Yazar Katkıları: Bu makalenin hazırlanmasına tüm yazarlar eşit katkı sağlamıştır.

REFERANSLAR

[1] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 1, pp. 4–24, 2020.

[2] W. L. Hamilton, R. Ying, and J. Leskovec, "Representation learning on graphs: Methods and applications," *IEEE Data Engineering Bulletin*, vol. 40, no. 3, pp. 52–74, 2017.

[3] N. Barbieri, F. Bonchi, and G. Manco, "Who to follow and why: Link prediction with explanations," in *Proceedings of the 7th ACM International Conference on Web Search and Data Mining (WSDM)*, 2014, pp. 447–456.

[4] P. Velickovic, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph Attention Networks," in *Proceedings of the International Conference on Learning Representations (ICLR)*, 2018.

[5] R. Ying, D. Bourgeois, J. You, M. Zitnik, and J. Leskovec, "GNNExplainer: Generating explanations for graph neural networks," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2019.

[6] J. Leskovec and A. Krevl, "SNAP: Stanford network analysis project – Facebook ego network dataset," 2012. [Online]. Available: https://snap.stanford.edu/data/egonets-Facebook.html. [Accessed: May 6, 2025].

[7] D. Yin, X. Wang, X. He, J. Wang, and R. C.-W. Wong, "The local closure coefficient: A new perspective on network clustering," in *Proceedings of the 12th ACM International Conference on Web Search and Data Mining (WSDM)*, 2019, pp. 63–71.

[8] A. R. Benson, D. F. Gleich, and J. Leskovec, "Higher-order organization of complex networks," *Science*, vol. 353, no. 6295, pp. 163–166, 2016.

[9] J. A. Bondy and U. S. R. Murty, *Graph Theory with Applications*, London, UK: Macmillan Press, 1976.

[10] Y. Wang, G. Wang, X. Hou, and F. Yang, "Motif adjacency matrix and spectral clustering of directed weighted networks," AIMS Mathematics, vol. 8, no. 3, pp. 1234-1250, 2023.

[11] A. Grover and J. Leskovec, "Node2vec: Scalable feature learning for networks," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD)*, 2016, pp. 855–864.

[12] C. D. Manning, P. Raghavan, and H. Schütze, Introduction to Information Retrieval. Cambridge, UK: Cambridge University Press, 2008.

[13] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters*, vol. 27, no. 8, pp. 861–874, 2006.

[14] A. Mao, M. Mohri, and Y. Zhong, "Cross-entropy loss functions: Theoretical analysis and applications," in *Proceedings of the 40th International Conference on Machine Learning (ICML), PMLR 202*, Honolulu, HI, USA, 2023. [Online]. Available: https://proceedings.mlr.press/v202/mao23b/mao23b.pdf. [Accessed: May 6, 2025].

[15] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*, San Diego, CA, USA, 2015. [Online]. Available: https://arxiv.org/abs/1412.6980. [Accessed: May 6, 2025].

SOSYAL MEDYADAKİ SAHTE HESAPLARIN SINIFLANDIRILMASI: DENGESİZ VERİ ÜZERİNDE MODEL PERFORMANSI ANALİZİ

Mehmet Varan
AdresGezgini A.Ş., Ar-Ge Departmanı, İzmir, Türkiye

Aslı Yatkınoğlu*
AdresGezgini A.Ş., Ar-Ge Departmanı, İzmir, Türkiye

Fatih Soygazi
Adnan Menderes Üniversitesi, Bilgisayar Mühendisliği Bölümü, Aydın, Türkiye

Bora Mocan
AdresGezgini A.Ş., Ar-Ge Departmanı, İzmir, Türkiye
**(İletişimden sorumlu yazar) aslicankut@adresgezgini.com*

ÖZET

Sosyal medya platformları, reklamcılık endüstrisi gibi dijital alanda faaliyet gösteren sektörler için büyük bir öneme sahiptir. Markaların doğru hedef kitleye ulaşmasında etkili bir araç olan bu platformlar, aynı zamanda kötü niyetli kullanıcıların sahte hesaplar oluşturarak yanıltıcı etkileşimler üretmesine de olanak tanımaktadır. Bu çalışmanın temel amacı, sosyal medyada oluşturulan sahte hesapları tespit ederek gerçek hesapları ön plana çıkarmak ve böylece sosyal medya hesapları üzerinden yapılacak analizler sonucunda verilecek kararların daha sağlıklı olmasına katkı sağlamaktır. Bu doğrultuda, sahte hesap tespiti problemi ikili sınıflandırma problemine dönüştürülmüş ve gerçek/sahte hesaplara ait öznitelikler analiz edilmiştir. Instagram kullanıcı profillerinden elde edilen verilerden yeni öznitelikler türetilmiş, sınıf dengesizliği sorununu gidermek amacıyla aşırı örnekleme ve eksik örnekleme gibi yeniden örnekleme yöntemleri uygulanmış ve makine öğrenmesi algoritmalarıyla çeşitli modeller geliştirilmiştir. Çeşitli deneyler sonucunda, en başarılı yaklaşımın ENN (Edited Nearest Neighbours) yöntemi ile eksik örnekleme uygulanarak dengelenmiş veri seti üzerinde hiper parametre optimizasyonu yapılmış XGBoost modeliyle elde edildiği gözlemlenmiştir. Bu yöntemde, 0.96 doğruluk oranı ile en yüksek başarıya ulaşılmıştır. Bulgular, sosyal medya verisi üzerinden sahte hesap tespitinde makine öğrenmesi ve yeniden örnekleme yöntemlerinin etkinliği ortaya koymaktadır.

Anahtar Kelimeler: Sosyal medya, sahte hesap tespiti, makine öğrenmesi, aşırı örnekleme, eksik örnekleme.

Giriş

Sosyal medya platformları, günümüzün dijital dünyasında büyük bir etkiye sahip olup, kullanıcılar arasındaki etkileşimi ve bilgi paylaşımını kolaylaştırmaktadır. Ancak, bu platformlar aynı zamanda kötü niyetli kullanıcılar için fırsatlar sunmakta ve bu durum, sahte hesapların yaratılmasıyla sonuçlanmaktadır. Sahte hesaplar, spam, kimlik hırsızlığı, sahte etkileşimler ve diğer kötü amaçlı faaliyetler için kullanılabilmektedir. Bu tür hesapların varlığı, sosyal medya platformlarında kullanıcı deneyimini olumsuz yönde etkileyebilir ve güvenlik risklerine yol açabilir. Bu sebeple, sahte hesapların tespiti ve engellenmesi, platformların güvenliğini sağlamak ve kullanıcı güvenini artırmak için kritik bir gereklilik haline gelmiştir.

Sahte hesapların tespiti, sosyal medya platformları üzerinde güvenliği artırmak ve kullanıcı deneyimini iyileştirmek amacıyla önemli bir araştırma alanı olmuştur. Sahte hesaplar, genellikle spam, kimlik hırsızlığı, sahte etkileşimler veya diğer kötü niyetli faaliyetler için kullanılır. Bu nedenle, sahte hesapların doğru bir şekilde tespit edilmesi, platformlar için kritik öneme sahiptir. Sahte hesapların tespiti, genellikle iki ana kategoriye ayrılabilir; kural tabanlı yöntemler ve makine öğrenmesi tabanlı yöntemler. Kural tabanlı yöntemler, kullanıcı verilerinde belirli eşiklerin aşılması veya spesifik desenlerin varlığı gibi deterministik (belirleyici) kurallara dayanarak sahte hesapları tespit etmeye çalışır. Bu yöntemler genellikle kolay anlaşılabilir ve uygulanabilir olmalarıyla öne çıkar; ancak, yeni veya öngörülemeyen sahtecilik davranışlarını yakalamakta yetersiz kalabilirler. Öte yandan, makine öğrenmesi tabanlı yöntemler, verilerdeki karmaşık örüntüleri öğrenerek daha esnek ve olasılıksal (probabilistic) tahminler yapma yeteneğine sahiptir. Bu yönüyle, veri çeşitliliği yüksek ve sahtecilik davranışları sürekli değişen durumlarda makine öğrenmesi yöntemleri daha avantajlı olabilirken, tanımlı ve sabit kurallarla açıklanabilecek senaryolarda kural tabanlı yaklaşımlar daha hızlı ve yorumlanabilir çözümler sunabilir.

Bu çalışmanın amacı, Instagram üzerindeki sahte hesapları tespit etmek için makine öğrenmesi tekniklerinden yararlanarak bir model geliştirmektir. Kullanıcı profillerinden elde edilen veriler üzerinden, sahte hesapların tipik davranış ve özelliklerine dayalı yeni öznitelikler türetilmiş, ve sınıflar arasındaki dengesizlik sorununu çözmek için çeşitli yeniden örnekleme yöntemleri uygulanmıştır. Çalışmanın sonunda, XGBoost algoritması ve hiper parametre optimizasyonları ile geliştirilen model ile yüksek doğruluk oranı elde edilmiştir.

Bu çalışma, sahte hesap tespiti için kullanılan yöntemlerin etkinliğini tartışırken, aynı zamanda veri setindeki dengesizliklerin ve farklı makine öğrenmesi tekniklerinin model performansına etkilerini incelemektedir. Ayrıca, gelecekteki çalışmalarda bu modelin daha dinamik öğrenme teknikleri ile entegrasyonunun önerildiği bir yol haritası sunulmaktadır.

Literatür Taraması

Sosyal medya platformları, kullanıcı etkileşimi açısından büyük veri kaynakları sunarken, bu veriler arasında sahte hesapların varlığı önemli bir güvenlik sorunu teşkil etmektedir. Sahte hesaplar; manipülatif davranışlar, spam içerikler ve kötü niyetli aktiviteler yoluyla kullanıcı deneyimini olumsuz etkilemektedir. Son yıllarda, sahte hesapların otomatik olarak tanımlanmasında makine öğrenmesi yöntemleri öne çıkmaktadır. Bu yöntemler, kullanıcı profilleri ve etkileşim davranışlarından elde edilen çok boyutlu verileri analiz ederek sahteci-

lik örüntülerini keşfetmeye olanak tanır. Ancak, sahte hesaplar çoğunlukla azınlık sınıfı olarak temsil edildiğinden, bu dengesiz veri yapısı model başarımını olumsuz etkileyebilmektedir. Bu nedenle, sınıflar arası dengeyi sağlamak amacıyla SMOTE, ADASYN gibi oversampling ve ENN gibi undersampling teknikleri, sahte hesap tespitine yönelik modellerin başarımını artırmak için sıklıkla kullanılmaktadır.

Çok sayıda çalışma, sahte hesapları tespit etmek için farklı yöntemler geliştirmiştir [1]. Bu alanda öne çıkan bazı çalışmalar şu şekildedir.

Akoglu ve arkadaşları [2] sosyal medya verileri üzerinde sahte hesapları tespit etmek için kural tabanlı yöntemler önermiştir. Çalışmada kullanıcıların profillerindeki belirli özellikler (örneğin takipçi sayısı, gönderi sayısı, etkileşim oranları) dikkate alınarak, şüpheli hesaplar tespit edilmiştir. Ancak bu yöntemler genellikle sınırlı doğruluk oranlarına sahiptir ve daha karmaşık durumları tespit etmek zor olabilmektedir.

Sahte hesapların tespiti için makine öğrenmesi yöntemlerinin kullanımı son yıllarda önemli bir araştırma konusu olmuştur. Özellikle, Rastgele Orman ve Destek Vektör Makineleri (SVM) gibi sınıflandırma algoritmaları sahte hesapları tespit etmek için yaygın olarak kullanılmıştır [3-4]. Bu yöntemler ile, farklı öznitelikler (takipçi sayısı, biyografi uzunluğu, profil fotoğrafı varlığı gibi) üzerinden öğrenme yaparak 0.99'a ulaşan doğruluk oranları elde edilmiştir.

Örneğin Akyon ve arkadaşları [5] çalışmalarında Instagram'daki sahte hesapların tespiti için Naive Bayes, Lojistik Regresyon, Destek Vektör Makineleri (SVM) ve Yapay Sinir Ağları gibi makine öğrenmesi algoritmaları uygulamıştır. SVM ve Yapay Sinir Ağlarının 0.96 doğruluk ile sahte hesap tespitinde en başarılı sonuçları verdikleri gözlemişlerdir.

Sahte hesap tespiti alanındaki en önemli zorluklardan biri, veri dengesizliği sorunudur. Gerçek hesaplar genellikle sahte hesaplara kıyasla çok daha fazladır, bu da sınıflandırma modellerinin çoğunluk sınıfına odaklanmasına ve azınlık sınıfını göz ardı etmesine yol açabilir. Çeşitli çalışmalar, SMOTE (Synthetic Minority Over-sampling Technique) ve ADASYN (Adaptive Synthetic Sampling) gibi aşırı örnekleme (oversampling) teknikleri ile bu dengesizliğin üstesinden gelmeye çalışmıştır [6-7]. Bu yöntemler, azınlık sınıfındaki örnekleri artırarak modelin doğruluğunu iyileştirmeyi hedeflemektedir.

Eksik örnekleme (undersampling) yöntemlerinden ENN [8], hatalı sınıflandırılmış örnekleri çıkararak modelin daha temiz ve anlamlı bir veri seti üzerinde öğrenmesini sağlamaktadır.

Ancak, her iki yöntem de tek başına yeterli olmayabilir. Dolayısıyla, sınıflandırma algoritmalarının doğru seçimleri ve hiper parametre optimizasyonları da büyük önem taşımaktadır. Rastgele Orman ve XGBoost gibi güçlü modellerin dengesiz veri setlerinde 0.90 üzerinde başarı gösterdiği rapor edilmiştir [9]. Bu modellerin, aşırı örnekleme teknikleri ile birleştiğinde sınıflandırma doğruluğunu önemli ölçüde artırabildiği gözlemlenmiştir.

Örneğin Mohammed ve arkadaşları [10] dengesiz veri kümesi üzerinde aşırı örnekleme ve eksik örnekleme teknikleri uygulayarak farklı makine öğrenmesi sınıflandırma modellerini karşılaştırmıştır.

Aşırı örnekleme tekniğinin farklı sınıflandırma modelleri üzerinde eksik örnekleme tekniğine göre daha başarılı sonuç verdiği gözlemlenmiştir. Denenen sınıflandırma modellerinden

(Destek Vektör Makineleri (SVM), Naive Bayes (NB), Karar Ağacı (DT), Rastgele Orman (RF), vb.) aşırı örnekleme yapılmış RF modeli 0.99 doğruluk oranı ile en başarılı sonucu vermiştir.

Brandt ve arkadaşları [11] iki aşırı örnekleme tekniği olan SMOTE ve ADASYN'in performanslarını karşılaştırmıştır. Üç farklı dengesiz veri kümesi üzerinde (Banka Müşterileri İçin Müşteri Kaybını Tahmin Etme, Ev Kredisi Temerrüt Riski ve Kredi Kartı Dolandırıcılığı Tespiti), üç farklı sınıflandırma modeli (Destek Vektör Makineleri (SVM), Logistic Regresyon (LR) ve Rastgele Orman (RF)) kullanılan çalışmada çoğu durumda hem SMOTE hem de ADASYN'in sınıflandırıcıların performansını artırdığı gözlenmiştir.

Yapılan literatür taraması, sosyal medya üzerindeki sahte hesap tespitinin, özellikle makine öğrenmesi ve derin öğrenme yöntemleri ile daha etkili hale geldiğini göstermektedir [12-14]. Ancak, bu alanda hâlâ bazı zorluklar bulunmaktadır. Özellikle veri dengesizliği, platformların API kısıtlamaları ve değişken etkileşim oranları gibi faktörler, doğru ve güvenilir sonuçlar elde etmeyi zorlaştırabilir. Bu bağlamda, sosyal medya verileri üzerinde SMOTE, ADASYN gibi aşırı örnekleme (oversampling) ve ENN gibi eksik örnekleme (undersampling) yöntemlerinin uygulanarak sınıflandırma performansının artırıldığı çeşitli çalışmalar literatürde yer almaktadır [15-16]. Bu tür yöntemlerin, özellikle dengesiz sınıf dağılımının söz konusu olduğu sahte hesap tespiti problemlerinde model başarımını önemli ölçüde iyileştirdiği gösterilmiştir. Bu nedenle, sahte hesap tespiti için yalnızca güçlü sınıflandırma algoritmalarının değil, aynı zamanda etkili veri ön işleme ve dengeleme tekniklerinin de birlikte kullanılması büyük önem taşımaktadır [17-18]. Gelecekte, bu yaklaşımların dinamik öğrenme teknikleri ile entegre edilmesi, sahte hesap tespitinde daha güvenilir ve sürdürülebilir çözümler sunabilecektir. Özellikle, çevrimiçi öğrenme (online learning), pekiştirmeli öğrenme (reinforcement learning) ve hibrit modellerin kullanımı, sahte hesapların evrilen yapısına karşı modellerin adaptasyon yeteneğini artırabilir [19-21].

Bu kapsamda yapılan çalışmalar incelendiğinde, sahte hesap tespiti problemine yönelik kullanılan farklı makine öğrenmesi algoritmaları, veri dengeleme teknikleri ve model iyileştirme stratejileri değerlendirilmiştir. Elde edilen bulgular doğrultusunda, mevcut yöntemlerin avantaj ve sınırlamaları göz önünde bulundurularak bu çalışma kapsamında uygulanabilir, etkili ve doğruluğu yüksek bir yaklaşımın geliştirilebileceği öngörülmektedir.

Materyal ve Metod

Bu bölümde, çalışmanın gerçekleştirilmesinde kullanılan veri toplama yöntemleri, öznitelik mühendisliği adımları, veri ön işleme süreci ve uygulanan makine öğrenmesi teknikleri detaylı olarak açıklanmıştır. İlk olarak, Instagram platformuna ait herkese açık kullanıcı verileri Apify üzerinde yer alan bir web kazıyıcı aracılığıyla toplanmıştır. Ardından, sahte ve gerçek hesapların ayırt edilebilmesi için çeşitli öznitelikler belirlenmiş ve veri seti bu doğrultuda yapılandırılmıştır. Veri setindeki dengesizlik problemiyle başa çıkmak amacıyla farklı yeniden örnekleme (resampling) yöntemleri uygulanmıştır. Son olarak, çeşitli makine öğrenmesi modelleri eğitilerek, sahte hesapları sınıflandırmadaki başarıları karşılaştırılmıştır. Bu süreç boyunca model doğruluğunu artırmak için hiper parametre optimizasyonları ve değerlendirme metrikleri de dikkate alınmıştır.

Apify Platformu ve Instagram Profil Verisi Toplama

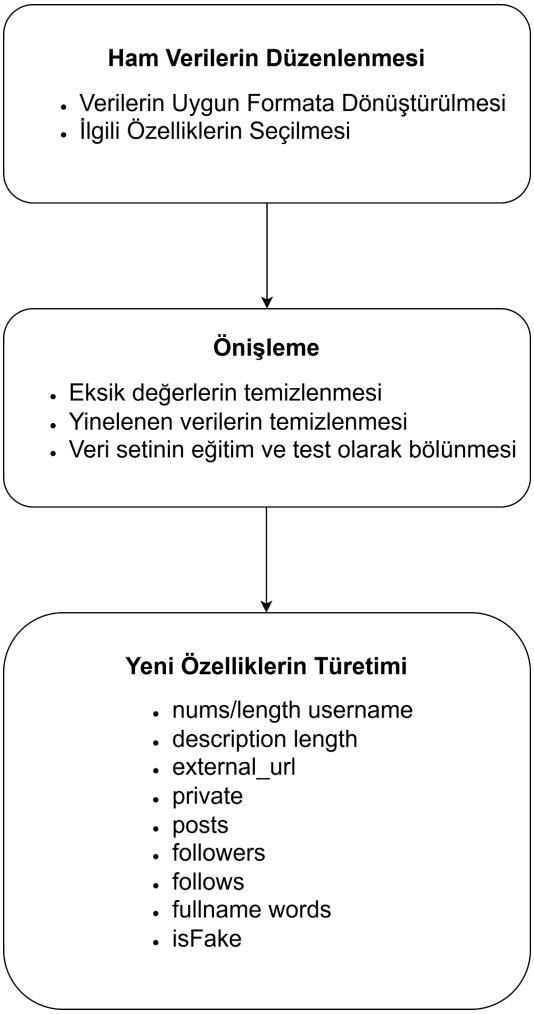
Instagram üzerindeki sahte hesapların tespitine yönelik bir model geliştirmek için geniş ve çeşitli bir veri setine ihtiyaç duyulmaktadır. Ancak, bu tür bir veri seti genel olarak hazır halde bulunmamaktadır. Özellikle sosyal medya platformlarından veri toplamak, platformların API limitleri veya erişim politikaları nedeniyle zorlu olabilmektedir. Bu nedenle, gerekli verileri toplamak amacıyla uygun bir web veri kazıma aracının kullanılması gerekmektedir.

Bu çalışmada, Instagram profillerine ait herkese açık verilerin toplanması için Apify platformu tercih edilmiştir. Apify, web veri kazıma (web scraping), otomasyon ve veri işleme için kullanılan güçlü bir platformdur [22]. Apify platformu, farklı web sitelerinden veri toplamak amacıyla çeşitli önceden tanımlanmış ve özelleştirilebilir aktörler sunar. Aktörler, belirli bir web sitesine veya veri türüne yönelik olarak yapılandırılmış otomasyon araçlarıdır. Kullanıcılar, kendi aktörlerini oluşturabilecekleri gibi, Apify kütüphanesinde bulunan hazır aktörleri de kullanabilirler. Bu aktörler sayesinde, karmaşık veri toplama süreçleri kolaylıkla yönetilebilir ve hızlandırılabilir.

Instagram verilerini toplamak amacıyla Apify üzerinde yer alan farklı aktörler bulunmaktadır. Bunlar, Instagram profil bilgilerini, gönderilerini veya takipçi listelerini çekmeye yönelik çeşitli işlevler sunar. Bu çalışmada, Instagram profillerine ait herkese açık bilgileri otomatik olarak toplayabilen Instagram Profile Scraper aktörü tercih edilmiştir. Söz konusu aktör; kullanıcı adı, biyografi, takipçi sayısı, takip edilen kişi sayısı, gönderi sayısı ve profil fotoğrafı gibi birçok temel bilgiyi hızlı ve güvenilir bir şekilde elde edebilmektedir. Bu sayede el ile müdahaleye gerek kalmadan kapsamlı bir veri seti oluşturulmuştur. Çalışma kapsamında toplam 1654 Instagram hesabından veri toplanmıştır. Bu hesapların 1255'i gerçek kullanıcı hesapları olarak etiketlenmişken, 399'u sahte hesap olarak etiketlenmiştir. Elde edilen veriler, model geliştirme ve analiz süreçlerine geçmeden önce ön işlemeye tabi tutulmuştur.

Önişleme ve Öznitelik Mühendisliği

Toplanan veriler, analiz için uygun hale getirilmeden önce bir dizi önişleme sürecinden geçirilmiştir. Bu aşama, verilerin temizlenmesi, eksik veya hatalı verilerin düzeltilmesi ve yeni özniteliklerin türetilmesini içerir. Şekil 1'de bu sürecin detaylı açıklaması yer almaktadır.



Şekil 1: Veri işlemleri

Ham Verinin Düzenlenmesi

Toplanan veriler, Apify platformu üzerinden Instagram profil bilgileri şeklinde elde edilmiştir. Ham veri seti, çeşitli alanlar içermekte olup, her bir profilin username (kullanıcı adı), biography (biyografi), followersCount (takipçi sayısı), followsCount (takip edilen kişi sayısı), postsCount (gönderi sayısı), fullName (tam ad) gibi bilgileri barındırmaktadır. Bu veriler, öncelikle analiz ve modelleme için uygun formatta düzenlenmiştir.

Yeni Özniteliklerin Türetimi

Veri setindeki mevcut özniteliklerden, kullanıcı hesabının sahte olup olmadığını tahmin edebilmek için çeşitli yeni öznitelikler türetilmiştir. Bu türetilen öznitelikler, sahte hesapların tipik özelliklerinden yola çıkılarak, kullanıcı adı, biyografi, profil durumu gibi alanlara dayanmaktadır. Aşağıda bu özniteliklerin nasıl türetildiği açıklanmıştır:

• **nums/length username:** Kullanıcı adındaki sayı karakterlerinin, toplam kullanıcı adı uzunluğuna oranı hesaplanmıştır. Sahte hesapların kullanıcı adlarında genellikle daha fazla sayı içerdiği gözlemiyle bu öznitelik türetilmiştir. Bu oran, sahte hesapların daha yüksek bir değere sahip olabileceği varsayımıyla eklenmiştir.

• **description length:** Profil biyografisinin uzunluğu hesaplanmıştır. Gerçek hesapların biyografileri genellikle daha anlamlı ve uzun olurken, sahte hesaplar genellikle kısa biyografiler kullanmaktadır.

• **external url:** Profilde harici bir bağlantının olup olmadığı kontrol edilmiştir. Eğer profilin biyografisinde bir URL varsa, bu öznitelik 1 olarak işaretlenmiş, yoksa 0 olarak belirlenmiştir.

• **private:** Hesabın gizli olup olmadığı kontrol edilmiştir. Gizli hesaplar genellikle daha az etkileşim alırken, açık hesaplar daha fazla görünürlük ve etkileşimde bulunur. Bu öznitelik, sahte hesapların gizlilik durumu ile ilişkilendirilmiştir.

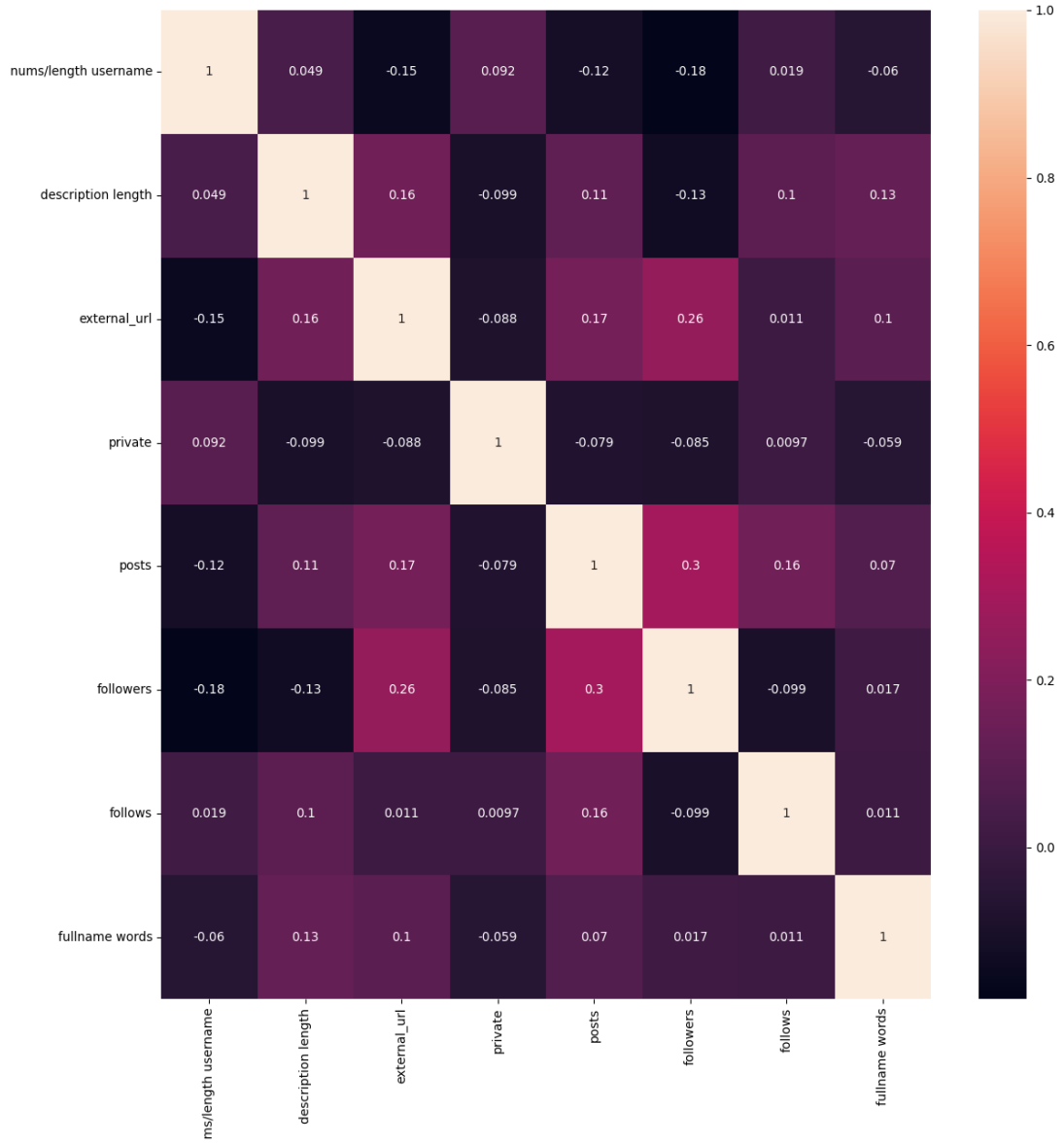
• **posts:** Kullanıcı hesabının gönderi sayısı alınmıştır. Sahte hesaplar, genellikle çok az sayıda gönderiye sahiptir. Bu nedenle gönderi sayısı, sahte hesapları tespit etmede etkili bir öznitelik olarak kullanılmıştır.

• **followers:** Takipçi sayısı, sahte hesapların düşük takipçi sayısına sahip olma eğiliminde olması nedeniyle dikkate alınan bir diğer önemli öznitelik. Gerçek hesaplar genellikle daha fazla takipçiye sahip olur.

• **follows:** Takip edilen kişi sayısı, sahte hesapların genellikle çok sayıda kişiyi takip etme eğiliminde olduğu gözlemiyle türetilmiştir.

• **fullname words:** Kullanıcı adının kaç kelimeden oluştuğu hesaplanmıştır. Gerçek hesaplarda genellikle bir veya iki kelimeden oluşan kullanıcı isimleri görülürken, sahte hesaplar daha rastgele ve uzun kullanıcı isimleri kullanabilmektedir.

Şekil 2'de verilen korelasyon matrisi, sahte ve gerçek Instagram hesaplarını ayırt etmede kullanılan öznitelikler arasındaki ilişkilerin genel olarak çok yüksek düzeyde olmadığını göstermektedir. Öznitelikler arasında güçlü doğrusal bir ilişki bulunmaması, her bir özelliğin modele farklı açılardan katkı sağladığını göstermesi açısından olumlu bir durumdur. En dikkat çekici korelasyon, gönderi sayısı (posts) ile takipçi sayısı (followers) arasında olup 0.30 düzeyindedir; bu durum, daha fazla gönderi paylaşan hesapların genellikle daha fazla takipçiye sahip olabileceğini düşündürmektedir. Benzer şekilde, harici linke sahip olma (external_url) ile takipçi sayısı (followers) (0.26) arasındaki pozitif ilişki, harici bağlantı paylaşan hesapların da daha fazla takipçiye sahip olma eğiliminde olduğunu göstermektedir. Diğer taraftan, kullanıcı adındaki numara sayısının kullanıcı adının uzunluğuna oranı (nums/length username), hesap adının kaç kelimeden oluştuğu (fullname words), hesabın gizli olup olmaması (private) gibi bazı özniteliklerin diğerleriyle olan korelasyonlarının düşük olması, bu özniteliklerin birbirinden bağımsız ve modele katkı sağlayabilecek ayırt edici nitelikte olduğunu ortaya koymaktadır. Bu durum, öznitelik setinin çeşitliliğini desteklemekte ve özellikle sahte hesapları tanımlamada karmaşık örüntüleri yakalayabilecek modeller için avantaj sağlamaktadır.



Şekil 2: Korelasyon matrisi

Ön İşleme

Veri seti oluşturulduktan ve yeni öznitelikler türetildikten sonra, veri kalitesinin artırılması ve modelin doğruluğunun garanti altına alınabilmesi için birkaç önişleme adımı uygulanmıştır. İlk adım olarak, veri setindeki eksik değerler temizlenmiştir. Ham veriler toplandıktan sonra, bazı satırlarda eksik ya da NaN (Not a Number) değerler bulunmaktaydı. Bu eksik veriler, modelin doğru sonuçlar vermesini engelleyebilir ve yanlış sınıflandırmalar yapmasına sebep olabilirdi. Bu nedenle, eksik veri içeren satırlar veri setinden tamamen çıkarılmıştır. Bu işlem, modelin eğitimi sırasında sadece geçerli ve tamamlanmış verilere dayanarak analiz yapılmasını sağlamak amacıyla gerçekleştirilmiştir.

Bunun yanı sıra, veri setinde yinelenen kayıtlar da tespit edilmiştir. Çoğu zaman, verilerin toplama sürecinde aynı profillerin birden fazla kez kaydedilmesi söz konusu olabilir. Bu yinelenen veriler, modelin öğrenme sürecini olumsuz etkileyebilir, çünkü aynı bilgilerin tekrarı

modelin eğitiminde yanlılık oluşturabilir. Bu sebeple, veri setinde bulunan tüm yinelenen kayıtlar kaldırılmıştır. Bu adım, her bir verinin eşsiz olmasını sağlayarak, modelin her örneğe ayrı ayrı odaklanmasını mümkün kılmıştır.

Veri setinin temizlenmesi, modelleme sürecine başlamadan önce gerçekleştirilmiş olup, eksik değerler ve yinelenen kayıtlar tamamen veri setinden çıkarılmıştır. Bu adım, modelin eğitimine uygun, tutarlı ve güvenilir bir veri zemini oluşturmak amacıyla yapılmıştır. Böylece verideki olası gürültülerden arındırılmış, algoritmaların daha sağlıklı öğrenmesini sağlayacak bir yapı elde edilmiştir. Veri temizliği sonucunda, sahte hesapları tespit etmeye yönelik geliştirilen modellerin doğruluğunu artıracak güçlü bir temel hazırlanmış ve analiz sürecinin güvenilirliği sağlanmıştır. Tablo-1'de veri setine ait örnek verilmiştir.

Tablo 1: Veri seti örneği

Öznitelik	Örnek-1	Örnek-2	Örnek-3	Örnek-4	Örnek-5
nums/length username	0.0	0.0	0.0	0.0	0.14
description length	49	85	62	82	137
external url	1	1	0	0	0
private	0	0	0	0	0
posts	225	105	189	796	49
followers	533,922	8,122	44,941	151,257	1,263
follows	499	591	62	991	497
fullname words	1	2	1	2	1
is fake	0	1	1	0	1

Model Eğitimleri ve Yeniden Örnekleme Teknikleri

Veri seti ön işleme aşamalarının tamamlanmasının ardından, model eğitimi sürecine geçilmiştir. Bu aşamada, verinin eğitimi ve değerlendirilmesi amacıyla %70 eğitim (train) ve %30 test olmak üzere ikiye bölünmesi sağlanmıştır. Eğitim verisi, modelin öğrenmesini sağlamak için kullanılırken, test verisi modelin genel performansını değerlendirmek için kullanılmıştır.

Veri setlerinde dengesizlik problemi, sınıflar arasında önemli sayısal farkların bulunmasından kaynaklanır ve bu durum makine öğrenmesi modellerinin performansını ciddi şekilde olumsuz etkileyebilir. Dengesiz veri setleri, modelin azınlık sınıfını görmezden gelmesine, yalnızca çoğunluk sınıfına odaklanmasına ve dolayısıyla genelleme gücünün düşmesine neden olabilir [23-24]. Bu nedenle, modelin her iki sınıfı da doğru bir şekilde ayırt edebilmesi için veri dengesizliğini gidermek amacıyla çeşitli yeniden örnekleme teknikleri uygulanmıştır.

Çalışmada öncelikle Rastgele Orman Sınıflandırma algoritmasıyla temel bir sınıflandırma gerçekleştirilmiş ve veri setindeki sınıf dengesizliğinin model performansını olumsuz etkilediği gözlemlenmiştir. Bu nedenle, SMOTE ve ADASYN gibi aşırı örnekleme, ENN gibi eksik örnekleme teknikleri uygulanarak veri dengelenmiş ve model başarıları karşılaştırılmıştır. Son aşamada, farklı makine öğrenmesi algoritmaları denenmiş ve hiper parametre optimizasyonları gerçekleştirilmiştir.

Deneysel Sonuçlar

Başlangıçta, veri setine herhangi bir dengeleme işlemi uygulanmadan geliştirilen modelde, Rastgele Orman Sınıflandırma algoritması kullanılarak 0.77 doğruluk değeri elde edilmiştir. Ancak bu doğruluk değeri, veri setindeki belirgin sınıf dengesizliği nedeniyle yanıltıcı olabilmektedir; zira 1.255 gerçek hesap (sınıf 0) ve yalnızca 399 sahte hesap (sınıf 1) bulunmaktadır. Kesinlik, duyarlılık ve F1-skoru gibi sınıf bazlı metrikler incelendiğinde, sahte hesapları temsil eden sınıfın duyarlılık değeri yalnızca 0.37 olarak ölçülmüş; buna karşılık, gerçek hesap sınıfı 0.91 gibi oldukça yüksek bir duyarlılık değeri göstermiştir. Bu durum, modelin sahte hesapları tespit etmekte yetersiz kaldığını, yani çok sayıda sahte hesabı gerçek olarak sınıflandırdığını ortaya koymaktadır. Bu hatalı sınıflandırmaların temel nedeni, yüksek sayıda gerçek hesap bulunması nedeniyle modelin tahminlerinin gerçek sınıfa doğru eğilim göstermesidir. Özellikle dengesiz veri setlerinde bu tür bir yanlılık, azınlık sınıfının (bu çalışmada sahte hesapların) doğru şekilde tanınmasını zorlaştırmakta ve modelin genel başarısını sınırlamaktadır. Tablo-2'de geliştirilen modele ait detaylı sınıflandırma raporu verilmiştir.

Tablo-2: Rastgele Orman Sınıflandırma modeli sonuçları

Sınıf	Kesinlik	Duyarlılık	F1	Doğruluk
0	0.80	0.91	0.85	0.91
1	0.61	0.37	0.46	0.37
Ort.	0.75	0.77	0.75	0.77

Daha sonra, veri setindeki dengesizlik problemini gidermek amacıyla SMOTE (Synthetic Minority Over-sampling Technique) yöntemi uygulanarak azınlık sınıfı olan sahte hesaplar üzerinde sentetik örneklerle aşırı örnekleme yapılmıştır. Bu işlem sonucunda, veri seti dengelenmiş ve toplamda 2.510 hesaba ulaşılmıştır; bu hesapların 1.255'i sahte, 1.255'i ise gerçek hesaptır. SMOTE uygulaması sonrasında eğitilen modelin doğruluk değeri 0.83'e yükselmiş ve özellikle sınıf 1 (sahte hesaplar) için duyarlılık (recall) değeri 0.84'e çıkarak, modelin sahte hesapları tanıma yeteneğinde önemli bir iyileşme sağlanmıştır. Bu dengeleme tekniği, azınlık sınıfına ait örneklerin daha iyi öğrenilmesini mümkün kılarak, modelin genel sınıflandırma performansını artırmıştır. Tablo-3'te geliştirilen modele ait detaylı sınıflandırma raporu verilmiştir.

Tablo 3: SMOTE + Rastgele Orman Sınıflandırma modeli sonuçları

Sınıf	Kesinlik	Duyarlılık	F1	Doğruluk
0	0.82	0.82	0.82	0.82
1	0.84	0.84	0.84	0.84
Ort.	0.83	0.83	0.83	0.83

ADASYN (Adaptive Synthetic Sampling) yöntemiyle gerçekleştirilen aşırı örnekleme sürecinde de sınıflar arası dengesizlik başarılı bir şekilde giderilmiş ve toplam veri sayısı 2.490'a ulaşmıştır; bunların 1.255'i gerçek, 1.235'i ise sahte hesaptır. ADASYN, özellikle azınlık sınıfın daha karmaşık bölgelerine odaklanarak sentetik örnekler üreten bir yöntem olduğu için, sahte hesapların daha etkili öğrenilmesine katkı sağlamıştır. Bu yöntemle eğitilen modelin doğruluk değeri 0.81 olarak ölçülmüş, bu da SMOTE ile elde edilen başarıya oldukça yakın bir sonuç sunmuştur. Her ne kadar ADASYN'in doğruluk oranı SMOTE'e göre biraz

daha düşük olsa da, sahte hesapların tanınmasındaki gelişme dikkate değerdir ve azınlık sınıfının model tarafından daha iyi öğrendiğini göstermektedir. Tablo-4'te geliştirilen modele ait detaylı sınıflandırma raporu verilmiştir.

Tablo 4: ADASYN + Rastgele Orman Sınıflandırma modeli sonuçları

Sınıf	Kesinlik	Duyarlılık	F1	Doğruluk
0	0.81	0.79	0.80	0.79
1	0.81	0.82	0.82	0.82
Ort.	0.81	0.81	0.81	0.81

Son aşamada, veri setinin dengesini sağlamak amacıyla ENN (Edited Nearest Neighbors) eksik örnekleme (undersampling) yöntemi uygulanmıştır. ENN, her bir örneği en yakın komşularıyla karşılaştırarak, sınıf yapısıyla uyuşmayan örnekleri veri setinden çıkartır. Başlangıçta 1.654 hesap örneğinden oluşan veri setinde 1.255 gerçek (sınıf 0) ve 399 sahte (sınıf 1) hesap bulunurken, ENN uygulaması sonrasında bu sayı 806'ya düşmüş; 407 gerçek ve 399 sahte hesap kalmıştır. Bu işlem sayesinde sınıflar dengelenmiş ve modelin her iki sınıfı da daha adil şekilde öğrenmesi sağlanmıştır. ENN sonrası eğitilen modelin doğruluk değeri 0.94'e ulaşmış, hem sahte hem de gerçek hesaplar için yüksek kesinlik, duyarlılık ve F1-skoru değerleri elde edilmiştir. ENN'nin bu denli başarılı sonuçlar vermesinin temel nedeni, veri setindeki gürültülü ve sınıflar arası sınırda yer alan örnekleri ortadan kaldırarak daha net bir karar yüzeyi oluşturmasıdır. Bu durum, modelin aşırı öğrenmesini (overfitting) engellemiş ve genelleme yeteneğini artırmıştır. Öte yandan, SMOTE ve ADASYN gibi aşırı örnekleme yöntemleri sınıf dengesini sağlasa da, gürültülü veya örtüşen bölgelerde oluşturulan sentetik örnekler karar sınırlarını belirsizleştirerek modelin daha karmaşık ve hata yapmaya açık yapılar geliştirmesine yol açabilir. Bu bağlamda, ENN'nin sadeleştirilmiş ve gürültüden arındırılmış veri kümesiyle daha başarılı bir performans sergilemesi beklenen bir sonuç olmuştur. Tablo-5'te geliştirilen modele ait detaylı sınıflandırma raporu verilmiştir.

Tablo 5: ENN + Rastgele Orman Sınıflandırma modeli sonuçları

Sınıf	Kesinlik	Duyarlılık	F1	Doğruluk
0	0.95	0.95	0.95	0.95
1	0.94	0.94	0.94	0.94
Ort.	0.94	0.94	0.94	0.94

Tüm bu sonuçlar, veri setindeki dengesizliklerin modelin performansını olumsuz etkilediğini ve yeniden örnekleme tekniklerinin başarıyı artırmada büyük rol oynadığını göstermektedir. Özellikle ENN yöntemi, tüm teknikler arasında en yüksek doğruluğu elde ederek modelin genellenebilirliğini artırmıştır. Ancak, model performansını daha da iyileştirmek amacıyla farklı algoritmaların denenmesi gerektiği düşünülmüştür. Bu kapsamda, ENN yöntemiyle dengelenmiş veri seti üzerinde farklı modeller eğitilmiş ve hiper parametre optimizasyonları gerçekleştirilmiştir. Yapılan denemeler sonucunda, XGBoost modeli, güçlü öğrenme yetenekleri ve esnekliği sayesinde doğruluk değeri 0.96'ya kadar yükseltmiştir. Bu sonuç, eksik örnekleme tekniklerinin dengesiz veri setlerinde etkili bir çözüm sunduğunu ve doğru algoritma seçimleriyle birlikte performansın önemli ölçüde artırılabileceğini ortaya koymaktadır. Farklı modellere ait sınıflandırma sonucu Tablo-6'da verilmiştir.

Tablo 6: ENN + diğer model sonuçları

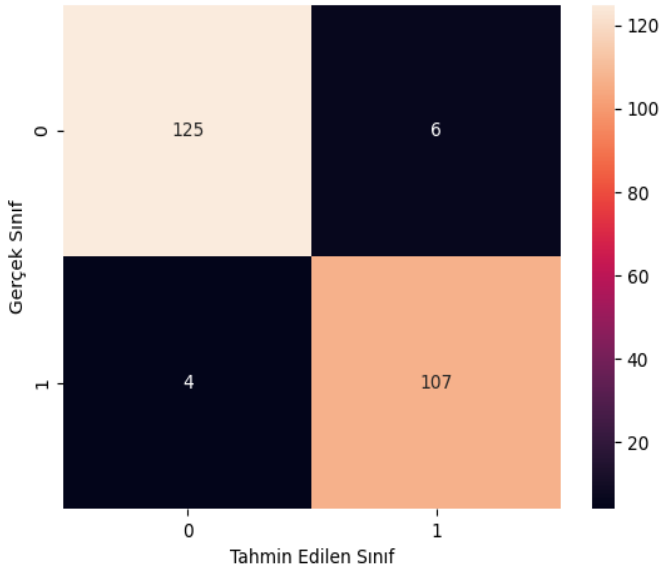
Model	Kesinlik	Duyarlılık	F1	Doğruluk
Random Forest	0.94	0.94	0.94	0.94
Decision Tree	0.95	0.95	0.95	0.95
AdaBoost	0.93	0.93	0.93	0.92
Gradient Boosting	0.96	0.96	0.96	0.96
CatBoost	0.95	0.95	0.95	0.95
XGBoost	0.96	0.96	0.96	0.96

Sonuç olarak, elde edilen en yüksek doğruluk oranı, ENN ile gerçekleştirilmiş olan yeniden örnekleme işleminden sonra XGBoost modelinin geliştirilmesi ile elde edilmiştir ve modelin doğruluk değerini 0.96'ya çıkarmıştır. Tablo-7'de bu modelin parametrelerine ilişkin detaylar verilmiştir.

Tablo 7: XGBoost hiper parametreleri

Parametre	Değer
learning rate	0.1
max_depth	3
n_estimators	200
tree_method	"exact"

Bu başarı, yeniden örnekleme tekniklerinin doğru şekilde kullanıldığında, dengesiz veri setlerinde sınıflandırma başarısını önemli ölçüde artırabileceğini göstermektedir. Modelin başarısını daha ayrıntılı bir şekilde değerlendirebilmek için, doğruluk oranı dışında, karışıklık matrisi (confusion matrix) kullanılmıştır. Karışıklık matrisi, modelin sınıflandırma performansını detaylı bir şekilde görselleştiren bir araçtır. Bu matris, doğru ve yanlış sınıflandırılmış örneklerin sayısını, her bir sınıf için doğru ve yanlış tahminleri gösterir. Gerçek Pozitif (True Positive), Yanlış Pozitif (False Positive), Gerçek Negatif (True Negatif) ve Yanlış Negatif (False Negative) değerleri üzerinden, modelin hangi sınıfları doğru tahmin ettiğini ve hangi sınıflarda hata yaptığını net bir şekilde görmek mümkündür. En iyi performans gösteren model olan XGBoost modeline ait karışıklık matrisi Şekil 3'te verilmiştir.



Şekil 3: En iyi sonuçları veren XGBoost+ENN modelinin karmaşıklık matrisi

Geliştirilen XGBoost modeli, 125 gerçek hesabı ve 107 sahte hesabı doğru bir şekilde sınıflandırırken, yalnızca 6 gerçek hesabı yanlışlıkla sahte olarak ve 4 sahte hesabı gerçek olarak etiketlemiştir. Bu sonuçlar, modelin sahte hesapları tespit etme yeteneğinin (hassasiyet) ve doğru bir şekilde tanımlama oranının (duyarlılık) yüksek olduğunu ortaya koymaktadır. Yanlış sınıflandırmaların büyük çoğunluğunun gerçek hesapların sahte olarak etiketlenmesinden kaynaklandığı görülmekte olup, bu durum modelin hatalı pozitifleri önlemekte sınırlı bir performansa sahip olduğunu göstermektedir. Ancak, genel doğruluk oranı ve başarılı sınıflandırma sayıları dikkate alındığında, ENN yöntemiyle dengelenmiş veri seti üzerinde XGBoost algoritmasının etkili bir şekilde çalıştığı ve sınıflandırma başarısını önemli ölçüde artırdığı sonucuna ulaşılabilir.

Sonuçlar

Bu çalışmada, Instagram'daki sahte hesapların tespitine yönelik bir makine öğrenmesi modeli geliştirilmiştir. Herkese açık kullanıcı profillerinden elde edilen veriler üzerinden, sahte hesapların tipik davranış ve özelliklerine dayalı yeni öznitelikler türetilmiş; ardından sınıflar arasındaki dengesizlik sorununu gidermek için çeşitli yeniden örnekleme yöntemleri uygulanmıştır. Özellikle ENN (Edited Nearest Neighbours) yöntemi ile dengelenmiş veri seti üzerinde yapılan analizlerde, modelin doğruluğu önemli ölçüde artmıştır. Son aşamada XGBoost algoritması ve hiper parametre optimizasyonları ile doğruluk değeri 0.96'ya ulaşmıştır. Elde edilen sonuçlar, hem kullanılan özniteliklerin anlamlılığını hem de veri ön işleme ve dengeleme adımlarının sınıflandırma başarısına etkisini ortaya koymakta; sahte hesap tespiti konusunda önerilen yaklaşımın etkili bir çözüm sunduğunu göstermektedir.

Gelecek çalışmalara yönelik olarak, bu çalışmanın statik bir veri seti üzerinde gerçekleştirilmiş olması bir sınırlılık olarak değerlendirilmektedir. Elde edilen veri, belirli bir dönemde toplanmış sabit bir kümeden oluştuğu için modelin, sahte hesapları belirlemedeki genel başarısını yalnızca bu veri seti üzerinden değerlendirmek mümkündür. Ancak, sahte hesap

üreten botlar algoritmalarını sürekli olarak güncellediklerinden, modelin bu dinamik yapıya karşı duyarsız kalma riski bulunmaktadır. Bu durum, zaman içerisinde modelin başarımının düşmesine neden olabilir. Bu nedenle, modelin güncel kalabilmesi ve adaptasyon sağlayabilmesi için pekiştirmeli öğrenme (reinforcement learning) [25] gibi sürekli öğrenme yaklaşımlarının uygulanması dikkate değer bir geliştirme alanı olarak öne çıkmaktadır. Özellikle çevrimdışı öğrenme yerine, çevrimiçi öğrenme [26] tekniklerinin uygulanması ile modelin yeni ortaya çıkan sahte hesap türlerine karşı kendini sürekli güncellemesi sağlanabilir. Bu yaklaşım, modelin doğruluğunu uzun vadede sürdürülebilir hale getirebilir ve gerçek zamanlı veri işleme yeteneklerini geliştirebilir. Teorik olarak, bu sorunun üstesinden gelmek için çevrimiçi öğrenme ve pekiştirmeli öğrenme gibi sürekli güncellenebilen modellerin kullanılması önerilmektedir. Ancak pratikte, bu yaklaşım çeşitli zorluklar barındırmaktadır. Özellikle Meta'nın veri erişim politikalarında yaptığı kısıtlamalar nedeniyle gerçek zamanlı ve sürekli veri çekimi oldukça zorlaşmıştır. API üzerinden yapılan veri toplama işlemleri hem hız hem de kapsam açısından sınırlı olduğundan, sahte hesaplara yönelik güncel ve büyük hacimli veri setleri oluşturmak mümkün olamamaktadır. Bu sebeple literatürde bu alanda yapılan çalışmaların büyük bir kısmı hâlâ statik ve toplu (batch) veriler üzerinden yürütülmektedir. Modelin dinamik sahte hesap stratejilerine uyum sağlayabilmesi için zaman içinde veri setinin ve modelin güncellenmesi kritik olsa da, mevcut teknik kısıtlar bu sürecin otomatik ve sürekli hale gelmesini zorlaştırmaktadır.

Diğer yandan, bu çalışmada makine öğrenmesi algoritmalarının tercih edilmesinin temel nedeni, eldeki verinin derin öğrenme yöntemlerini eğitmek için yeterince büyük olmamasıdır. Derin öğrenme yöntemleri genellikle büyük miktarda veriye ihtiyaç duyar ve mevcut veri setinin sınırlı büyüklüğü bu tür modellerin potansiyelini tam olarak ortaya koymayı zorlaştırmaktadır [27]. Bu nedenle, aşırı örnekleme (SMOTE, ADASYN) ve eksik örnekleme (ENN) teknikleri kullanılarak veri seti dengelenmiş ve makine öğrenmesi algoritmalarının başarımı artırılmıştır. Ancak, makine öğrenmesi yöntemleriyle elde edilen sonuçların ötesine geçebilmek adına daha büyük ve çeşitli bir veri setine ihtiyaç duyulmaktadır. Bu bağlamda, sentetik veri üretilmesi, veri miktarını ve çeşitliliğini artırmak için önemli bir çözüm sunabilir [28]. Bu yaklaşım, verinin kapsamını genişletebilir ve daha karmaşık derin öğrenme modellerinin eğitilmesine olanak tanıyabilir.

Son olarak, veri setinin çeşitliliği artırıldığında transfer öğrenimi (transfer learning) [29-30] gibi tekniklerin uygulanması da dikkate alınabilir. Önceden eğitilmiş modellerin, bu özel problem üzerine yeniden eğitilmesiyle başarı oranlarının artırılması hedeflenebilir. Tüm bu yaklaşımlar, sahte hesap tespitinin dinamik yapısını yakalayarak daha dayanıklı ve etkili modeller geliştirilmesini mümkün kılabilir.

Çıkar Beyanı**:** Yazarlar, herhangi bir çıkar çatışması olmadığını beyan etmektedir.

Teşekkürler: Bu çalışma Küçük ve Orta Ölçekli İşletmeleri Geliştirme ve Destekleme İdaresi Başkanlığı (KOSGEB) tarafından Ar-Ge, Ür-Ge ve İnovasyon projesi kapsamında 1105433 numaralı proje ile desteklenmiştir.

Yazar Katkıları: Bu çalışmada Mehmet Varan, çalışmanın tasarımının ve yöntemlerinin planlanması, veri toplama ve işleme süreçlerinin yürütülmesi, yazım sürecine katkı sağlanması ve araştırma fikrinin geliştirilmesinde aktif rol oynamıştır. Aslı Yatkınoğlu, literatür ta-

ramasının gerçekleştirilmesi, elde edilen bulguların analizi ve yorumlanması, bildirinin yazım sürecinin yürütülmesi ile dil, içerik ve bütünlük açısından eleştirel incelemesinden sorumlu olmuştur. Fatih Soygazi, proje sürecinin akademik denetimi ve yönlendirilmesi, analizlerin değerlendirilmesi ve yorumlanması, bildirinin teslim öncesi içerik ve bütünlük açısından gözden geçirilmesi ile araştırma fikrinin oluşturulmasına katkı sağlamıştır. Bora Mocan ise proje organizasyonu ve yürütülmesinin gÖZETimi, gerekli kaynakların (veri, araç-gereç vb.) sağlanması, çalışma kapsamında kullanılan malzemelerin temin edilmesi ve araştırma fikrinin oluşturulmasına yönelik katkılar sunmuştur. Tüm yazarlar, çalışmanın her aşamasında aktif rol almıştır.

REFERANSLAR

[1] Aljabri, M., Zagrouba, R., Shaahid, A., Alnasser, F., Saleh, A. and Alomari, D.M., 2023. Machine learning-based social media bot detection: a comprehensive literature review. Social Network Analysis and Mining, 13(1), p.20.

[2] Akoglu, L., Tong, H. and Koutra, D., 2015. Graph based anomaly detection and description: a survey. Data mining and knowledge discovery, 29, pp.626-688.

[3] Ali, A.K. and Abdullah, A.M., 2022. Fake accounts detection on social media using stack ensemble system. International Journal of Electrical and Computer Engineering (IJECE), 12(3), pp.3013-3022.

[4] Soygazi, F., Aydoğan, M.E., Taşkıran, H.C. and Kaya, Ö., 2021. Matching potential customers and influencers for social media marketing. Journal of Artificial Intelligence and Data Science, 1(2), pp.150-159.

[5] Akyon, F.C. and Kalfaoglu, M.E., 2019, October. Instagram fake and automated account detection. In 2019 Innovations in intelligent systems and applications conference (ASYU) (pp. 1-7). IEEE.

[6] Chawla, N.V., Bowyer, K.W., Hall, L.O. and Kegelmeyer, W.P., 2002. SMOTE: synthetic minority over-sampling technique. Journal of artificial intelligence research, 16, pp.321-357.

[7] He, H., Bai, Y., Garcia, E.A. and Li, S., 2008, June. ADASYN: Adaptive synthetic sampling approach for imbalanced learning. In 2008 IEEE international joint conference on neural networks (IEEE world congress on computational intelligence) (pp. 1322-1328). Ieee.

[8] Wilson, D.L., 1972. Asymptotic properties of nearest neighbor rules using edited data. IEEE Transactions on Systems, Man, and Cybernetics, (3), pp.408-421.

[9] Wang, C., Deng, C. and Wang, S., 2020. Imbalance-XGBoost: leveraging weighted and focal losses for binary label-imbalanced classification with XGBoost. Pattern recognition letters, 136, pp.190-197.

[10] Mohammed, R., Rawashdeh, J. and Abdullah, M., 2020, April. Machine learning with oversampling and undersampling techniques: overview study and experimental results. In 2020 11th international conference on information and communication systems (ICICS) (pp. 243-248). IEEE.

[11] Brandt, J. and Lanzén, E., 2021. A comparative review of SMOTE and ADASYN in imbalanced data classification.

[12] Tunç, Ü., Atalar, E., Gargı, M.S. and Aydın, Z.E., 2022. Classification of fake, bot, and real accounts on instagram using machine learning. Politeknik Dergisi, 27(2), pp.479-488.

[13] Chelas, S., Routis, G. and Roussaki, I., 2024. Detection of Fake Instagram Accounts via Machine Learning Techniques. Computers, 13(11), p.296.

[14] Mujahid, M., Kına, E.R.O.L., Rustam, F., Villar, M.G., Alvarado, E.S., De La Torre Diez, I. and Ashraf, I., 2024. Data oversampling and imbalanced datasets: An investigation of performance for machine learning and feature engineering. Journal of Big Data, 11(1), p.87.

[15] Banerjee, A., Bhattacharjee, M., Ghosh, K. and Chatterjee, S., 2020. Synthetic minority oversampling in addressing imbalanced sarcasm detection in social media. Multimedia Tools and Applications, 79(47), pp.35995-36031.

[16] Shah, A., Varshney, S. and Mehrotra, M., 2024. Detection of fake profiles on online social network platforms: performance evaluation of artificial intelligence techniques. SN Computer Science, 5(5), p.489.

[17] Nguyen, T.T., Hwang, D. and Jung, J.J., 2017. Handling imbalanced classification problem: A case study on social media datasets. Journal of Intelligent & Fuzzy Systems, 32(2), pp.1437-1448.

[18] Wu, B., Liu, L., Yang, Y., Zheng, K. and Wang, X., 2020. Using improved conditional generative adversarial networks to detect social bots on Twitter. IEEE Access, 8, pp.36664-36680.

[19] Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., Santamaría, J., Fadhel, M.A., Al-Amidie, M. and Farhan, L., 2021. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. Journal of big Data, 8, pp.1-74.

[20] Akpan, E., 2022. Fake Profile Detection on social media using Generative Adversarial Networks (GANs).

[21] Siddiqui, F. and Suaib, M., 2023. Enhancing Spammer Fake Profile Detection on Social Media Platforms using Artificial Neural Networks. International Journal of Engineering and Management Research, 13(4), pp.1-6.

[22] Web scraping, data extraction and automation. Apify, 2022. from: https://apify.com/

[23] Kotsiantis, S., Kanellopoulos, D. and Pintelas, P., 2006. Handling imbalanced datasets: A review. GESTS international transactions on computer science and engineering, 30(1), pp.25-36.

[24] Mujahid, M., Kına, E.R.O.L., Rustam, F., Villar, M.G., Alvarado, E.S., De La Torre Diez, I. and Ashraf, I., 2024. Data oversampling and imbalanced datasets: An investigation of performance for machine learning and feature engineering. Journal of Big Data, 11(1), p.87.

[25] Kaelbling, L.P., Littman, M.L. and Moore, A.W., 1996. Reinforcement learning: A survey. Journal of artificial intelligence research, 4, pp.237-285.

[26] Hoi, S.C., Sahoo, D., Lu, J. and Zhao, P., 2021. Online learning: A comprehensive survey. Neurocomputing, 459, pp.249-289.

[27] Zhang, Q., Yang, L.T., Chen, Z. and Li, P., 2018. A survey on deep learning for big data. Information Fusion, 42, pp.146-157.

[28] Fonseca, J. and Bacao, F., 2023. Tabular and latent space synthetic data generation: a literature review. Journal of Big Data, 10(1), p.115.

[29] Weiss, K., Khoshgoftaar, T.M. and Wang, D., 2016. A survey of transfer learning. Journal of Big data, 3, pp.1-40.

[30] Goyal, B., Gill, N.S., Gulia, P., Prakash, O., Priyadarshini, I., Sharma, R., Obaid, A.J. and Yadav, K., 2023. Detection of fake accounts on social media using multimodal data with deep learning. IEEE Transactions on Computational Social Systems.

TIME-TO-COLLISION PREDICTION IN AUTONOMOUS VEHICLES USING MACHINE LEARNING ALGORITHMS

Elif Cesur*

Industrial Engineering, Faculty of Engineering and Natural Sciences, İstanbul Medeniyet University, İstanbul, Türkiye

Sena Cicek

Industrial Engineering, Faculty of Engineering and Natural Sciences, İstanbul Medeniyet University, İstanbul, Türkiye

**(Corresponding author) elif.karakaya@medeniyet.edu.tr*

ABSTRACT

With the increasing prevalence of autonomous vehicles, studies aiming to enhance traffic safety and promote the societal acceptance of these technologies have gained significant importance. This study aims to estimate the time-to-collision (TTC) of autonomous vehicles, enabling them to respond to environmental risks more accurately and in a timely manner. The dataset used in the study contains information about the distance between the vehicle and the detected obstacle, as well as the current speed of the vehicle. Using this data, new variables such as target speed and time-to-collision were derived. To predict TTC, five machine learning algorithms were applied: Decision Tree, Random Forest, XGBoost, Gradient Boosting Decision Tree, and Extra Trees. The performance of these algorithms was evaluated using MAE, MSE, and R^2 metrics. Subsequently, hyperparameter optimization was applied to the algorithms considered as potential solutions, and the best performance was achieved with the Extra Trees algorithm. The results demonstrate that accurate TTC predictions can significantly enhance the environmental awareness of autonomous vehicles, enabling them to foresee potential collision scenarios earlier and more reliably.

Keywords: Autonomous vehicle, machine learning, artificial ıntelligence, time-to-collision.

Introductıon

More than one billion conventional vehicles used to meet mobility needs have become one of the most important research topics in both industrial and academic fields today, due to the environmental damage they cause and the loss of life and property resulting from traffic accidents [1]. In this context, autonomous vehicle technologies have gained significant importance in minimizing the environmental damage and accident-related losses caused by conventional vehicles and in enhancing transportation safety. At this point, autonomous vehicles represent an innovative transportation intervention that can impact public health [2]. Autonomous vehicles offer many potential advantages and safety bene-

fits. These advantages include reduced costs, enhanced safety, greater mobility, increased customer satisfaction, and lower crime rates. Additionally, autonomous vehicles can reduce traffic collisions, thereby decreasing injuries and other expenses such as insurance costs associated with these accidents [3]. However, for this technology to gain societal acceptance and be used safely, autonomous vehicles must operate reliably.

In complex systems such as autonomous vehicles, achieving the desired performance depends on effectively solving numerous subproblems [4]. Among the most critical of these subproblems are the decision-making processes that ensure the reliable operation of autonomous vehicles and the escape strategies they employ in potential risk situations. It is of great importance for an autonomous vehicle to act correctly in rapidly changing and unpredictable traffic conditions without human intervention, thereby preventing material and moral damages [5]. In this context, to ensure the reliability of autonomous vehicles, this study predicts the time-to-collision of autonomous vehicles using machine learning algorithms applied to data collected from sensors.

As a result, in this study, the time-to-collision of autonomous vehicles was predicted, aiming to enable autonomous vehicles to adapt to traffic conditions, respond safely, and improve road safety. In this way, it is expected that autonomous vehicle technology will gain societal acceptance and be used reliably.

Method and Case Study

In this study, the prediction of the time-to-collision of autonomous vehicles using machine learning algorithms was aimed. The dataset used in the study consists of measurements including the current speed of the vehicle and the distance between the vehicle and the detected obstacle. It was determined that there were no missing values in the dataset. However, considering the possibility of sensor-induced measurement errors, graphs were plotted for each feature and the data were visually analyzed. As a result of these analyses, outliers suspected to carry sensor errors were identified and removed from the dataset, thereby increasing the accuracy of the dataset.

After the data preprocessing steps, two new variables were added to the dataset: target speed and time-to-collision. The target speed was calculated using the current speed and the distance between the vehicle and the obstacle, while the time-to-collision was obtained by dividing the distance by the difference between the current speed and the target speed. In addition, as a result of the correlation analysis conducted on the dataset, it was observed that there is a strong positive relationship between the current speed and the target speed, and a strong negative relationship between the target speed and the time-to-collision. Weaker relationships were observed among the other variables. This indicates that the variables of speed and target speed are more decisive factors in collision scenarios.

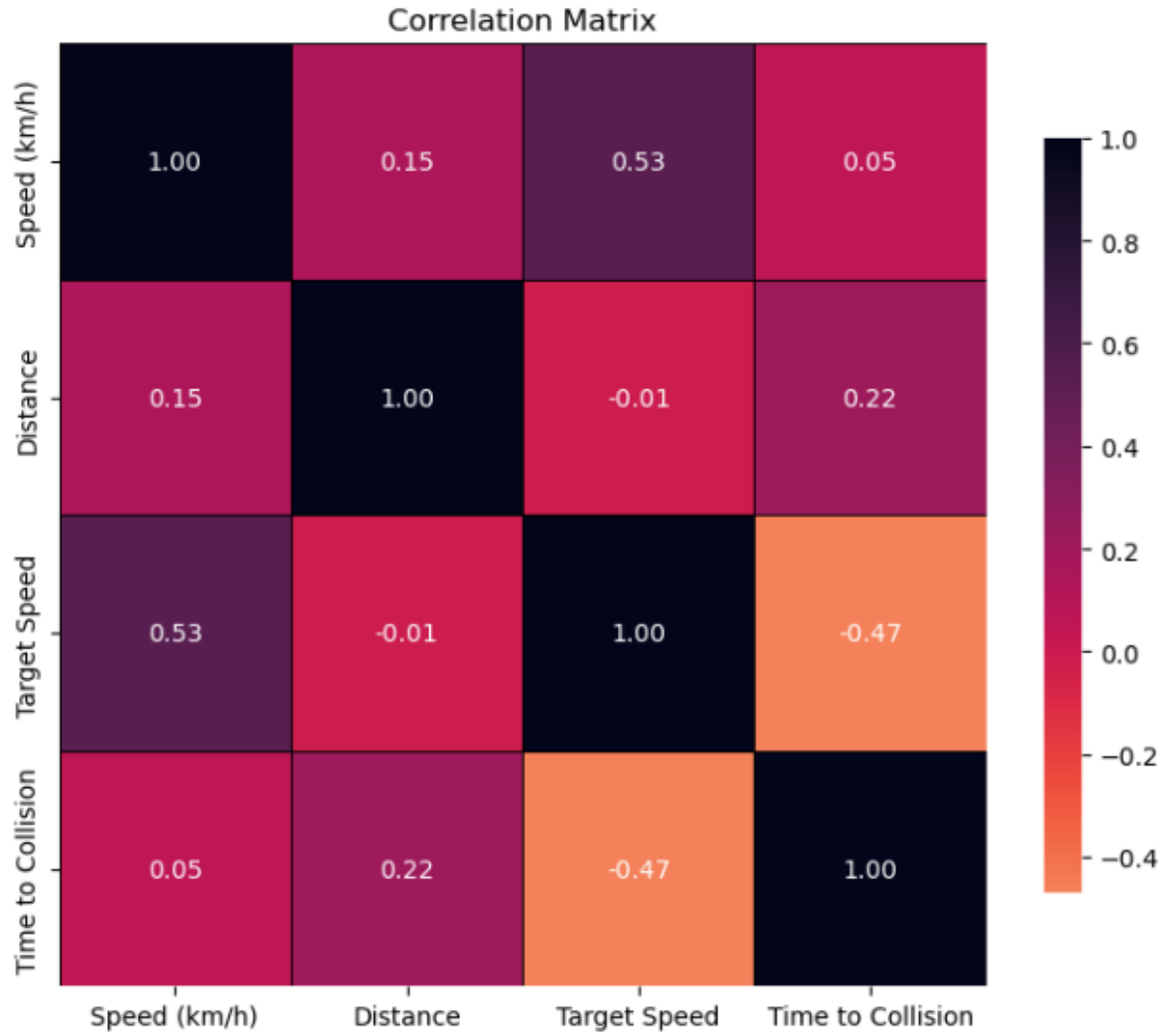


Figure 1: Heatmap of Feature Correlations

In the modeling phase, five different machine learning algorithms were used to predict the time-to-collision: Decision Tree, Random Forest, XGBoost, Gradient Boosting Decision Tree (GBDT), and Extra Trees. During the modeling process, the dataset was split into two parts, with 80% used for training and 20% for testing. The performance of the models was evaluated using the metrics of Mean Absolute Error (MAE), Mean Squared Error (MSE), and R^2 score. Based on the performance metrics obtained at the end of the modeling process, the Decision Tree, Random Forest, and XGBoost algorithms were not considered sufficient solutions for predicting the time-to-collision due to their low performances. The GBDT and Extra Trees algorithms produced significantly better results compared to the other models and explained the target variable more successfully. Considering that the high performances of these two algorithms could be further improved, parameter optimization was carried out on these models in the next stage, and the models were retrained. As a result of the parameter optimization studies, the highest accuracy rate and the lowest error values were achieved with the Extra Trees algorithm. Consequently, the Extra Trees algorithm was determined to be the most suitable model for predicting the time-to-collision in autonomous vehicles.

Discussion and Conclusion

The safe and effective operation of autonomous vehicles depends on their ability to accurately perceive environmental conditions and to intervene promptly in potential risk situations. In this study, the aim is to predict the time-to-collision (TTC) of autonomous vehicles using machine learning algorithms, and in this context, the performances of different algorithms were compared.

The results obtained indicate that machine learning algorithms can be effectively used for TTC prediction. In particular, the Extra Trees algorithm achieved the highest accuracy and the lowest error values compared to other methods, making it the most successful model. Furthermore, after the parameter optimization was carried out, the performance of the Extra Trees algorithm was further improved, and it was concluded that this model is the most suitable solution for predicting the time-to-collision. These findings will contribute to enabling autonomous vehicles to detect potential collision risks at an earlier stage and to make safer driving decisions, thereby enhancing traffic safety. In line with these results, it will also be possible to increase public trust in autonomous vehicle technologies and to promote their societal acceptance.

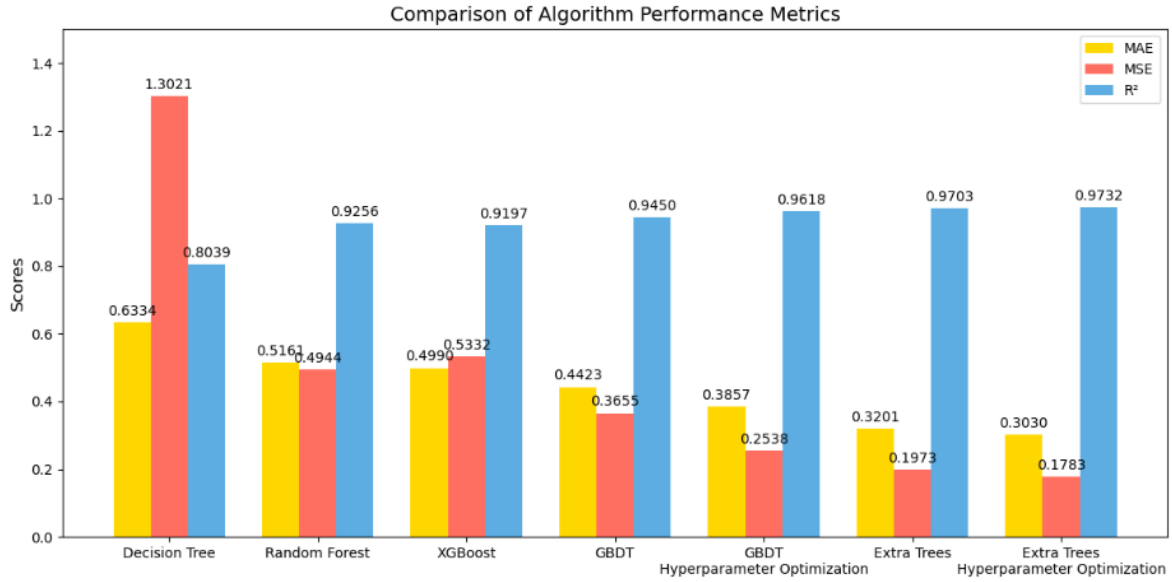


Figure 2: Performance Comparison of Machine Learning Algorithms

In conclusion, this study has demonstrated that the time-to-collision can be predicted with high accuracy to enhance the safety of autonomous vehicles. The developed model contributes to enabling autonomous vehicles to make safer and more proactive decisions in traffic environments and holds the potential to serve as a fundamental building block for future collision avoidance systems.

REFERENCES

[1] Tastan, Y., & Kaymaz, H. (2021). Otonom Araçların Önündeki Zorluklar. International Journal of Advances in Engineering and Pure Sciences, 33(2), 195–209. https://doi.org/10.7240/JEPS.741594

[2] Akkaya, S., & Özbay, H. (2022). Otonom Araçların Akıllı Ulaşım Politikaları Üzerindeki Etkileri. Akıllı Ulaşım Sistemleri ve Uygulamaları Dergisi, 5(2), 200–210. https://doi.org/10.51513/JITSA.1160891

[3] Gürtaş, S. (2020). Otonom Araç Sürüş Destek Sistemleri Ve Yapay Zeka Uygulamaları. (Yüksek lisans tezi). Bursa Uludağ Üniversitesi, Fen Bilimleri Enstitüsü, Otomotiv Mühendisliği Anabilim Dalı.

[4] Mutlu, İ. (2021). Otonom Araçlar İçin Optimizasyon Tabanlı Yanal ve Doğrusal Yörünge Planlama. Avrupa Bilim ve Teknoloji Dergisi, 27, 539–548. https://doi.org/10.31590/EJOSAT.932390

[5] Köylü, F., & Atılkan, Y. (2024). Otonom araçlarda derin pekiştirmeli öğrenme yöntemleri ile sollama. Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi, 13(2), 429–439. https://doi.org/10.28948/NGUMUH.1331354

TOMATO IS LIFE: A PROTOTYPE FOR SORTING TOMATOES BASED ON SIZE, COLORS, AND QUALITY

Charl's Benedick Martell W. Inoferio*
Department of Computer Science, Western Mindanao State University, Zamboanga City, Philippines

Rhea Jenn D. Sala,
Department of Computer Science, Western Mindanao State University, Zamboanga City, Philippines

Mary Harolhette O. Rom,
Department of Computer Science, Western Mindanao State University, Zamboanga City, Philippines

Jaydee C. Ballaho,
Department of Computer Science, Western Mindanao State University, Zamboanga City, Philippines
(Corresponding author) qb202100128@wmsu.edu.ph

ABSTRACT

This study presents an automated tomato sorting system designed to classify tomatoes based on size (small and large), color (red and green), and bad quality. The research addresses the challenges and drawbacks of traditional manual sorting, which is labor-intensive, prone to errors, and causes inconsistencies. The objective of this study is to develop an IoT model with computer vision techniques integrated with a machine learning algorithm and measure the segregation and time to complete the sorting. The system utilizes a web camera, conveyor belt, breadboard, power supply, and servo motors as the main components of the system. Images of tomatoes are captured in real-time and processed using Roboflow's Vision Transformer (ViT) model, which classifies them according to predefined parameters. Lastly, the classification results determine the movement of servo motors, which direct the tomatoes into designated containers. The findings reveal that the system achieves a high accuracy of more than 80% in sorting tomatoes in their correct bins with an average sorting time of 28 seconds per tomato. The study concludes that automated sorting systems using machine learning can provide a scalable and reliable solution for agricultural industries, minimize labor costs, and improve post-harvest management.

Keywords: Agriculture, automated sorting, computer vision, machine learning, tomato sorting.

Introduction

Tomato (Solanum lycopersicum) is one of the most widely cultivated fruit crops in the Philippines due to its versatility in culinary applications and high nutritional value. As a warm-season crop, it thrives in various soil types and climatic conditions [1]. The Philippines, being an agricultural country, relies heavily on farming, particularly crop cultivation, which serves as a primary livelihood for many Filipinos [2]. However, traditional farming methods require significant human effort and time, making agriculture labor-intensive and inefficient.

Tomatoes are among the most profitable crops both locally and globally due to their high market demand and economic viability. Ensuring the delivery of high-quality tomatoes is essential for maintaining customer satisfaction, strengthening market competitiveness, and increasing profitability [3]. Despite its economic potential, the tomato industry faces challenges in sorting processes, often resulting in inefficiencies, increased labor costs, and post-harvest losses.

Currently, tomato sorting is primarily performed manually, which presents several drawbacks, including high labor costs, worker fatigue, inconsistencies, and low accuracy [4]. Additionally, relying solely on human judgment for classification increases the risk of sorting errors, leading to decreased product quality and economic losses for farmers [5]. These inefficiencies highlight the need for technological interventions to improve sorting precision and efficiency.

The Internet of Things (IoT) offers promising solutions for optimizing agricultural processes. Computer vision, in particular, enables the objective analysis of tomatoes based on predefined criteria such as color, size, and quality. Therefore, this study aims to develop a Smart Sorting Solution that integrates IoT to enhance sorting accuracy, reduce labor dependency, and minimize product wastage. By modernizing traditional sorting methods, this research seeks to address existing challenges in tomato classification, ultimately improving efficiency and quality control in the industry.

Related Studies

Image Processing and Computer Vision in Tomato Sorting

The use of image processing and machine vision has been effective in assessing tomato quality and automating the sorting process. A study by Zuo [6] highlighted the advantages of automation in reducing manual labor while ensuring consistency in classification. However, challenges such as sensitivity to lighting variations and the need for advanced algorithms for real-time processing were identified. To address these issues, the study suggested incorporating deep learning models and improved sensor technologies.

A neural network-based tomato sorting system was developed using a conveyor belt, camera, and microcontroller, demonstrating a 96.47% accuracy rate in classifying tomato quality. Their findings emphasized the effectiveness of image processing in agricultural automation, proving that computer vision can significantly enhance sorting precision [7].

A tomato sorting system utilizing the HSV color model for color detection, object inner points for size measurement, and a load cell sensor for weight classification achieved

100% accuracy in color detection and 95% accuracy in weight measurement. However, the system developed by Sari et al. [8] struggled with size classification, achieving only 5% accuracy. These results suggest that further improvements in size detection algorithms are necessary to enhance overall classification performance.

Machine Learning and Deep Learning in Tomato Classification

Machine learning techniques have gained significant traction in tomato classification due to their ability to process large datasets, recognize patterns, and enhance classification accuracy over time. Various models, including Artificial Neural Networks (ANNs), Support Vector Machines (SVMs), and Convolutional Neural Networks (CNNs), have been utilized to automate the grading process, improving efficiency in agricultural applications.

An automated tomato grading and classification system was developed by Behera et al. [9], utilizing digital imaging and MATLAB for preprocessing. The system classified tomatoes based on three size categories (small, medium, and large) and three varieties (cherry, classic, and cylindrical). The correlation coefficients (R^2 values) of 0.98 for length and 0.97 for width demonstrated a strong relationship between predicted and actual tomato measurements. These results indicate the high precision of the system in measuring tomato dimensions and classifying them accordingly.

The implementation of Convolutional Neural Networks (CNNs) has significantly improved classification accuracy in tomato sorting. A study conducted by Abekoon et al. [10] developed a CNN-based classification model that achieved 99% accuracy in both training and validation phases for binary classification (ripe and unripe tomatoes). Furthermore, a multi-class classification model attained 99% training accuracy and 98% validation accuracy, demonstrating its robustness in identifying different tomato categories. The study also examined the impact of ripeness on post-harvest longevity, concluding that red tomatoes exhibited the longest shelf life due to their optimal physiological and chemical properties.

Another approach involved integrating K-Means clustering for segmentation with SVM and ANN models for classification by Pati et al. [11], effectively distinguishing fresh tomatoes from defective ones. The study demonstrated that K-Means clustering improved image segmentation, allowing the machine learning models to more accurately differentiate between fresh and defective tomatoes. Despite the system's high performance, researchers recommended further refinements in feature selection to enhance classification accuracy, particularly in distinguishing minor defects from natural imperfections.

Artificial Neural Networks (ANNs) have also been employed in tomato classification by leveraging texture, size, and color features. Girma et al. [12] developed an ANN-based model that successfully classified tomatoes with high accuracy. The dataset comprised 180 images collected from multiple agricultural regions, ensuring diversity in tomato samples. MATLAB was used for preprocessing and feature extraction, contributing to the model's reliability in identifying different tomato characteristics.

Meanwhile, a CNN-based tomato grading system further advanced classification performance by implementing transfer learning to detect surface defects such as rotting and swelling. Tee et al. [13] reported a 97.61% accuracy rate in defect detection, highlighting

the potential of deep learning models in identifying various tomato imperfections. Additionally, an OpenCV-based size measurement algorithm was developed to enhance sorting accuracy. The system achieved a Root Mean Square Error (RMSE) of 0.258 mm in size classification, indicating high precision in determining tomato dimensions. These findings suggest that integrating machine learning with computer vision significantly improves tomato classification and defect detection.

Hybrid and Alternative Approaches in Tomato Sorting

In addition to machine learning techniques, hybrid approaches combining different methodologies have been explored to improve tomato sorting accuracy. A fuzzy logic and machine vision-based sorting system demonstrated enhanced classification performance by incorporating uncertainty handling mechanisms. Unlike conventional machine learning models, fuzzy logic allows for more flexible decision-making by considering varying degrees of ripeness and defect severity. The study, conducted using MATLAB, showed promising results; however, scalability remained a challenge for large-scale applications. The researchers suggested further optimization of fuzzy-based sorting algorithms to facilitate commercial implementation [14].

An alternative method combining K-Means clustering and pixel area subtraction was introduced by Yong et al. [15] to improve classification accuracy. The system incorporated a conveyor belt, cameras, and a controlled lighting setup to enhance image acquisition quality. Experimental results demonstrated that the system achieved 90.00% accuracy in defect detection, 83.33% sorting accuracy, and an improved 95.24% detection coverage for damaged tomatoes. These findings underscore the importance of controlled lighting conditions in maximizing the performance of machine vision-based sorting systems. The researchers emphasized that optimizing illumination and background settings could further reduce misclassification errors, particularly in distinguishing similar-colored tomatoes.

Methodology

Dataset

The dataset used for training and testing was collected from market-purchased tomatoes to ensure a diverse representation of real-world variations. A total of 11,398 high-resolution images were captured from eight different angles under controlled lighting conditions, maintaining consistency and reducing the impact of shadows and reflections. These images were then categorized to facilitate effective model training. As shown in Table 1, the dataset is distributed across different classes. Table 2 outlines the classification of tomatoes, categorizing them into three distinct labels: Red, Green, and Bad. Meanwhile, Table 3 groups tomatoes into different categories based on predefined size ranges.

Table 1: Dataset Distribution

Category Name	Number of Images
Red	3,438
Green	3,473
Bad	4,487

Table 2: Tomato Color Classification

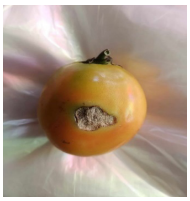
Class Name	**Pictures**	
Green (both green and yellow)		
Red (both red and orange)		
Bad		
		

Table 3: Tomato Size Classification

Class Name	**Picture**	**Max. Diameter**	**Max. Circumference**
Small		<5 cm	≈ 15 cm
Large		> 5 cm	≈ 40 cm

Computer Vision Model

This study undertook two tasks: one is classification (identifying what type of tomato), and the other is object detection (what the camera detects). Pre-trained models were selected from the Roboflow Platform, a cloud-based computer vision tool.

Size Detection

To determine the size of a tomato, a bounding box is created around it. The bounding box measures the width of the tomato in centimeters using a floating-point decimal function. The formula used for computing the real-world width from the bounding box width in pixels is:

$$size = (bbpx \div 225) * 15 \quad (1)$$

Equation 1 converts the bounding box width from pixels to centimeters. The bbpx is the associated bounding box measured in pixels, 225 is the set camera frame after calibration, and 15 is the referencing real-world size of a small tomato. After repeated trial, the set threshold for small tomato is 6.55, otherwise, it will be sorted as large tomato.

Materials and Cost

This research built a prototype using lightweight and alternative materials as illustrated in Table 4 amounting to 12,843.00 pesos.

Table 4: Materials and Cost in Building the Prototype

Material	Quantity	Cost
Steel Foil	1 piece by roll	₱ 432.00
Plywood	1	₱ 880.00
Nail	50	₱ 5.00
Wood Screw	100	₱ 130.00
Jumper Wire • Male – Male • Female – Female Male – Female	by set	₱ 190.00
Dc – Dc Buck Step Down Converter • Adjustable • 5 Volts Fixed	2 pcs adjustable 1 pc 5 volt fixed	₱ 100.00 ₱ 60.00
Acdc 12 Volts 5 Ampere	1 pc	₱ 295.00
16 Channel 12-Bit PWM/Servo Driver-I2C Interface PCA9685 Module	1 pc	₱ 176.00
Raspberry Pi 4b	1 pc	₱ 3,750.00
Raspberry Pi 4 Case With Micro Fan	1pc	₱ 536.00
Web Camera 720p	1pc	₱ 1895.00
Servo Motor • Sg 92 • Sg 90	9 pcs overall all	sg 92- ₱ 160.00 sg 90 – ₱ 105.00
Bred Board	1 pc	₱ 85.00

Material	Quantity	Cost
Stepper Motor	1 pc	₱ 110.00
Stepper Motor Driver	1 pc	included with the stepper motor
Wire	1 pc 1 meter	₱ 35.00
Flash Light (LED Cube Light) Rechargeable	1 pc	₱ 160.00
Embossed Lnoli	1 pc 1 meter	₱ 99.00
No Nails	2 pc	₱ 85.00
Prekit Arduino	1 pc	₱ 1,179.00
Bearing	8 pc	₱ 55.00
Pipe 3 Inch	1pc	1,248.00
Pipe 2 Inch	1 pc	1,000.00
Steel Rod	4 pc	₱ 65.00
Washer	6 pc	₱ 8.00

Statistical Tools

This study employs statistical tools to evaluate the performance, accuracy, and efficiency of the tomato sorter. The key metrics used include overall accuracy both in size and colors, and average sorting time per tomato, as defined in 2 and 3, respectively.

(2)

(3)

Results and Discussion

Evaluation of Roboflow Models

After annotation, the training model was initiated, and customization options such as source image selection, train/test split (default), and preprocessing (augmentation: grayscale and saturation percentage) were applied. In Table 5, a comparison between the Vision Transformer (ViT) and Residual Network (ResNet) models is presented based on their validation accuracy. The results indicate that the ViT model achieved a validation accuracy of 99.5%, whereas the ResNet model attained 98.3%. This comparison highlights that ViT performed better in classifying tomatoes based on their color and quality.

Table 5: Validation Accuracy between Roboflow Models

Model	Validation Accuracy
ViT	99.5%
ResNet	98.3%

Experiment Conducted in the Tomato Sorting Machine

To assess the performance of the developed tomato sorting machine as assembled in Fig. 1, a series of experiments were conducted to measure its accuracy in classifying tomatoes based on size, color, and quality. The experiment involved processing three batches of tomatoes, each containing ten samples. The input parameters recorded included the

actual number of small and large tomatoes, their colors (red, green, or bad), and the start time of sorting. The output results in Table 7 were then compared against the actual values in Table 6 to determine the accuracy of classification.

Figure 1: Prototype Model of the Tomato Sorting

Table 6: Input to the Tomato Sorter before the experiment

Tomato (pcs)	Size		Color			Start Time
	Small	***Large***	***Red***	***Green***	***Bad***	
10	3	4	3	4	3	09:27 AM
10	4	2	4	2	4	09:42 AM
10	5	2	4	3	3	10:18 AM

Table 7: Output of the Tomato Sorter after the experiment

Tomato (pcs)	Size		Color			End Time
	Small	***Large***	***Red***	***Green***	***Bad***	
10	3	2	3	2	3	09:32 AM
10	3	2	3	2	3	09:46 AM
10	4	2	4	2	1	10:23 AM

Evaluation of the Tomato Sorting Machine

In Table 8, the sorting system demonstrated varying levels of accuracy across different criteria, with the highest accuracy observed in identifying red tomatoes at 91.67%, indi-

cating a strong performance in distinguishing them from other categories. Sorting small tomatoes also showed a high accuracy of 85%, while large tomatoes followed closely at 83.33%, suggesting that the system effectively differentiates based on size. However, the accuracy for green tomatoes was lower at 72.33%, indicating some difficulty in correctly classifying them. The lowest accuracy was recorded for bad tomatoes at 69.33%, suggesting that the system struggled to accurately identify defective or rotten tomatoes, likely due to the single-sided camera angle, which may have limited the system's ability to detect defects on all sides of the tomatoes.

Table 8: Accuracy score in terms of size, color, and quality

Criteria in Sorting	Accuracy
Small	85.0%
Large	83.33%
Red	91.67%
Green	72.33%
Bad	69.33%

Additionally, the results indicate that the machine processes each tomato in approximately 28 seconds, demonstrating a reasonable sorting time for real-world applications. This efficiency is crucial in ensuring the system's practicality and usability in large-scale operations.

Overall, the average accuracy of the sorting system was calculated to be 80.33%, demonstrating a generally reliable performance. However, the lower accuracy for green and bad tomatoes suggests areas for improvement, particularly in enhancing the system's ability to distinguish between these categories. Further optimizations, such as refining image recognition algorithms or improving sensor accuracy, may help increase the overall sorting efficiency and reliability, making the system more effective for agricultural and industrial applications.

Conclusion

This study successfully developed an automated tomato sorting system that classifies tomatoes based on size, color, and quality using computer vision and machine learning. By integrating IoT technology with a Vision Transformer (ViT) model, the system achieved an average sorting accuracy of 80% and processed each tomato in approximately 28 seconds. These findings demonstrate the potential of automation in reducing labor dependency, improving classification precision, and minimizing post-harvest losses in the agricultural sector.

The prototype provides a scalable and cost-effective solution for modernizing traditional sorting methods, making it a valuable tool for farmers and agribusinesses. However, certain limitations, such as occasional misclassification and variations in accuracy, highlight areas for improvement. Future work may explore enhancing model performance with larger datasets, incorporating multi-spectral imaging for better defect detection, and optimizing sorting speed through improved hardware integration. Additionally, the system could be expanded to support real-time monitoring and data analytics, enabling farmers

to make informed decisions about crop management and market distribution.

Declaration of Interest: The authors declare that there is no conflict of interest.

Acknowledgements: The authors acknowledge the support of the College of Computing Studies – Western Mindanao State University.

REFERENCES

[1] J. Magsasaka, "Tomato (Kamatis) Cultivation Guide: All You Need to Know in Planting, Growing and Harvesting Tomato." Accessed: Mar. 09, 2025. [Online]. Available: https://www.juanmagsasaka.com/2020/12/tomato-kamatis-cultivation-guide-all.html

[2] L. C. Dorado, J. I. C. Aguila, and R. B. Caldo, "Smart Farm : Automated Classifying and Grading System of Tomatoes using Fuzzy Logic," LPU-Laguna Journal of Engineering and Computer Studies, vol. 3, no. 3, Oct. 2016.

[3] FasterCapital, "The Benefits of Creating a Quality Product", Accessed: Mar. 09, 2025. [Online]. Available: https://fastercapital.com/content/The-Benefits-of-Creating-a-Quality-Product.html

[4] O. O. Arjenaki, P. A. Moghaddam, And A. M. Motlagh, "Online tomato sorting based on shape, maturity, size, and surface defects using machine vision," Turkish Journal of Agriculture and Forestry, Jan. 2013, doi: 10.3906/tar-1201-10.

[5] R. G. de Luna, E. P. Dadios, A. A. Bandala, and R. R. P. Vicerra, "Tomato Fruit Image Dataset for Deep Transfer Learning-based Defect Detection," in 2019 IEEE International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM), IEEE, Nov. 2019, pp. 356–361. doi: 10.1109/CIS-RAM47153.2019.9095778.

[6] H. Zuo, "Analysis and Detection of Tomatoes Quality using Machine Learning Algorithm and Image Processing," International Journal of Advanced Computer Science and Applications, vol. 13, no. 12, 2022, doi: 10.14569/IJACSA.2022.0131250.

[7] Megha. P. Arakeri and Lakshmana, "Computer Vision Based Fruit Grading System for Quality Evaluation of Tomato in Agriculture industry," Procedia Comput Sci, vol. 79, pp. 426–433, 2016, doi: 10.1016/j.procs.2016.03.055.

[8] M. I. Sari, R. Fajar, T. Gunawan, and R. Handayani, "The Use of Image Processing and Sensor in Tomato Sorting Machine by Color, Size, and Weight," JOIV : International Journal on Informatics Visualization, vol. 6, no. 1–2, p. 244, May 2022, doi: 10.30630/joiv.6.1-2.944.

[9] S. Behera, A. Mahapatra, A. Rath, and P. Sethy, "Classification & Grading of Tomatoes using Image Processing Techniques," Mar. 2019.

[10] T. Abekoon, H. Sajindra, J. Jayakody, E. Samarakoon, and U. Rathnayake, "Image processing techniques to identify tomato quality under market conditions," Smart Agricultural Technology, vol. 7, p. 100433, Mar. 2024, doi: 10.1016/j.atech.2024.100433.

[11] K. Patil, "Identifying the Quality of Tomatoes in Image Processing," Int J Res Appl Sci Eng Technol, vol. 10, no. 1, pp. 780–782, Jan. 2022, doi: 10.22214/ijraset.2022.39909.

[12] B. Girma, B. S. Goshu, E. Mengistu, And M. R. Bodke, "Inspection Of Tomatoes Using Image Processing Techniques," Rom J Biophys, vol. 33, no. 4, pp. 133–147, Dec. 2023, doi: 10.59277/RJB.2023.4.01.

[13] S. J. Tee and A. R. bin Mohamed Shariff, "Tomato Grading System With Computer Vision," in Asian Conference on Remote Sensing, 2020.

[14] Engr. E. D. Yong et al., "Tomato Quality Identifier Applying Color-Based Segmentation Using MATLAB with K-Means Clustering and Pixel Area Subtraction," in 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM), IEEE, Nov. 2019, pp. 1–6. doi: 10.1109/HNICEM48295.2019.9072748.

TÜRKÇE E-TİCARET YORUMLARI ÜZERİNDE DUYGU ANALİZİ: BERT İLE KLASİK YÖNTEMLERİN KARŞILAŞTIRILMASI

Melike Çiftçi*
Erciyes Üniversitesi, Bilgisayar Mühendisliği, Kayseri, Türkiye

Nesibe Yalçın
Erciyes Üniversitesi, Bilgisayar Mühendisliği, Kayseri, Türkiye

Celal Öztürk
Erciyes Üniversitesi, Yazılım Mühendisliği, Kayseri, Türkiye
(İletişimden sorumlu yazar) mlkcftc01@gmail.com

ÖZET

Doğal dil işleme (Natural Language Processing, NLP), son yıllarda özellikle müşteri geri bildirimlerinin değerlendirilmesinde önemli bir rol üstlenmiştir. NLP uygulamalarından biri olan duygu analizi, metinlerdeki duygusal eğilimleri belirleyerek işletmelerin stratejik karar alma süreçlerini desteklemektedir. Bu çalışmada, müşteri geri bildirimlerini otomatik olarak analiz edebilmek amacıyla Türkçe e-ticaret sitelerinden toplanan kullanıcı yorumları, duygu durumlarına göre olumlu, olumsuz ve nötr olarak sınıflandırılmıştır. Veri setindeki yorumlar; eksik verilerin temizlenmesi, noktalama işaretlerin kaldırılması ve veri setindeki dengesizliği gidermek amacıyla geri çeviri (back translation) ile veri artırma gibi çeşitli ön işleme adımlarına tabi tutulmuştur. Ön işleme sonrası yorumlar, BERT tabanlı gömme yöntemiyle vektörleştirilmiş ardından çeşitli makine öğrenmesi algoritmaları ve ince ayar yapılmış BERT modeli kullanılarak sınıflandırılmıştır. Kullanılan modellerin başarımları F1-skoru, kesinlik ve duyarlılık metrikleriyle değerlendirilmiştir. Ayrıca modellerin genel başarısını test edebilmek için k-katlı çapraz doğrulama uygulanmıştır. Kullanılan tüm yöntemler arasından ince ayar yapılmış BERT modeli %90 F1-skoru ile en yüksek başarıyı elde ederek Türkçe gibi eklemeli ve bağlama duyarlı dillerde duygu analizinin doğruluğunu belirgin biçimde artırabildiğini göstermiştir.

Anahtar Kelimeler: BERT, Doğal Dil İşleme, Duygu Analizi, Makine Öğrenmesi.

Giriş

E-ticaret siteleri, çevrimiçi depolama sağlayıcıları ve internet kullanıcılarının her geçen gün artmasıyla birlikte, üretilen veri miktarı da hızla çoğalmaktadır. Bu veriler çoğunlukla yapılandırılmamış biçimdedir ve bu nedenle doğru şekilde analiz edilerek düzenlenmeleri büyük önem taşımaktadır. Bu noktada duygu analizi, verilerin anlamlandırılması ve sınıflandırılması sürecinde kritik bir rol üstlenmektedir [1]. Duygu analizi; bireylerin organizasyonlar,

ürünler, hizmetler ve olaylar gibi çeşitli konular hakkındaki görüşlerini ifade ettikleri metinlerde gizli olan duygu, düşünce ve fikirlerin ortaya çıkarılmasına yönelik çalışmalardır [2]. Bu analiz sırasında metin içerisindeki duygu belirten kelimelerin doğru şekilde tespit edilmesi gereklidir. Bu amaçla, önceden tanımlanmış duygu sözlükleri kullanılabileceği gibi öznitelik seçimi yöntemleri ile her bir kelimenin taşıdığı duygu da belirlenebilmektedir [3]. Özellikle e-ticaret sitelerinde kullanıcılar, ürünlerin marka, model ve fiyat bilgilerine ek olarak diğer kullanıcıların deneyimlerine ve değerlendirmelerine büyük önem vermektedir. Bu platformlarda yapılan kullanıcı yorumları, ürün ve hizmetlere dair duygu ve düşüncelerin önemli bir yansımasıdır [4], [5].

Çeşitli alanlarda kullanıcıların ürünler hakkında yaptıkları yorumlar ve görüşler, duygu analizi yöntemleri aracılığıyla anlamlı verilere dönüştürülmektedir. Yılmaz vd. [6] tarafından yapılan çalışmada, Türkiye'nin üç büyük e-ticaret platformundaki kullanıcı yorumları metin madenciliği ve duygu analizi yöntemleriyle incelenmiştir. Çalışmada, kelime bulutu ve bigram analizi gibi görselleştirme teknikleri kullanılmış, yorumların %60'ının olumlu duygular içerdiği, olumsuz yorumların ise genelde kargo ve iade süreçleriyle ilgili olduğu tespit edilmiştir. Kılıçer vd. [7], Türkçe e-ticaret sitelerindeki kullanıcı yorumları üzerinde duygu analizi çalışması yapmışlardır. Çalışmada, Lojistik Regresyon (LR), Destek Vektör Makinesi (DVM), Naïve Bayes (NB), NB Multinomial (NBM), Karar Ağacı (J48) ve Rastgele Orman (RO) algoritmaları uygulamışlar ve NBM algoritması ile en iyi performansı (%69 doğruluk ve %72 F1-skoru) elde etmişlerdir. Çalışmada, nötr yorumların sınıflandırılmasındaki zorluklara ve veri setlerinin genişletilmesi gerekliliğine dikkat çekilmiştir. Görmez vd. [8], Türkçe metinlerde duygu analizi için derin öğrenme modellerinin performanslarını incelemişlerdir. Çalışmada, Beyazperde ve Winvoker veri setlerini kullanılmış, Evrişimli Sinir Ağı (Convolutional Neural Network, CNN), Transformer ve karma modeller geliştirilmiş ve performans değerlendirmesi doğruluk, kesinlik, F1-skor ve duyarlılık gibi metriklerle yapılmıştır. Winvoker veri seti ile eğitilen karma model, %94.16 doğruluk, %91.75 kesinlik, %89.58 duyarlılık ve %90.58 F1-skor değerlerine ulaşarak en iyi performansı göstermiştir. Beyazperde veri seti ile eğitilen modellerde ise CNN, %89.08 doğruluk, 88.82 kesinlik, %89.42 duyarlılık ve %89.12 F1-skor değerleriyle öne çıkmıştır. Ayrıca, ön işleme süreçlerinin etkisi analiz edilmiş ve daha büyük örnek sayısına sahip veri setlerinde ön işlemin model performansına etkisinin azaldığı tespit edilmiştir. Demir vd. [9], Türkçe metinlerde duygu analizi için sözlük tabanlı bir yaklaşım geliştirmişlerdir. Çalışmada, Afinn, Bing ve NRC gibi sözlüklerden yararlanılmış ve Türkçe için özel olarak birleştirilmiş kapsamlı bir sözlük kullanılmıştır. Duygu analizi, kelime seviyesinde çalışmakta ve metinlerin polaritesini (olumlu, olumsuz) belirlemektedir. Araştırma üç farklı veri seti üzerinde yürütülmüştür. İlk veri seti bir film yorum sitesinden elde edilen 70 yorumu (35 olumlu, 35 olumsuz) içerirken, ikinci veri seti bir kitap e-ticaret sitesinden alınan 91309 kullanıcı yorumu içermektedir. Üçüncü veri seti ise 20000 metinden (10000 olumlu, 10000 olumsuz) oluşmaktadır. Çalışmada iki farklı yöntem geliştirilmiş ve yöntemlerin doğruluk oranları değerlendirilmiştir. İlk yöntem %74.17 ile %77.14 arasında doğruluk sağlarken, ikinci yöntem %77.50 ile %82.85 arasında doğruluk oranına ulaşmıştır. Sonuçlar, kelime düzeyindeki hatalar, kinayeli ifadeler ve birden fazla anlam taşıyan deyimlerin sistem performansını etkilediğini göstermiştir. Çalışma, Türkçe metinlerde duygu analizi için sözlük tabanlı yaklaşımların etkinliğini ortaya koymakta ve gelecekteki çalışmalarda mecazi ifadeleri içeren daha kapsamlı sözlüklerin geliştirilmesinin önemini vurgulamaktadır. Tuzcu [10] gerçekleştirdiği çalışmada, çevrimiçi bir kitap satış

sitesinden toplanan 91309 okuyucu yorumu kullanarak duygu analizi yapmıştır. Verilerin sınıflandırılması için Çok Katmanlı Algılayıcı (Multi-Layer Perceptron, MLP), NB, DVM ve LR algoritmalarını kullanmıştır. Sonuçlara göre, MLP algoritması %89 doğruluk oranı ile en iyi performansı göstermiştir. Pozitif ve negatif yorumlara göre başarı oranları sırasıyla %92 ve %86 olarak elde edilmiştir. DVM algoritması, pozitif sınıf için %96.14 oranında yüksek bir başarı göstermesine rağmen negatif sınıflandırmada geride kalmıştır. LR algoritması %84.07 doğruluk oranı ile MLP'ye yakın sonuçlar üretirken, NB algoritması %77.57 doğruluk oranı ile diğer algoritmaların gerisinde kalmıştır. Durak ve Hark [1], e-ticaret sitelerinde kullanıcılar tarafından yapılan yorumlar üzerinde duygu analizi gerçekleştirmişlerdir. Çalışmada, toplamda 76478 İngilizce yorum içeren bir veri seti kullanılmış ve bu yorumlar, pozitif ve negatif olarak iki sınıfa ayrılmıştır. Özellik çıkarımı için CountVectorizer ve TF-IDF Vectorizer yöntemleri uygulanmış ve yorumlar sayısal verilere dönüştürülmüştür. Duygu analizi için NB, LR, RO ve XGBoost gibi makine öğrenmesi algoritmaları kullanılmıştır. Yapılan deneylerde, CountVectorizer yöntemiyle NB algoritması, %84,9 doğruluk oranı ile en yüksek performansı sergilemiştir. LR, %82.4 doğruluk oranı ile ikinci sırada yer alırken, RO %81.1 ve XGBoost %63.8 doğruluk oranlarına ulaşmıştır. TF-IDF yönteminde kelime bazında en yüksek doğruluk oranı, %83.1 ile NB algoritması tarafından elde edilmiştir.

Gerçekleştirdiğimiz çalışmada, Türkçe e-ticaret sitelerinden elde edilen kullanıcı yorumları; olumlu, olumsuz ve nötr olmak üzere üç farklı duygu durumuna ayrılarak sınıflandırılmıştır. Yapılandırılmamış metin verileri üzerinde doğal dil işleme (Natural Language Processing, NLP) ve metin madenciliği teknikleri uygulanmış; çeşitli ön işleme adımlarının ardından Transformatörlerin BERT tabanlı gömme yöntemleri ile metinlerden anlamlı özellikler çıkarılmıştır. Ardından hem klasik makine öğrenmesi algoritmaları hem de ince ayar yapılmış BERT modeli kullanılarak sınıflandırma işlemi gerçekleştirilmiştir. Modellerin değerlendirilmesinde, k-katlı çapraz doğrulama (k-fold cross validation) tekniği kullanılmış ve elde edilen performans sonuçları karşılaştırılmıştır.

Materyal ve Yöntem

Bu bölümde, çalışmada kullanılan veri seti, BERT modeli ve uygulanan sınıflandırma yöntemleri detaylı bir şekilde ele alınmaktadır.

Veri Seti

Bu çalışmada, Kaggle platformunda yayımlanan "E-Ticaret Ürün Yorumları" veri seti [11] kullanılmıştır. Veri seti, **çeşitli** Türkçe e-ticaret platformlarından toplanan kullanıcı yorumlarını içermekte olup yorumların olumlu, olumsuz veya nötr olarak sınıflandırıldığı duygu durumlarını barındırmaktadır. Toplamda 15170 yorum içeren bu veri setinde, duygu sınıflarına göre dağılım şu şekildedir:

- **Olumlu (1):** 6799 yorum (%44.9)
- **Olumsuz (0):** 6978 yorum (%46)
- **Nötr (2):** 1393 yorum (%9.2)

Veri seti, farklı ürün kategorilerinden toplanan çeşitli yorumları içermesi nedeniyle, Türkçe dilinin morfolojik yapısını yansıtan ve duygu analizine uygun geniş bir veri kaynağıdır.

Ancak, nötr yorumların diğer sınıflara kıyasla düşük bir orana sahip olması, sınıflandırma modellerinin performansını dengeleme açısından zorluk teşkil etmiştir.

BERT

Veri ön işleme sürecinin ardından, kullanıcı yorumlarından anlamlı özellikler çıkarmak amacıyla Transformatörlerin Çift Yönlü Kodlayıcı Gösterimleri (Bidirectional Encoder Representations from Transformers, BERT) [12] modeli kullanılmıştır. Bu model, cümleleri yalnızca soldan sağa değil, aynı zamanda sağdan sola da okuyarak bağlamı çift yönlü olarak öğrenebilme özelliğine sahiptir. **Çalışmada, Türkçe diline özel olarak eğitilmiş "dbmdz/bert-base-turkish-cased"** adlı BERT modeli tercih edilmiştir. Bu model, Hugging Face Transformers kütüphanesi aracılığıyla yüklenmiş ve kullanılmıştır. Metinler, BERT tokenizer ile parçalanmış (tokenize edilmiş), her bir yorum sabit uzunlukta olacak şekilde maksimum 512 token'a kadar kısıtlanmış ve giriş tensörlerine dönüştürülmüştür. Bu işlem sonucunda model, her yorum için 768 boyutlu bir özellik vektörü (embedding) üretmiştir. Özellikle, her cümlenin genel temsili olarak BERT'in özel [CLS] token'ına karşılık gelen vektör alınmıştır.

BERT tabanlı gömme yöntemi, kelimeler arasındaki bağlamı öğrenme yeteneği sayesinde geleneksel kelime sıklığı tabanlı yöntemlere kıyasla daha anlamlı ve etkili bir vektörsel karşılık sunmaktadır. Bu yöntem, yorumların duygu durumlarının doğru bir şekilde sınıflandırılmasında önemli bir rol oynamıştır.

Makine Öğrenmesi Algoritmaları

Bu çalışmada, kullanıcı yorumlarının duygu durumlarına göre (olumlu, olumsuz ve nötr) sınıflandırılması amacıyla DVM, LR ve RO makine öğrenmesi algoritmaları kullanılmıştır.

DVM

Destek Vektör Makineleri (DVM) [13], sınıflandırma problemlerinde yaygın olarak kullanılan, etkili ve matematiksel temeli güçlü bir algoritmadır. Temel amacı, farklı sınıflara ait verileri maksimum marjin sağlayan bir hiper düzlem ile en iyi şekilde ayırmaktır. Lineer olarak ayrılmayan verilerde ise çekirdek (kernel) fonksiyonları ile veri daha yüksek boyutlara yansıtılarak ayrılabilir hale getirilmektedir.

LR

Lojistik Regresyon (LR) [14], adında "regresyon" geçmesine rağmen, sınıflandırma problemlerinde yaygın olarak kullanılan bir genelleştirilmiş doğrusal model türüdür. Olasılıksal temellere dayanan bu yöntem, bağımlı değişkenin kategorik olduğu durumlarda tercih edilir. Model, bağımlı değişken ile bir veya daha fazla bağımsız değişken arasındaki ilişkiyi bir lojistik (sigmoid) fonksiyon aracılığıyla ifade eder. Elde edilen çıktı, bir sınıfa ait olma olasılığını verir.

RO

Rastgele Orman (RO) [15], topluluk (ensemble) öğrenme yöntemlerinden biridir ve çok sayıda karar ağacından oluşan bir sınıflandırma algoritmasıdır. Her bir ağaç, eğitim verisinin rastgele seçilen bir örneklemi (bootstrap) üzerinden inşa edilir ve her düğümde özellik-

lerin rastgele bir alt kümesi dikkate alınarak ayrım yapılır. Bu çeşitlilik, modelin aşırı öğrenme (overfitting) riskini azaltırken genelleme performansını artırır. Tahmin sürecinde her ağaç bağımsız olarak karar verir ve nihai sınıf, çoğunluk oylaması ile belirlenir.

Uygulama

Türkçe kullanıcı yorumlarının duygu analizine hazır hale getirilmesi için ilk olarak veri seti üzerinde çeşitli veri ön işleme adımları uygulanmıştır. Daha sonra sayısal vektörlere dönüştürülmüş ve model eğitimleri gerçekleştirilmiştir. Çalışmada, Python programlama dili kullanılmıştır. Uygulama ortamı olarak Google Colab tercih edilmiş ve GPU desteği sayesinde modellerin eğitim süreci hızlandırılmıştır.

Veri Ön İşleme

Veri setindeki kullanıcı yorumlarının **sınıflandırmaya** uygun hale getirilmesi amacıyla çeşitli veri ön **işleme adımları uygulanmıştır**:

1. Veri Artırma (Back Translation): Azınlıkta kalan nötr sınıfa ait örneklerin sayısını artırmak amacıyla, yorumlar Türkçe'den İngilizce'ye, ardından tekrar Türkçe'ye çevrilerek veri artırımı gerçekleştirilmiştir.

2. Eksik Verilerin Kaldırılması: Yorum sütununda bulunan eksik veya boş veriler temizlenmiştir. Yorum içermeyen satırlar analiz sürecine dahil edilmemiştir.

3.Noktalama İşaretlerinin Kaldırılması: Yorumlarda bulunan noktalama işaretleri kaldırılmış ve metinler sadeleştirilmiştir. Bu adım, modelin yorumlardaki dilbilgisel yapıdan bağımsız olarak anlam odaklı çalışmasına yardımcı olmuştur.

4. Küçük Harfe Dönüştürme: Yorumlar tamamen küçük harfe dönüştürülerek metinler standardize edilmiştir.

5.Sayıların Kaldırılması: Yorumlarda geçen sayılar kaldırılarak metnin sadece kelime temsiliyle modellenmesi sağlanmıştır

6.HTML Etiketlerinin Kaldırılması: HTML formatından gelen metinlerde bulunan <br>, <p> gibi etiketler temizlenmiştir.

7.Fazla Boşlukların Temizlenmesi: Kelimeler arasında birden fazla boşluk varsa, bu boşluklar kaldırılarak metinler daha düzenli hale getirilmiştir.

8. Türkçe Olmayan Yorumların Temizlenmesi: Dil tespiti ile Türkçe dışındaki yorumlar veri setinden ayıklanmıştır.

9.Tekrarlayan Satırların Silinmesi: Veri setinde yer alan yinelenen yorumlar tespit edilerek çıkarılmıştır.

Bu adımların ardından, metinlerin sınıflandırma için daha temiz ve düzenli bir yapıya dönüştürülmesi sağlanmıştır. **Sınıflandırma modellerin** etkinliğini değerlendirmek amacıyla veri seti, **eğitim** veri seti (%70) ve test veri seti (%30) olmak üzere ikiye ayrılmıştır. Modellerin genelleme yeteneğini değerlendirmek amacıyla 5-katlı **çapraz doğrulama** tekniği de uygulanmıştır.

Sınıflandırma

Çalışmada önerilen model, önceden eğitilmiş BERT modeli temel alınarak geliştirilmiştir. Duygu analizi görevine özel olarak model **üzerin**de ince ayar (fine-tuning) uygulanmıştır. Modelin eğitimi öncesinde tüm metinler, BERT tokenizer yardımıyla 512 uzunluğa kadar sabitlenmiş ve özel [CLS] ve [SEP] token'ları ile biçimlendirilmiştir. Bu işlem sırasında her örnek için bir dikkat maskesi (attention mask) de oluşturulmuştur. İnce ayar **işlemi sırasında kullanılan başlıca hiperparametreler** aşağıda verilmiştir:

- Döngü Sayısı (Epoch): 3
- Grup Boyutu (Batch Size): 8
- **Öğrenme Oranı (Learning Rate): 2e-5**
- Ağırlık Kaybı (Weight Decay): 0.01

İnce ayar yapılmış BERT (FINE-TUNED BERT) modelinin başarımı, aynı veri seti üzerinde BERT ile vektörleştirilen girdiler üzerinden eğitilmiş üç farklı makine öğrenmesi modeli ile karşılaştırılmıştır: BERT+LR, BERT+DVM ve BERT+RO. Tüm modeller, ağırlıklı ortalamalara göre F1-skoru (F1-score), kesinlik (precision) ve duyarlılık (recall) metrikleri üzerinden değerlendirilmiştir. Karşılaştırmalı sonuçlar Tablo 1'de sunulmuştur. Önerilen FINE-TUNED BERT modeli, en yüksek duyarlılık (0.903), kesinlik (0.906) ve F1-skoru (0.904) değerlerini vermiştir. BERT+LR modeli; 0.825 duyarlılık, 0.820 kesinlik ve 0.822 F1-skoru ile ikinci en yüksek başarımı göstermiştir.

Tablo 1: Sınıflandırma Modellerinin Başarımı

Model	Duyarlılık	Kesinlik	F1-skoru
BERT+LR	0.825	0.820	0.822
BERT+DVM	0.820	0.816	0.817
BERT+RO	0.760	0.752	0.716
FINE-TUNED BERT	0.903	0.906	0.904

Sonuç ve Tartışma

Bu çalışmada, Türkçe e-ticaret yorumlarının duygu durumlarını sınıflandırmak amacıyla **çeşitli makine öğrenme yöntemleri tabanlı modeller ile BERT modeli** karşılaştırmalı olarak incelenmiştir. Çalışmada kullanılan veri setinde, sınıf dengesizliği problemi özellikle nötr sınıfta belirgin olduğundan, geri çeviri temelli veri artırma stratejisi uygulanmış ve başarım artırılmıştır. **Özellikle BERT tabanlı gömme yöntemi ile elde edilen vektör temsilleri, L**R, DVM ve RO makine öğrenmesi algoritmalarıyla değerlendirilmiştir. Yapılan deneyler sonucunda, LR ve DVM tabanlı modellerin doğruluk ve F1-skoru gibi metriklerde literatürde aynı veri seti ile yapılan çalışmalara kıyasla daha iyi sonuçlar verdiği gözlemlenmiştir.

Bununla birlikte, önceden eğitilmiş BERT modeline ince ayar uygulanarak oluşturulan sınıflandırıcı, tüm modeller arasında en yüksek başarım**ı göstermiştir. İnce ayar yapılmış** BERT modeli, test verisi üzerinde %90 F1-skoru ile öne çıkmıştır. Özellikle bağlamsal yapıya duyarlı mimarisi sayesinde, nötr sınıf dahil olmak üzere tüm sınıflarda dengeli ve yüksek başarım sağlamıştır. Literatürdeki farklı duygu analizi **çalışmaları ile** karşılaştırmalı sınıflandırma başarımlarına dair detaylı bilgi Tablo 2'de verilmiştir.

Tablo 2: Literatür Karşılaştırması

Kaynak	Veri Seti	Sınıflar	Yöntem	En iyi Sonuç
[1]	**İngilizce yorumlar**	Olumlu, Olumsuz	NB, LR, RO, XGBoost	NB - %84.9
[6]	Google Play	Olumlu, Olumsuz	Görselleştirme, Bigram	-
[7]	Türkçe e-ticaret yorumları	Olumsuz, Olumlu, Nötr	NB, NBM, LR, DVM, RO, J48	NBM - %72.0
[8]	Winvoker / Beyazperde	Olumlu, Olumsuz	CNN, Transformer, Karma Model	CNN - %89.1
[9]	Film, kitap, genel yorumlar	Olumlu, Olumsuz	Sözlük Tabanlı Yöntemler	%82.8
[10]	Kitap satış sitesi	Olumlu, Olumsuz	MLP, NB, DVM, LR	MLP - %89.0
Bu çalışma	Türkçe e-ticaret yorumları	Olumsuz, Olumlu, Nötr	BERT, LR, DVM, RO	FINE-TUNED BERT - %90

Değerlendirme ve Öneriler

Bu çalışmada duygu analizine yönelik, makine öğrenmesi tabanlı modeller geliştirilmiş ve önceden eğitilmiş bir BERT modeli ince ayar yöntemi ile optimize edilmiştir. Çalışma kapsamında elde edilen bulgular, geleneksel makine öğrenmesi yöntemlerinin Türkçe duygu analizi problemlerinde etkili bir şekilde uygulanabileceğini göstermiştir. Bununla birlikte, BERT gibi bağlamsal dil modelleri üzerine yapılan ince ayar, güçlü ön işleme ve veri artırma stratejileriyle birleştirildiğinde, duygu sınıflandırmasında makine öğrenmesi modellerinden daha üstün performans sergilemiştir. Önerilen modelin olumlu, olumsuz ve nötr yerine daha ince duygu durumları için analizleri yapılabilir ve farklı metin türlerinde başarımı incelenebilir.

Gelecekte yapılacak çalışmalarda, daha büyük ve çeşitli veri setlerinin kullanılması, sınıf dengesizliklerinin daha gelişmiş yöntemlerle giderilmesi ve alternatif transformer tabanlı modellerin (örneğin RoBERTa, ELECTRA, DeBERTa gibi) kullanılması önerilmektedir.

Yazar Katkıları: Melike Çiftçi; Planlama ve Tasarım, Literatür Taraması, Veri Analizi, Model Geliştirme, Parametre Optimizasyonu, Sonuçları Analiz Etme, Bildiriyi Yazma. Nesibe Yalçın; Genel Tasarım, Literatür İncelemesi, Sonuçları Değerlendirme, Bildiriyi Gözden Geçirme. Celal Öztürk; **Çalışmanın Akademik Yönlendirmesi, Metodolojik Çerçevenin Oluşturulması ve Genel Araştırma Sürecinin Danışmanlığı.**

REFERANSLAR

[1] Durak, M. H., Hark, C. (2024) Makinelerin Öğrenmesi Yöntemlerinin Duygu Okuma Yeteneği: E-Ticaret Yorumlarını Anlamak. Computer Science, 9(2), 133-141.

[2] Özyurt, B., Akcayol, M. A. (2018). Fikir Madenciliği ve Duygu Analizi, Yaklaşımlar, Yöntemler Üzerine Bir Araştırma. Selçuk Üniversitesi Mühendislik, Bilim ve Teknoloji Dergisi, 6(4), 668-693.

[3] Yavuz, G. (2023). Web kazıma ve duygu analizi temelli ürün analiz sistemi. Aksaray Üniversitesi Sosyal Bilimler Enstitüsü, Yüksek Lisans Tezi.

[4] Uzun, E., Uçar, M. (2023). E-Ticaret Sitelerinde Yer Alan Tüketici Yorumlarının Metin Madenciliği Yöntemleri ile Analizi. 4. International Cappadocia Scientific Research Congress, Nevşehir, Türkiye, 1010-1018.

[5] Bozkurt, A. H., Yalçın, N. (2024). Topluluk Öğrenmesi Algoritmaları Kullanarak Amazon Yemek Yorumları Üzerine Duygu Analizi. Bilecik Şeyh Edebali Üniversitesi Fen Bilimleri Dergisi, 11(1), 128-139.

[6] Yılmaz, F., Adalı, G. K. (2024). Metin Madenciliği Yaklaşımı ile E-Ticaret Sitesi Uygulamalarının Müşteri Yorumlarına Yönelik Duygu Analizi. Fenerbahçe Üniversitesi Sosyal Bilimler Dergisi, 4(2), 57-75.

[7] Kılıçer, S., Şamlı, R. (2023). E-Ticaret Sitelerindeki Türkçe Ürün Yorumları Üzerine Makine Öğrenmesi Algoritmaları ile Duygu Analizi. Veri Bilimi, 6(2), 15-23.

[8] Görmez, Y., Arslan, H., Atak, B. (2024). Türkçe Metinlerde Duygu Analizi: Derin Öğrenme Yaklaşımlarının ve Ön İşlem Süreçlerinin Model Performansına Etkisi. Fırat Üniversitesi Mühendislik Bilimleri Dergisi, 36(1), 509-520.

[9] Demir, Ö., Chawai, A. I. B., Doğan, B. (2019). Türkçe Metinlerde Sözlük Tabanlı Yaklaşımla Duygu Analizi ve Görselleştirme. International Periodical of Recent Technologies in Applied Engineering, 1(2), 58-66.

[10] Tuzcu, S. (2020). Çevrimiçi Kullanıcı Yorumlarının Duygu Analizi ile Sınıflandırılması. Eskişehir Türk Dünyası Uygulama ve Araştırma Merkezi Bilişim Dergisi, 1(2), 1-5.

[11] Çabuk, M. (2021). E-Ticaret Ürün Yorumları. Erişim Tarihi: 09/04/2025. https://www.kaggle.com/datasets/mujdatcabuk/eticaret-urun-yorumlari

[12] Aftan, S., Shah, H. (2023). A Survey on Bert and Its Applications. 20th Learning and Technology Conference (L&T), IEEE, 161-166.

[13] Hearst, M. A., Dumais, S. T., Osuna, E., Platt, J., Scholkopf, B. (1998). Support Vector Machines. IEEE Intelligent Systems and Their Applications, 13(4), 18-28.

[14] Nick, T. G., Campbell, K. M. (2007). Logistic Regression. Topics in biostatistics, 273-301.

[15] Biau, G., Scornet, E. (2016). A Random Forest Guided Tour. Test, 25(2), 197-227.

UZAKTAN ALGILAMA İÇİN TARIM ODAKLI ÇÖZÜNÜRLÜK ARTIRIMININ YEŞİL YOĞUN BÖLGELERDE İNCELENMESİ

Deniz Çelik*
Sakarya Uygulamalı Bilimler Üniversitesi, Bilgisayar Mühendisliği, İstanbul, Türkiye

Ali Furkan Kamanlı
Sakarya Uygulamalı Bilimler Üniversitesi, Elektrik ve Elektronik Mühendisliği, Sakarya, Türkiye

Buket Bingöl
Sakarya Uygulamalı Bilimler Üniversitesi, Bilgisayar Mühendisliği, İstanbul, Türkiye
(İletişimden sorumlu yazar) denizcelik2@outlook.com

ÖZET

Bu çalışmada, uzaktan algılama görüntü verilerinin tarım odaklı analizlerine yönelik kullanımında çözünürlük artırımı etkisinin incelenmesi için bir çözüm alternatifi sunulması hedeflenmiştir. Düşük çözünürlüklü uzaktan algılama görüntülerinin, özellikle tarımsal planlama ve kaynak yönetimi açısından verimliliğini arttırmak amacıyla, derin öğrenme tabanlı SwinIR modeli kullanılarak bir çözünürlük artırımı çözümü oluşturulmuş ve incelenmiştir. Cezayir Uzay Ajansı'na ait Alsat-2B uydusundan elde edilen görüntülerden oluşan bir veri seti üzerinde, 4x ölçekleme oranıyla düşük ve yüksek çözünürlüklü görüntü çiftleri oluşturulmuş, bu veri seti modelin ince ayarlama sürecinde kullanılmıştır. Swin Transformer mimarisi üzerine kurulu SwinIR modeliyle, görüntülerdeki bağlamsal bilgiler hem yerel hem küresel ölçekte değerlendirilebilmiş ve görüntü detayları başarılı şekilde yeniden yapılandırılabilmiştir. Modelin başarımı, PSNR ve SSIM metrikleri ile ölçülmüştür. Elde edilen sonuçlarla, sınırlı veri ve eğitim süresine rağmen modelin tarımsal alanların daha detaylı izlenmesine katkı sunabileceği gösterilmiştir. Çalışmanın özellikle düşük çözünürlüklü uydu verilerinin etkin kullanımını mümkün kılması ile, uzaktan algılama tabanlı tarım uygulamaları için bir alternatif oluşturabileceği düşünülmektedir.

Anahtar Kelimeler: Süper çözünürlük, uzaktan algılama, yapay zekâ.

Giriş

Uzaktan algılama teknolojileri, tarım arazilerinin izlenmesi, planlaması ve sürdürülebilir kullanımı açısından yenilikçi ve verimli yaklaşımlar sunmaktadır. Günümüzde geniş alanların uydu kaynaklı izlenmesi ve bu sayede mümkün olan büyük ölçekli veri temini ile tarım alanında birçok analizin daha hassas biçimde yapılabilmesine imkân tanınmaktadır. Ancak bu yöntemlerin kullanımında bazı kısıtlayıcı faktörler mevcuttur. En önemli faktörlerden biri, yaygın ve düşük maliyetli görüntüleme kaynaklarında görece düşük olabilen görüntü çözünürlüğüdür. Düşük çözünürlüklü görüntülerden edinilen bilgilerin; hassas analizlerde yetersizliklere, verimsiz tarımsal planlamalara, kaynak israfına ve küçük alanların yanlış sınırlarla

belirlenmesine yol açabileceği düşünülmektedir. Bu bağlamda yapılmış olan çözünürlük arttırımı daha kaliteli veriler sunması ile, mevcut kaynaklardan daha yüksek bilgi çıkarılmasına, hassas tarımsal uygulamalara imkân tanınmasına ve doğruluk oranlarının iyileştirilmesine zemin hazırlayarak, kısıtlayıcı etmenleri büyük ölçüde ortadan kaldırabilmektedir. Aynı zamanda sık ve kolay veri temini, alana gitme ve manuel yöntemlere başvurma ihtiyacı doğurmadan sağlanabildiği için sağladığı ekonomik artılardandır.

Nüfus artışı, iklim değişiklikleri ve diğer küresel sorunlar; tarım arazilerinin daha doğru kullanılması ve verimliliğin artması konularında çalışmaların genişletilmesinde tetikleyici rol oynamıştır. Sağlanan çözüm, tarım verimliliği alanındaki çalışmaların derinleşmesi için gerekli veri kaynaklarının edinimi ve mevcut veri kaynaklarının zenginleştirilmesi konusunda yardımcı olabilmektedir. Bu katkı, yeni uydu sistemlerine olan bağımlılığı da azaltmaktadır. Elde edilen detaylı mekânsal veriler mahsul türü analizi, bitkilerde hastalık tespiti gibi farklı tarımsal uygulamalara katkı sağlayabilir. Uzaktan algılamada çözünürlük arttırma konularında farklı çalışmalar bulunmaktadır. Zhang ve arkadaşları (2025), SwinFR, SwinIR ve Hızlı Fourier Dönüşümü tekniklerini birleştirerek, görüntülerin çözünürlüğünü artırmada etkili bir yöntem sunmuşlardır. Bu yaklaşım, yüksek kaliteli görüntü iyileştirmeleri sağlayarak, süper çözünürlük alanında önemli bir katkı sağlamaktadır [1]. Bu konuda yapılan başka bir çalışmada Tu ve arkadaşları evrişimli katmanlar ile swin transformer mimarisinin avantajlarını birleştiren ve SWCGAN olarak adlandırılan yeni bir üretici adversaryal ağ (GAN) yapısı önermiştir. Çalışmada ilk olarak sığ özellik çıkarımı için bir evrişimli katman kullanılmıştır. Sonrasında yüksek çözünürlüklü görüntüler üretmek üzere derin özellikleri çıkaran swin transformer blokları geliştirilmiştir. Son olarak da adversaryal eğitim için ayrıştırıcı Swin Transformer yapısı kullanılmıştır. Çalışmanın sonucunda önerilen yöntemin problemin çözümünde başarılı olduğu görülmektedir [2]. Görüntü verilerinin süper çözünürlüğü üzerine bir diğer çalışma Lei ve arkadaşları (2021) tarafından yapılmıştır. Çalışma ile önerilen Transformer tabanlı TransENet yapısı kullanılarak görüntüler üzerinde çok kademeli özellik çıkarımı kullanılmış ve düşük çözünürlüklü görüntülerin yüksek çözünürlüklü edinimi sağlanmıştır. Çalışmanın deneysel sonuçları, TransENet yapısı ile yüksek frekanslı ayrıntıların geri kazanımında ve görsel kalite metriklerinde kayda değer iyileşmeler elde edildiğini göstermektedir [3].

Materyal ve Metot

Veri Seti ve Ön İşleme

Çalışma içeriğinde kullanılan veri kümesi, bir literatür çalışması olarak çözünürlük artırımı çalışmalarına yönelik hazırlanmış bir çalışma yoluyla Kaggle platformu üzerinden edinilmiştir. Djerida ve arkadaşlarının (2021) yürüttüğü çalışma kapsamında Cezayir Uzay Ajansı tarafından işletilen Alsat-2B uydusundan edinilen görüntülerin derlenmesiyle oluşturulmuştur. Veri seti içeriğinde 2.5m pankromatik ve 10m çözünürlüklü çok bantlı (kırmızı, yeşil, mavi, yakın kızıl ötesi ve vb.) görüntü sağlayabilen uydu kaynağı kullanılmıştır [4]. İlgili çalışma içeriğinde ön işlemeler sonrası görüntü örnekleri için dört kat (4x) ölçekleme oranını yansıtacak biçimde 64x64 boyutlu düşük çözünürlüklü (LR) renkli (RGB) görüntü ve 256x256 boyutlu yüksek çözünürlüklü (HR) renkli (RGB) görüntü veri örneği çiftleri oluşturulmuştur [4].

Veri kümesindeki toplamda 2759 görüntü çiftinin 2182 adedi eğitim kümesinde ve 577 adedi test kümesinde olacak biçimde ayrılmıştır. Veri seti içeriğindeki görüntü parçaları

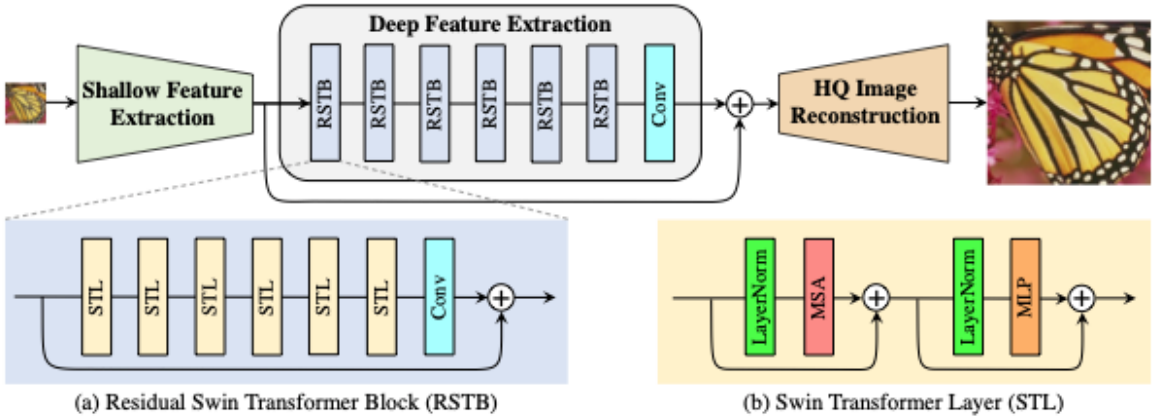
(patch); kentsel alanlar, tarım arazileri, stadyum ve köprü benzeri özel yapıları da kapsayan karmaşık örnekleri içerek biçimde edinilmiştir.

Model Mimarisi

Görüntü çözünürlüğü artırımı, yapay zekâ literatüründe uzun yıllardır çalışılan bir konu olarak derin öğrenme modellerinin yükselişinden bu yana birçok farklı yaklaşım altında incelenmiş ve çeşitli çalışmalar literatüre kazandırılmıştır. CNN tabanlı bir çözüm olarak Süper Çözünürlüklü Konvolüsyonel Sinir Ağları (SR-CNN) [5] yapısı bu alandaki ilk örneklerdendir. Bir diğer çalışma içeriğinde Generative Adversarial Network (GAN) yapısını kullanan SR-GAN [6] modeli ile GAN kullanımının çözünürlük artırımı odaklı işletilmesi incelenmiştir. Transformer mimarisinin öncelikle doğal dil problemleri, ardından bilgisayarlı görü uygulamalarındaki başarımının anlaşılması sonrası bu yapının çözünürlük artırımı çözümü olarak kullanılmasını inceleyen çeşitli çalışmalar [7] yayımlanmıştır.

Çözünürlük artırımının tarım odaklı uzaktan algılamada kullanımını inceleyen bu çalışma içeriğinde, güncel bir mimari ile uzaktan algılama görüntü verilerinin açık erişimi mümkün olan veya görece düşük maliyetli veri edinimi sağlayan Sentinel-2 ve benzeri uydu kaynaklarını hedefleyen bir çözünürlük artırımı sisteminin geliştirilmesi incelenmiştir. Bahsedilen standartlarda uydu veri kaynakları ile edinilebilen 10 metre/piksel veya daha düşük çözünürlüklü yer görüntülerinin 2 metre/piksel (ve yakın değerler) düzeyinde iyileştirilmesi ve bu sayede uzaktan algılama ile yürütülebilen tarım analizlerinin verimlileştirilmesi amaçlanmıştır.

Veri kümesinin modellenmesi ve çözünürlük artırımının tarım odaklı olarak incelenmesi için Transformer mimarisine dayalı güncel bir çözünürlük artırımı modeli olan SwinIR kullanılmıştır. SwinIR modeli, Şekil 1'de açıklandığı gibi içerisindeki Swin Transformer [8] mimarisi aracılığıyla görsel içerisindeki bağlamsal bilgiyi yerel ve küresel düzeyde çıkarmak için oluşturulmuş pencere tabanlı çok uçlu dikkat mekanizmasından (window-based multi-head attention) faydalanmaktadır [9].



Şekil 1: Görüntü restorasyonu için SwinIR mimarisi [9]

SwinIR modeli, tahmin üretmek için içerisinde barındırdığı üç ayrı modülden faydalanmaktadır. İlk aşamada yüzeysel bilgi edinilmesi için "yüzeysel özellik çıkarıcı" (shallow feature extractor) yapısını kullanır. Daha sonrasında yüksek düzeyde karmaşık bilginin edinimi

için Swin Transformer bloklarını içeren "derin özellik çıkarıcı" (deep feature extractor) yapısını kullanmaktadır. Son adımda "görüntü yeniden yapılandırma" (image reconstruction) bloğunda birleştirilen özellikler kullanılarak yüksek çözünürlüklü görüntü elde edilmiş olur. SwinIR, önerdiği "kaydırmalı pencere" odaklı dikkat mekanizması ile farklı çözünürlük seviyelerinde bilgiyi birleştirmektedir. Bu yapı sayesinde görüntü içerisinde ele alınan sahneye dair detaylar, doğru biçimde eşleştirilerek SwinIR modelinin genel başarımını önemli ölçüde yükseltmiştir [9].

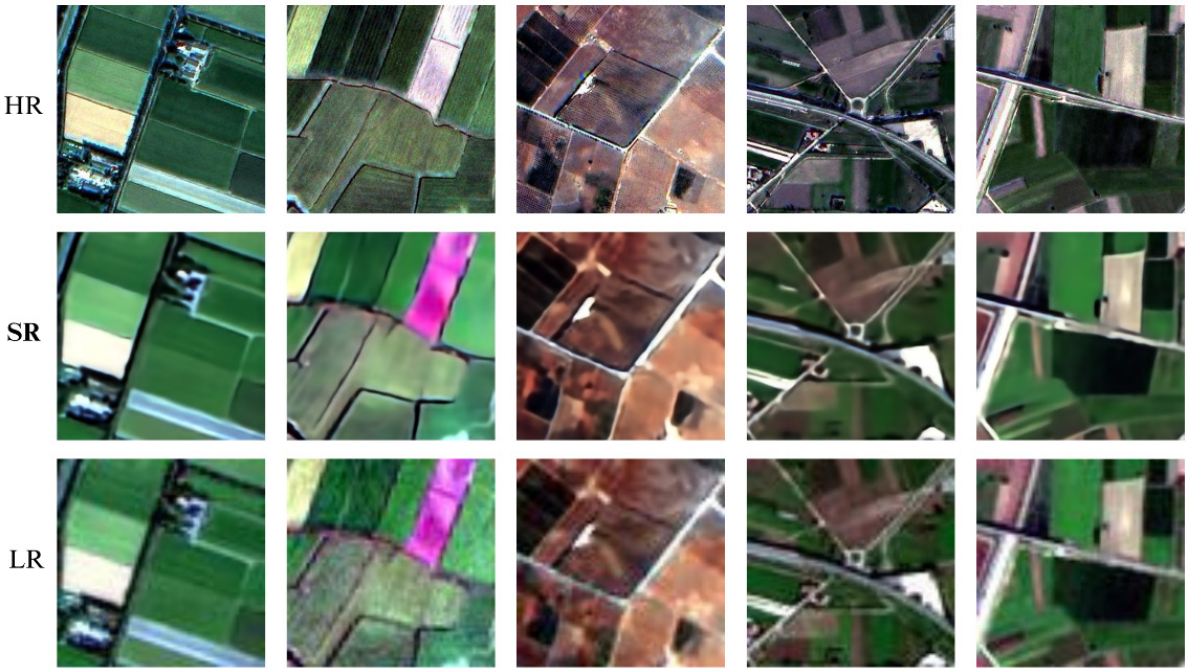
SwinIR modelinin çalışma içeriğinde elde edilen veri kümesi üzerindeki başarımının anlaşılması için model ince-ayarlama (fine-tuning) süreci yürütülmüştür. Bunun için SwinIR modelinin resmi kaynak kodlarını içeren GitHub sayfası kullanılmış ve gerekli konfigürasyonun tanımlanması sonrası ince ayarlama uygulanmıştır. Bu ince ayarlama sürecinde modelin yeniden eğitimi için varsayılan değerler olan L1 kayıp fonksiyonu ve Adam optimizasyon algoritması kullanılmıştır. Python dilinde ve PyTorch kütüphanesi ile oluşturulmuş bu kaynak kodu deposu, ince ayarlama yönergesi ve eldeki veriler kullanılarak işletilmiştir. Modelin eğitim adımındaki hesaplama işlem yükünün temini için AWS platformu üzerindeki A100 Graphics Processing Unit (GPU) makine örnekleri kullanılmıştır ve elde edilen ara model örnekleri bu platformda saklanmıştır.

Değerlendirme Metrikleri

Yürütülen görece kısa süreli ve küçük ölçekli ince ayarlama sürecinin sonucunda elde edilen model örneğinin başarımının anlaşılması için Peak Signal-to-Noise Ratio (PSNR) ve Structural Similarity Index (SSIM) metrikleri kullanılmıştır. PSNR metriği, işlenen görüntülerin orijinal ve yeniden yapılandırılmış halleri arasındaki farkı piksel düzeyinde ölçmek için oluşturulmuş bir değerlendirme kriteri olarak kullanılmıştır. SSIM metriği ise sadece piksel düzeyinde farkın ölçümüyle sınırlı kalmayarak sistemin iki durum arasındaki parlaklık, kontrast değerleri ve yapısal benzerliğinin değerlendirilmesi amacıyla kullanılmıştır.

İnce ayarlama sonucu elde edilen model örneğinin test kümesi içerisindeki 55 adet görüntü örneği üzerinde değerlendirilmesiyle 15.02 dB maksimum PSNR skoru elde edilmiştir. Bununla beraber yapısal tutarlılığın incelenmesinde kullanılan SSIM metriği için yine 0.347 skoru elde edilmiştir.

Kullanılan test kümesi içerisindeki "tarım" alt kümesi üzerinde görüntü çözünürlük artırımı / restorasyonu çıktıları Şekil 2'de SR satırı ile gösterilmiştir.



Şekil 2: Tarım test kümesi örnekleri üzerinde çözünürlük artırımı gösterimi

Tartışma

Önerilen sistemin problem için elde edilen veri kümesi üzerindeki başarımı, literatür standardı metrikler ile çalışma kapsamında incelenmiştir. Buna göre SwinIR modelinin veri seti üzerindeki başarımı, genel ölçekte görece düşük görünmesine karşın, veri kümesi ile yayımlanan çalışma çıktılarında [4] paylaşılan diğer alternatiflerle yakın ve tutarlı bir sonuç ürettiği gözlemlenmiştir.

Çalışmada elde edilen sonuçlar göz önünde bulundurulduğunda, model performansının kısıtlı bir veri kümesi ile ince ayar sürecinden önemli ölçüde etkilediği gözlemlenmektedir. Ayrıca, incelenen modelin uzaktan algılama alanında bir gerçek hayat çözümü olarak kullanılabilmesi için veri kısıtlarının giderilmesinin önemi ifade edilmiştir. Diğer yandan, modelin veri kümesi için incelenen diğer alternatif modellerle yakın ve tutarlı başarım eldesi ile Transformer tabanlı bir çözünürlük artırımı / görüntü restorasyonu modelinin uzaktan algılama problemlerinde değerlendirmeye açık bir alternatif olduğu ifade edilebilmektedir. Gelecek çalışmalarda mevcut veri kümesi kısıtlarını içermeyen farklı ve gelişkin bir veri kümesinin daha uzun bir ince ayarlama sürecinde değerlendirilmesi hedeflenmektedir. Bu sayede SwinIR modeli ve diğer Transformer tabanlı çözümlerin uzaktan algılama ve dahası tarım alanındaki potansiyelinin geniş perspektiften incelenmesi hedeflenmektedir.

Sonuç

Çalışma kapsamında uzaktan algılama verilerinin tarımsal incelemelere yönelik olarak kullanımda verimliliği artırmak hedefiyle bir derin öğrenme çözünürlük artırımı modeli incelenmiştir. Bu bağlamda elde edilen bir veri seti üzerindeki 10 metre/piksel çözünürlüklü uydu görüntülerinden 2,5 metre/piksel çözünürlükle görüntülerin elde edilmesi için Swin Transformer mimarisi tabanlı SwinIR modeli için ince ayarlama süreci işletilmiştir. Geliştirilen

sistem, veri kümesinin doğrulama alt grubu içerisinde yer alan «tarım» alt kümesi üzerinde test edilmiş ve 15.02 dB maksimum PSNR skoru, ardından 0.347 SSIM skoru metrikleri listelenmiştir. Elde edilen sonuçlara göre, incelenen sistemin veri kümesi limitleri düşünülerek değerlendirildiğinde uzaktan algılama temelli tarımsal incelemelerde verim artırıcı bir alternatif çözüm olarak değerlendirilebileceği gözlemlenmiştir.

Çıkar Beyanı: Yazarlar bu çalışmada herhangi bir çıkar çatışmasının bulunmadığını beyan etmektedir.

Teşekkürler: Bu çalışmanın tüm süreçlerinde bilgi birikimi ve rehberliği ile katkı sağlayan değerli danışmanımız Doç. Dr. Ali Furkan Kamanlı'ya en derin şükranlarımızı sunarız. Ayrıca, çalışmanın gerçekleştirilmesine olanak tanıyan ve gerekli akademik altyapıyı sağlayan Sakarya Uygulamalı Bilimler Üniversitesi'ne teşekkür ederiz.

Yazar Katkıları: Buket Bingöl bu çalışmanın kavramsal araştırmaları ve literatür taramasını, Ali Furkan Kamanlı algoritmanın kurulması ve çalışmanın düzenlenmesini, Deniz Çelik model implementasyonu çalışmalarını gerçekleştirmiştir.

REFERANSLAR

[1] J. Zhang, and Y. Tu, "SwinFR: Combining SwinIR and fast fourier for super-resolution reconstruction of remote sensing images," Digital Sensing Processing, vol. 159, p. 105026, 2025.

[2] J. Tu, G. Mei, Z. Ma, and F. Piccialli, "SWCGAN: Generative Adversarial Network Combining Swin Transformer and CNN for Remote Sensing Image Super-Resolution," IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, 2022

[3] S. Lei, Z. Shi, and W. Mo, "Transformer-Based Multistage Enhancement for Remote Sensing Image Super Resolution" IEEE Transactions on Geoscience and Remote Sensing, vol. 60, 2021

[4] A. Djerida, K. Djerriri, M. S. Karoui, and M. E. A. Larabi, "A new public Alsat-2B dataset for single-image super-resolution,", Mar. 2021.

[5] C. Dong, C. C. Loy, K. He, and X. Tang, "Image Super-Resolution Using Deep Convolutional Networks," Jul. 2015, arXiv:1501.00092.

[6] C. Ledig et al., "Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network," May. 2017, arXiv:1609.04802.

[7] Z. Lu, J. Li, H. Liu, C. Huang, L. Zhang, and T. Zeng, "Transformer for Single Image Super-Resolution," Apr. 2022, arXiv:2108.11084.

[8] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin ve B. Guo, "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows," Mar. 2021, arXiv:2103.14030.

[9] J. Liang, J. Cao, G. Sun, K. Zhang, L. V. Gool ve R. Timofte, "SwinIR: Image Restoration Using Swin Transformer," Aug. 2021, arXiv:2108.10257.

UZAKTAN ALGILAMADA SEGMENT ANYTHİNG MODEL İLE TARIM ALANLARININ TESPİTİ VE EKİLİ ALAN ANALİZİ

Buket Bingöl*
Sakarya Uygulamalı Bilimler Üniversitesi, Bilgisayar Mühendisliği, İstanbul, Türkiye

Ali Furkan Kamanlı
Sakarya Uygulamalı Bilimler Üniversitesi, Elektrik ve Elektronik Mühendisliği, Sakarya, Türkiye

Deniz Çelik
Sakarya Uygulamalı Bilimler Üniversitesi, Bilgisayar Mühendisliği, İstanbul, Türkiye
**(İletişimden sorumlu yazar) buketbingol2@gmail.com*

ÖZET

Bu çalışmada, uzaktan algılama verileri kullanılarak tarım alanlarının tespiti ve ekili alanların analizinde Segment Anything Model tabanlı bir sistem önerisi sunulmaktadır. Günümüzde yaşanan iklim değişikliği ve hızlı nüfus artışı gibi etkenler sebebiyle, tarım alanlarının izlenmesini ve verimliliğin artırılması konusunda yenilikçi fikirlere ihtiyaç duyulduğu düşünülmektedir. Bu kapsamda geliştirilen iki aşamalı sistemin ilk adımında, Segment Anything Model kullanılarak uzaktan algılama verilerindeki tarımsal alan sınırları tespit edilmiştir. Elde edilen segmentasyon maskeleri, ikinci adımda bir sınıflandırıcı mimariye aktarılmış ve bu alanların ekili olup olmadığı tespit edilmiştir. Önerilen sistemin başarımı, F1 skoru ile değerlendirilmiştir. Elde edilen sonuçlara göre, temel model olarak kullanılan Segment Anything Modelin uzaktan algılamada alan segmentasyonu konusunda kayda değer bir alternatif oluşturduğu ve sınıflandırma bölümü ile birlikte tarımsal analizlerde kullanışlı bir araç olabileceği gösterilmiştir. Çalışmanın içerdiği çıktılar ile akıllı tarım teknolojilerine katkı sağlayabileceği düşünülmektedir.

Anahtar Kelimeler: Tarım alanı tespiti, yapay zekâ, uzaktan algılama.

Giriş

Günümüzde yaşanan iklim krizi, mevsimsel değişkenlikler, nüfus artışı ve kentleşme gibi faktörler, tarım verimliliği konusunda yaşanabilecek negatif etkileri akıllara getirmektedir. Yaşanan iklim krizi, tarımın gelişimini zorlaştırabilmekte ve olumsuz etkisini göstermektedir [1]. Tarım alanlarının doğru şekilde kullanılması, mahsul verimliliğinin artması ve doğal kaynakların korunması, doğru tarım politikalarının uygulanması açısından önemli olabilecek konulardır. Ancak mevcut yöntemlerle yapılan bazı tarım analizleri altyapı yetersizliği sebebi ile sahaya gitme veya sahadan manuel görüntü alma gibi gerekliliklerle zaman alıcı, yüksek maliyetli ve bazı yaklaşımları sebebiyle ilkel olabilmektedir. Bu noktada uzaktan algılama görüntüleri kullanılarak yapılan tarım alanı tespiti ve ekili alan analizi, tarımsal incelemelerin ve doğru tarım politikalarının daha kolay uygulanabilirliğinin önünü açmaktadır. Uzaktan

algılama ile elde edilen veriler geniş alanları kapsaması, erişilebilir olması ve tekrarlanabilir veri eldesi yönleri ile daha kapsamlı ve daha doğru tarım analizlerine imkân tanımaktadır. Tarım alanlarının tespiti ve ekili alan analizleri, tarım alanlarındaki mahsul gelişim süreçlerinin daha iyi yönetilmesi ve ürün sağlığının izlenmesi açısından katkı sağlayabilmektedir [2]. Bu sayede sürdürülebilir tarımın önü daha da açılmış olacaktır.

Tarım alanlarının izlenmesi, ekosistemin bozulması ile ortaya çıkabilecek sorunlara yerinde ve doğru müdahale imkânı sağlayarak, ekonomik ve çevresel kaynakların doğru kullanılmasını sağlamaktadır. Bu sayede tarımda daha verimli ve sürdürülebilir adımlar atılabilecektir. Elde edilen sonuçlar, gerekli takiplerin yapılması ile tarım dışı faaliyetlerin kontrolsüz yayılımlarını önlemek için de kullanılabilir. Bu sayede tarım arazilerinin yanlış şekilde kullanımı önlenebilir. Bahsedilen artılar doğrultusunda hem çiftçilere hem de tarım analizi ile ilgilenen mercilere fayda sağlanabileceği ön görülmektedir.

Konu ile ilgili literatürde yapılan araştırmalarda, tarım alanlarında görüntü segmentasyonu üzerine yakın zamanda yapılan çalışmalar ilgi çekmektedir. Bu çalışmalardan biri Carraro ve arkadaşlarının (2023), Segment Anything Model (SAM)'in hassas tarım alanındaki uygulanabilirliğini inceledikleri çalışmadır. CWFID veri kümesi kullanılarak, SAM'in bitki örtüsünü arka plandan ayırma başarımı dört farklı segmentasyon modunda değerlendirilmişlerdir. Çalışmanın sonucunda tam otomatik segmentasyonun düşük doğruluk sağladığını, ancak yarı otomatik modların yüksek doğruluk ve zaman kazancı sunduğu gösterilmiştir. Çalışma, SAM'in hassas tarımda otomatik anotasyon sürecini hızlandırabilecek potansiyele sahip olduğunu göstermektedir [3].

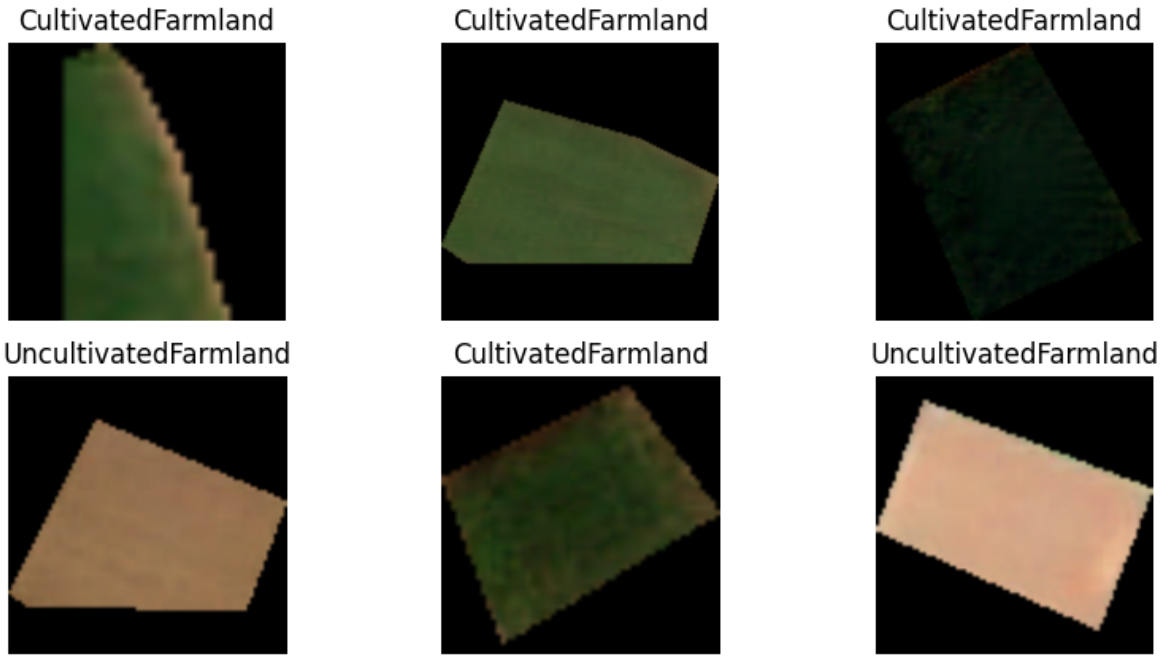
Başka bir çalışmada Tripathy ve arkadaşları (2024), SAM kullanarak, 2 metre çözünürlüklü SkySat görüntü verileri üzerinden Hindistan'ın Bihar bölgesindeki tarım alanı sınırlarını ek eğitim verisi olmadan belirlemek üzerine araştırma yapmışlardır. Farklı giriş boyutları, çok tarihli uydu görüntüleri ve kenar geliştirme tekniklerinin model performansı üzerindeki etkileri incelenmiştir. Çalışmanın sonucunda, SAM'in yaklaşık %58 oranında tarla sınırlarını doğru bir şekilde belirlediğini göstermiştir. Bu çalışmada, eğitim için gerekli veri kümelerinin sınırlı olduğu bölgelerde, tarım alanlarının sınırlarını belirleme konusunda bir potansiyel ortaya konulmaktadır. Bu bulgular, tarım analizleri ve tarımsal haritalandırma konuları için önemli bir referans oluşturmaktadır [4].

Tarım alanlarının sınırlarının belirlenmesi konusunda diğer bir çalışma Xie ve çalışma arkadaşları (2025) tarafından geliştirilen "farmland boundary Segment Anything Model (fabSAM)" adlı yöntemdir. Bu yöntem, SAM temel alınarak tarım arazisi sınırlarını tespit etmeye yönelik yenilikçi bir yaklaşım ortaya koymaktadır. FabSAM, Deeplabv3+ tabanlı bir "Prompter" ile SAM'ı birleştirerek çalışmaktadır. AI4Boundaries ve AI4SmallFarms veri kümeleri üzerinde yapılan çalışma sonuçları, fabSAM'ın tarım alanı tanımlama ve sınır belirleme konularında kayda değer iyileştirmeler sağladığını göstermektedir. Çalışma, fabSAM'ın etkinliğini ortaya koymakta ve Sentinel-2 gibi açık kaynaklı uydu görüntü veri setlerinden küresel tarım alanı bölge ve sınır haritalarının daha kolay elde edilmesine olanak tanımaktadır. Bu çalışma tarım alanı izleme ve analiz süreçleri açısından gelecekteki araştırmalar ve uygulamalar için önemli bir kaynak oluşturmaktadır [5].

Materyal ve Metot

Veri Eldesi ve Ön İşleme

Çalışma içerisinde önerilen sistemin incelenmesi için çok bantlı Sentinel-2 uydusu üzerinden levha biçiminde büyük boyutlu olarak edinilen renkli görüntüler, 388x388 boyutlu dilimlere ayrılmış ve ekili tarım alanı, ekili olmayan tarım alanı ve diğer olacak şekilde etiketlenmiştir. Görüntü kaynağının standart çözünürlüğü olan 10 metre/piksel ile elde edilen görüntüler üzerinden çıkarılan 1436 adet alan etiketi Şekil 1'de gösterildiği gibi çalışmada kullanılmak üzere kaydedilmiştir.



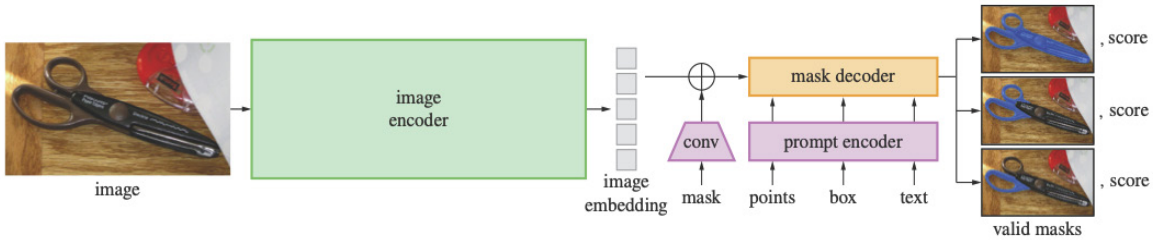
Şekil 1: Alan maskeleri ve etiketlerini içeren veri kümesi eldesi

Veri setinin modelleme adımındaki ön işlemesinde her bir görsel model girdisi olarak kullanılmak üzere 224x224 boyutlu olacak şekilde yeniden boyutlandırılmıştır. Ardından ince ayarlamada kullanılan modelin ön eğitiminde elde edilen ortalama kanal değerleri ile normalize edilmiş ve tensör çevrimi uygulanmıştır.

SAM Tabanlı Mimarinin Oluşturulması

Son yıllarda literatüre kazandırılan ve gelişimini sürdüren temel modeller (foundational models) çok büyük miktarda veri adedine sahip veri kümeleri üzerinde eğitilerek farklı görev türlerinde yüksek başarıma ulaşabilmektedir. Öğrenme aktarımı yardımıyla oluşturulan modeller, içerisinde tutulan genel amaçlı özellik setlerini içeren parametreler yardımıyla problemlerin çözümünde esnek bir alternatif oluşturmaktadır. Günümüzde Generative Pre-trained Transformer (GPT), Bidirectional Encoder Representation from Transformers (BERT), Contrastive Language-Image Pre-training (CLIP) gibi sıralanabilen birçok temel model örneği, doğal dil işleme ve bilgisayar görüsü alanlarında geleneksel modellerin üzerinde performans ile çözüm olarak kullanılabilmektedir.

Tanımlanan problemin çözümüne yönelik makine öğrenmesi mimarisi seçiminde problemin birden fazla alt adımı için farklı modeller tercih edilmiştir. Uydu görüntüleri üzerindeki alanların anlaşılabilmesi ve segmentasyonunda güncel bir temel modelin etkisinin incelenmesi için Meta tarafından geliştirilen SAM tercih edilmiş ve kullanılmıştır. SAM, genel amaçlı görüntü segmentasyonu görevleri ve geniş kapsamda kullanım amacıyla tasarlanmış, Transformer mimarisi tabanlı bir derin öğrenme temel modelidir [6]. Bu model, farklı türlerde ve geniş ölçekli görüntü veri kümeleri üzerinde eğitilmiştir. SAM, edindiği genel amaçlı parametre havuzu ve Şekil 2'de tariflenen Vision Transformer tabanlı çok katmanlı görüntü kodlayıcı (image encoder) sayesinde yeniden eğitim veya ince ayarlama (fine-tuning) ihtiyacı duymaksızın yüksek doğrulukta nesne ve alan segmentasyonu yapabilmektedir.

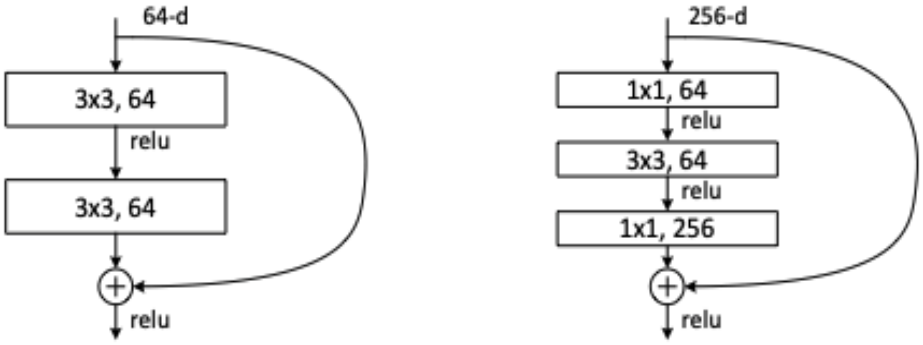


Şekil 2: Segment Anything Model Genel Görünümü [7]

Bu sayede SAM mimarisinin, uzaktan algılama görüntü verileri gibi karmaşık detayları içeren veri kümelerinde tarım alanları da dahil olmak üzere tüm ayrık bölgeler ve nesnelerin etkin biçimde tespitinde kullanılabilmesi planlanmıştır. Çalışma içeriğinde modelin ViT-L versiyonu temel alınmış ve SAM ile Şekil 3'te gösterildiği biçimde elde edilen alan sınırları, sonraki aşamada kullanılacak sınıflandırma modeli için girdi halinde kullanılmıştır.

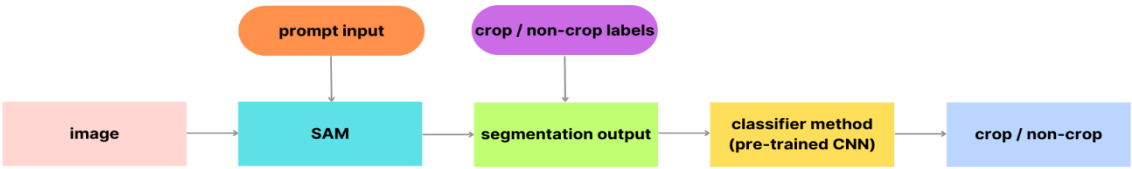
Şekil 3: Segment Anything Model Tarım Uydu Görüntülerindeki Örnek Çıktısı

Görüntü üzerinde maskeler ile elde edilen alanların sınıflandırılması için Convolutional Neural Network (CNN) yapısına sahip bir ön eğitimli görüntü sınıflandırma mimarisi olan ResNet50 modeli kullanılmıştır. ResNet50 modeli, derin sinir ağlarının eğitim sürecinde katman yapısı derinleştikçe oluşan "kaybolan gradyan" (vanishing gradients) problemi ve buna bağlı başarım kaybını çözmek amacıyla Şekil 4'te gösterilen "artık bağlantısı" (residual connection) ve "atlama bağlantısı" (skip connection) yapılarını kullanan bir sınıflayıcı mimarisidir. Bu yapıların kullanımı ile sinir ağının nispeten daha fazla sayıda derin katmanlara sahip olmasıyla ortaya çıkan bozucu etkilerin etkisi azaltılmış ve daha yüksek adetli parametre sayısı ve karmaşık öznitelik çıkarımı sayesinde literatüre kazandırıldığı dönemdeki alternatif mimarilere kıyasla daha yüksek başarım elde etmiştir. ResNet50 modeli, ImageNet veri kümesi üzerinde maliyet etkin başarımı sayesinde bu çalışmada sınıflayıcı olarak tercih edilmiş ve önerilen sistemin oluşturulmasında kullanılmıştır.



Şekil 4: Segment ResNet Mimarisi ile Gösterimi Yapılan Artık ve Atlama Bağlantıları [8]

Sistemin ilk adımında SAM ile alınan alanlar, bölgesel parçalar halinde sınıflandırıcı mimarisine girdi olarak verilmiştir. Bu alan sınırları ile eğitilmiş olan sınıflayıcı kullanılarak her bir görsel içerisindeki ilgilenilen alanların sınıf tahmini Şekil 5 ile gösterilen akış ile elde edilmiştir. Böylelikle alanların tarım niteliği bakımında işlenip işlenmediği, yani ekili bir mahsul bulunup bulunmadığının tespiti yapılmıştır.



Şekil 5: Önerilen SAM Tabanlı Sistemin Genel Görünümü

Eğitim Süreci

Önerilen sistemin sınıflandırıcı adımı için gerçekleştirilen eğitim adımında, modelin Python dilinde PyTorch kütüphanesi üzerinden ön eğitilmiş versiyonu kullanılmıştır. Modelin parametrelerinin güncellenmesinde yüksek doğruluk ve kararlı yakınsama amacıyla Adam iyileyici algoritması tercih edilmiştir. Modelin aşırı öğrenimini engellemek ve kararlığını korumak amacıyla erken durdurma (early-stopping) ve düzenlileştirme (regularization) kriterleri tanımlanmıştır.

Tartışma

Bu çalışma, derin öğrenme temel modellerinin bilgisayar görüsü problemlerinden biri olan uzaktan algılama üzerindeki başarımının anlaşılması konusunda bir referans oluşturmayı hedeflemektedir. Bu amaçla kullanılan segmentasyon temel modeli olan SAM yapısının başarımı tarım alanlarının segmente edilmesi için kullanılmış ve alan sınırları elde edilmiştir. Çalışma kapsamında önerilen sistem, hesaplama karmaşıklığı yönünden tek adımlı sistemlere kıyasla yüksek maliyetli yapısına karşın derin öğrenme çözümlerinin gerektirdiği karmaşık eğitim ve veri işleme süreçlerine bağlılığı azaltan alternatiflerin incelenmesi açısından kayda değer sonuçlar paylaşmaktadır. Bir temel modelin bilgisayarlı görü problemleri üzerindeki genel amaçlı başarımının, görece küçük ölçekli ve farklı dikeylerdeki diğer problemlerin çözümünde geliştirme karmaşıklığını azaltabilen bir alternatif olabileceği değerlendirilmiştir. Bu sayede yalnızca bir sınıflayıcı modelinin eğitimi ile uzaktan algılama gibi karmaşık problemlerin çözümünde alternatif oluşturulabilmiştir. Literatür çalışmalarıyla örneklendiği üzere, kullanılan temel modelin ince ayarlanması gibi ek geliştirmelerin sağlanmasıyla, tek adımlı genel amaçlı bir çözüm önerisinin elde edilebileceği öngörülmektedir. Bu yaklaşımın ileri çalışmalarda incelenmesi hedeflenmektedir. Bunun yanında, sınıflandırıcı adımında derin öğrenme modellerine alternatif olarak, problem özelindeki Normalised Difference Vegetation Index (NDVI) ve benzeri metriklerin kullanımı ile farklı ve yapısal olarak daha az hesaplama karmaşıklığına sahip sistem alternatiflerinin oluşturulabileceği görülmüştür.

Sonuç

Oluşturulan iki adımlı sistemin başarım ölçümünde, test için kullanılan 150 adet görüntü örneği içerisinden 1436 adet alan maskesi elde edilmiş ve sınıflandırılmıştır. Alan maskelerinin elde edilmesinde kullanılan yöntemin başarımı, oluşturulan alan maskeleri ve tahmin edilen alan maskelerinin karşılaştırmasını içeren IoU metriği ile doğrulanmıştır. Elde edilen alan sınırlarının örnek çıktıları Şekil 6 ile gösterilmiştir.

Şekil 6: SAM kullanılarak elde edilen alan sınırları

Sistemin ikinci adımı olan sınıflandırma başarımı için SAM ile elde edilen alan sınırı maskeleri ve gerçek etiketler değerlendirilmiştir. Oluşturulan sınıflayıcı algoritma ile %81,1 ortalama makro F1 skoru elde edilmiştir.

Çıkar Beyanı: Yazarlar bu çalışmada herhangi bir çıkar çatışmasının bulunmadığını beyan etmektedir.

Teşekkürler: Bu çalışmanın her aşamasında bilgi ve tecrübeleriyle yol gösteren, katkı sağlayan değerli danışmanımız Doç. Dr. Ali Furkan Kamanlı'ya en içten teşekkürlerimizi

sunarız. Ayrıca çalışmalarımız süresince gerekli akademik ortamı ve imkânları sağlayan Sakarya Uygulamalı Bilimler Üniversitesi'ne teşekkür ederiz.

Yazar Katkıları: Deniz Çelik bu çalışmanın kavramsal araştırmaları ve literatür taramasını, Ali Furkan Kamanlı algoritmanın kurulması ve çalışmanın düzenlenmesini, Buket Bingöl model implementasyonu çalışmalarını gerçekleştirmiştir.

REFERANSLAR

[1] U.S. Environmental Protection Agency, "Climate Change Impacts on Agriculture and Food Supply." epa.gov. Accessed: April 1, 2025. [Online.] Available: https://www.epa.gov/climateimpacts/climate-change-impacts-agriculture-and-food-supply

[2] M. Alizadeh, P. Ahmadi and M. Azimzadeh, "Boundary Detection in Agricultural Fields Using a Residual U-Net++ Architecture," 11th International Symposium on Telecommunications, 2024.

[3] A. Carraro, M. Sozzi, and F. Marinello, "The Segment Anything Model (SAM) for accelerating the smart farming revolution," Smart Agricultural Technology, vol. 6, p. 100367, 2023.

[4] P. Tripathy, K. Baylis, K. Wu, J. Watson, and R. Jiang, "Investigating the Segment Anything Foundation Model for Mapping Smallholder Agriculture Field Boundaries Without Training Labels," 2024, arXiv:2407.01846.

[5] Y. Xie, H. Wu, H. Tong, L. Xiao, W. Zhou, L. Li, and T. C. Wanger, "fabSAM: A Farmland Boundary Delineation Method Based on the Segment Anything Model," 2025, arXiv:2501.12487.

[6] A. Kirillov, E. Mintun, N. Ravi, H. Mao, C. Rolland, L. Gustafson, T. Xiao, S. Whitehead, A. C. Berg, W.-Y. Lo, P. Dollár, and R. Girshick, "Segment Anything," 2023, arXiv:2304.02643.

[7] S. Sen, "Segment Anything Model (SAM): Meta's Groundbreaking Segment Anything Model," Analytics Vidhya, Jul. 2024.

[8] S. R. Rath, "Residual Neural Networks – ResNets: Paper Explanation." debuggercafe.com. Accessed: Apr. 1, 2025. [Online.] Available: https://debuggercafe.com/residual-neural-networks-resnets-paper-explanation

VISUALIZATION AND EDGE INTERSECTION ANALYSIS OF POLYGONAL CYCLE GRAPHS

Ümit Sarp*
İzmir Kâtip Çelebi University, Continuing Education App. and Res. Cen., İzmir, Türkiye

Ahmet Emin
Karabük University, Dept. of Mathematics, Faculty of Science, Karabük, Türkiye

Bilal Demir
Balıkesir University, Dept. of Mathematics, Necatibey Faculty of Education, Balıkesir, Türkiye
*(Corresponding author) umit.sarp@ikcu.edu.tr

ABSTRACT

Polygonal numbers, which geometrically represent integers as nested polygons, bridge number theory and combinatorial geometry. This paper introduces *polygonal cycle graphs* $PC_m(n)$ as graph-theoretic analogs of m-gonal numbers and systematically evaluates four layout algorithms **Circular, Fruchterman-Reingold(FR), Kamada-Kawai(KK), and Spring** for minimizing edge intersections. A computational framework using Python's NetworkX library generates $PC_m(n)$ graphs, visualizes them, and quantifies crossings. Results demonstrate that force-directed algorithms (Kamada - Kawai, Fruchterman - Reingold) significantly reduce intersections compared to Circular layouts. The study provides insights into optimizing graph readability for nested cyclic structures, with applications in network visualization and mathematical modeling.

Keywords: Polygonal numbers, graph drawing, edge intersection, layout algorithms, NetworkX.

Introduction

Polygonal numbers $P_m(n)$, such as triangular $(m=3)$ and square $(m=4)$ numbers, have captivated mathematicians since antiquity. These numbers, formed by arranging points into regular polygons, encode geometric patterns into integer sequences. For instance, the n-th triangular number $P_3(n)=\frac{n(n+1)}{2}$ corresponds to a triangular lattice of points. Beyond their arithmetic properties, polygonal numbers inspire graph-theoretic representations, where points become vertices and adjacencies model geometric connections [1].

This paper uses *polygonal cycle graphs* $PC_m(n)$ as graphs derived from m-gonal numbers, where vertices correspond to points in the polygonal arrangement, and edges follow adjacency rules mirroring the geometric structure [2]. A key challenge in visualizing $PC_m(n)$ is minimizing edge intersections, which degrade readability. We address this by

evaluating four graph layout algorithms, analyzing their efficacy in reducing crossings. The computational framework combines mathematical rigor with practical visualization, advancing the interplay between number theory and graph drawing.

Background

Polygonal Numbers

Formally, the n-th m-gonal number is defined as:

$$P_m(n) = \frac{(m-2)n^2 - (m-4)n}{2}.$$

Graph Drawing Algorithms

Circular Layout [3]:

Logic and How it Works: The Circular Layout algorithm positions all the nodes of the graph at equal intervals along a circle. This is a straightforward approach often used for a quick visualization of the graph's overall structure.

Strengths: Simple and fast algorithm ($O(|V|)$ complexity). Guarantees zero edge crossings for simple cycles.

Weaknesses: High probability of edge intersections for complex and nested structures, and often does not provide a readable drawing.

Fruchterman-Reingold (FR) [4]:

Logic and How it Works: The Fruchterman-Reingold algorithm employs a physical analogy to draw graphs. Each node (vertex) in the graph behaves as if it were a charged particle repelling other nodes. Connected nodes (edges) act like springs attracting each other. The algorithm aims to find the optimal positions for nodes and edges by balancing these repulsive and attractive forces. Nodes close to each other repel, while nodes connected by an edge attract. These forces continuously update the positions of the nodes until the system reaches an equilibrium.

The force equations are as follows:

$$F_{\text{rep}}(d) = -\frac{k^2}{d}, \quad F_{\text{attr}}(d) = \frac{d^2}{k},$$

where d represents the distance between two nodes, and k is a constant scaling the optimal node spacing.

Strengths: Produces drawings with balanced edge lengths and generally aesthetically pleasing layouts.

Weaknesses: Has a high computational cost, especially for large graphs ($O(|V|^2)$ complexity).

Kamada-Kawai (KK) [5]:

Logic and How it Works: The Kamada-Kawai algorithm aims to minimize an energy function based on the graph-theoretic distances (shortest paths) between nodes. Ideally, nodes that are "close" in the graph should also be close in the drawing. The algorithm defines an energy function based on the difference between the current Euclidean distance and the ideal distance (derived from the shortest path length) between all pairs of nodes and iteratively reduces this energy.

The energy function is expressed as:

$$E = \sum_{i \neq j} \frac{1}{2} \left(\frac{\| p_i - p_j \|}{l_{ij}} - 1 \right)^2,$$

where p_i and p_j are the position vectors of nodes i and j, $\| p_i - p_j \|$ is the Euclidean distance between these nodes, and l_{ij} is the shortest path length in the graph.

Strengths: Yields good results, especially for certain types of graphs like hierarchical structures.

Weaknesses: Has a high computational complexity ($O(|V|^3)$).

Spring Layout [6]:

Logic and How it Works: The Spring Layout algorithm is a simplified version of Fruchterman-Reingold. It considers only the edges as linear springs and uses Hooke's law. A pulling force is applied between each pair of connected nodes, proportional to the distance between them. The goal is to find a node configuration where all springs reach an ideal "rest length."

The force equation is:

$$F_{\text{spring}} = -k \cdot (d - L),$$

where k is the spring constant, d is the distance between two nodes, and L is the rest length of the spring.

Strengths: Faster than Fruchterman-Reingold.

Weaknesses: May not produce results as precise and good as Fruchterman-Reingold for complex graphs. Its computational complexity is $O(|V|^2)$.

Polygonal Cycle Graphs: Construction

For m-gonal number $P_m(n)$, the polygonal cycle graph $PC_m(n)$ is constructed through:

Vertex Set: $V = \{1, \ldots, P_m(n)\}$ where $P_m(n) = \frac{(m-2)n^2 - (m-4)n}{2}$

Edge Set:

1. Linear edges: $i \sim i+1$ for non-polygonal vertices

2. Polygonal connections: $P_m(k) \sim P_m(k+1)$ for $1 \le k < n$
3. Radial links: $P_m(k-1)+1 \sim P_m(k)+1$ for $1 \le k < n$

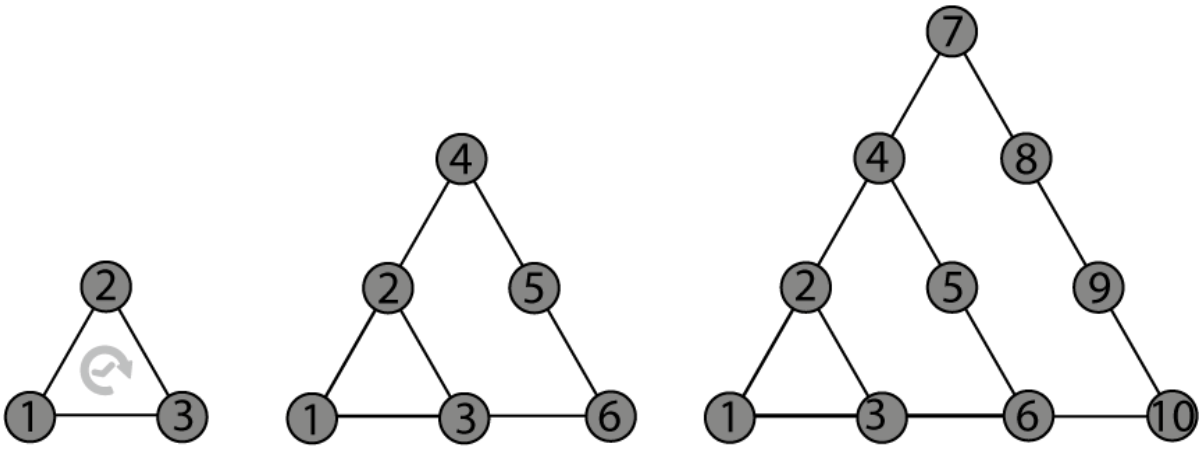


Figure 1: Adjacency rules for $P_3(n)$ with $n = 2,3,4$.

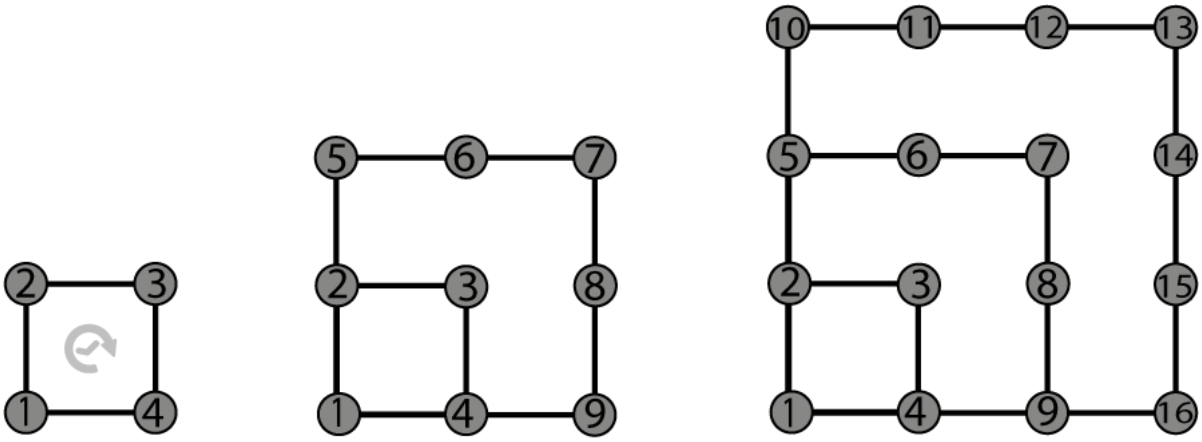


Figure 2: Adjacency rules for $P_4(n)$ with $n = 2,3,4$.

Edge Intersection Detection

The counter-clockwise (CCW) method [7] determines edge intersections. Consider two distinct edges $e_1 = (u_1, v_1) \in E$ and $e_2 = (u_2, v_2) \in E$ in a graph $G = (V, E)$ with a planar embedding, where the planar positions of the vertices are given by a function $pos : V \to \mathbb{R}^2$ such that $pos(u) = (x_u, y_u)$ for a vertex $u \in V$. For a "true" intersection, these two edges must have four distinct vertices, i.e., $|\{u_1, v_1, u_2, v_2\}| = 4$.

The intersection condition is based on the counter-clockwise orientation (ccw) of triplets of points. For three points $A = (x_A, y_A)$, $B = (x_B, y_B)$, and $C = (x_C, y_C)$, the ccw value is defined as:

$$ccw(A, B, C) = (y_C - y_A)(x_B - x_A) - (y_B - y_A)(x_C - x_A)$$

A positive value indicates that C is to the left of the directed line AB, a negative value indicates it's to the right, and zero indicates collinearity.

For the edges (u_1, v_1) and (u_2, v_2), let's define the orientations:

$$o_1 = \mathrm{CCW}(pos(u_1), pos(u_2), pos(v_2))$$

$$o_2 = \mathrm{CCW}(pos(v_1), pos(u_2), pos(v_2))$$

$$o_3 = \mathrm{CCW}(pos(u_1), pos(v_1), pos(u_2))$$

$$o_4 = \mathrm{CCW}(pos(u_1), pos(v_1), pos(v_2))$$

Edges (u_1, v_1) and (u_2, v_2) intersect if and only if $o_1 \neq o_2$ and $o_3 \neq o_4$. This condition is equivalent to stating that the endpoints of one edge lie on opposite sides of the line formed by the other edge.

The total number of edge intersections in the graph G can be determined by iterating through all distinct pairs of edges, checking if they have four distinct vertices and if the above intersection condition ($o_1 \neq o_2$ and $o_3 \neq o_4$) is satisfied. The total count is the sum of all such intersecting pairs.

Visual examples of edge intersections generated by different layout algorithms for $PC_3(n)$ and $PC_4(n)$ are presented. Figure 3 shows layouts for triangular cycle graphs ($m = 3$) and Figure 4 shows layouts for square cycle graphs ($m = 4$), both for n ranging from 2 to 7. These figures illustrate how layout choice significantly impacts the number of visual crossings, affecting the overall readability of the graph structure.

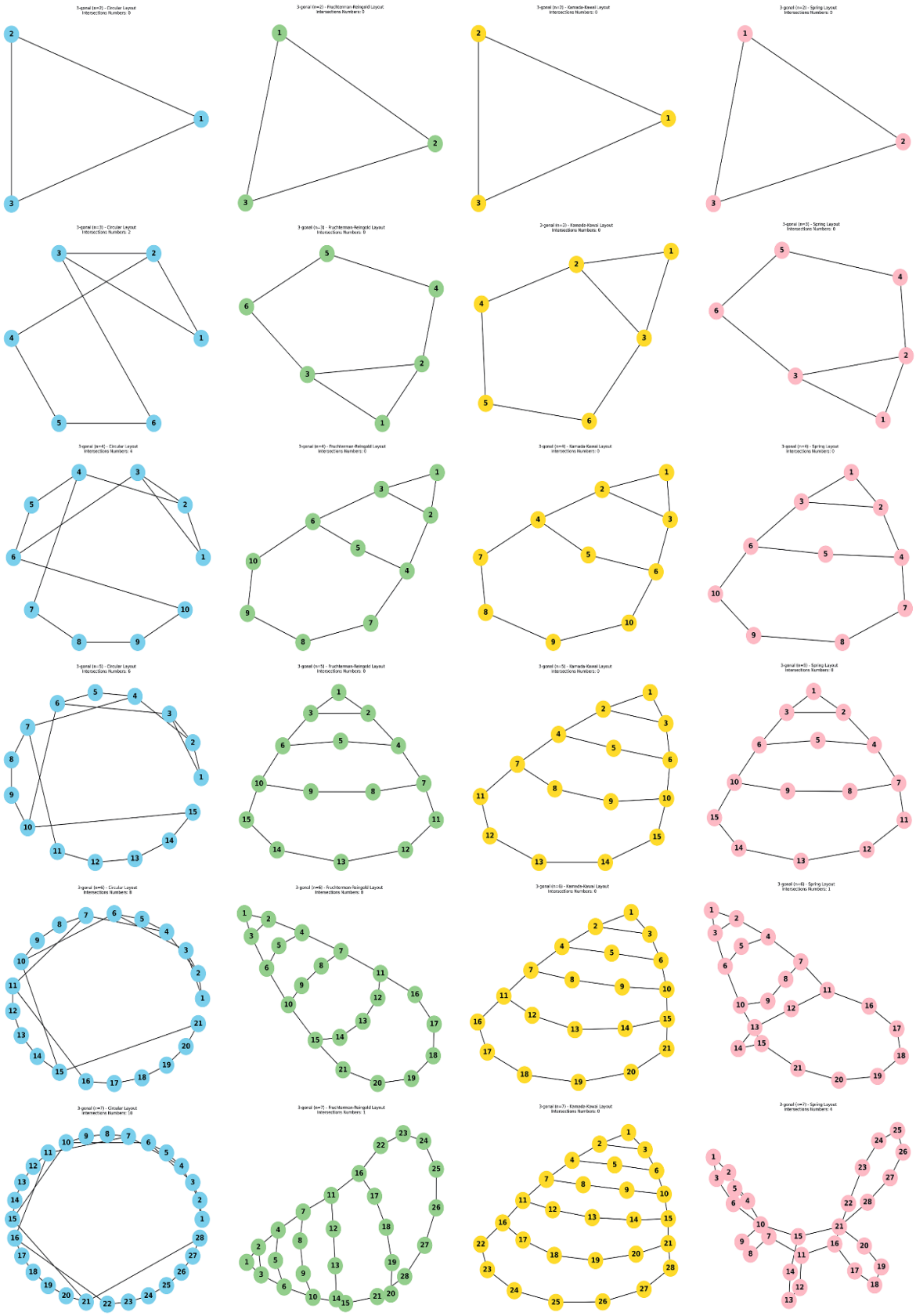


Figure 3: Layouts for $PC_3(3)$ with $n = 2,3,\ldots,7$

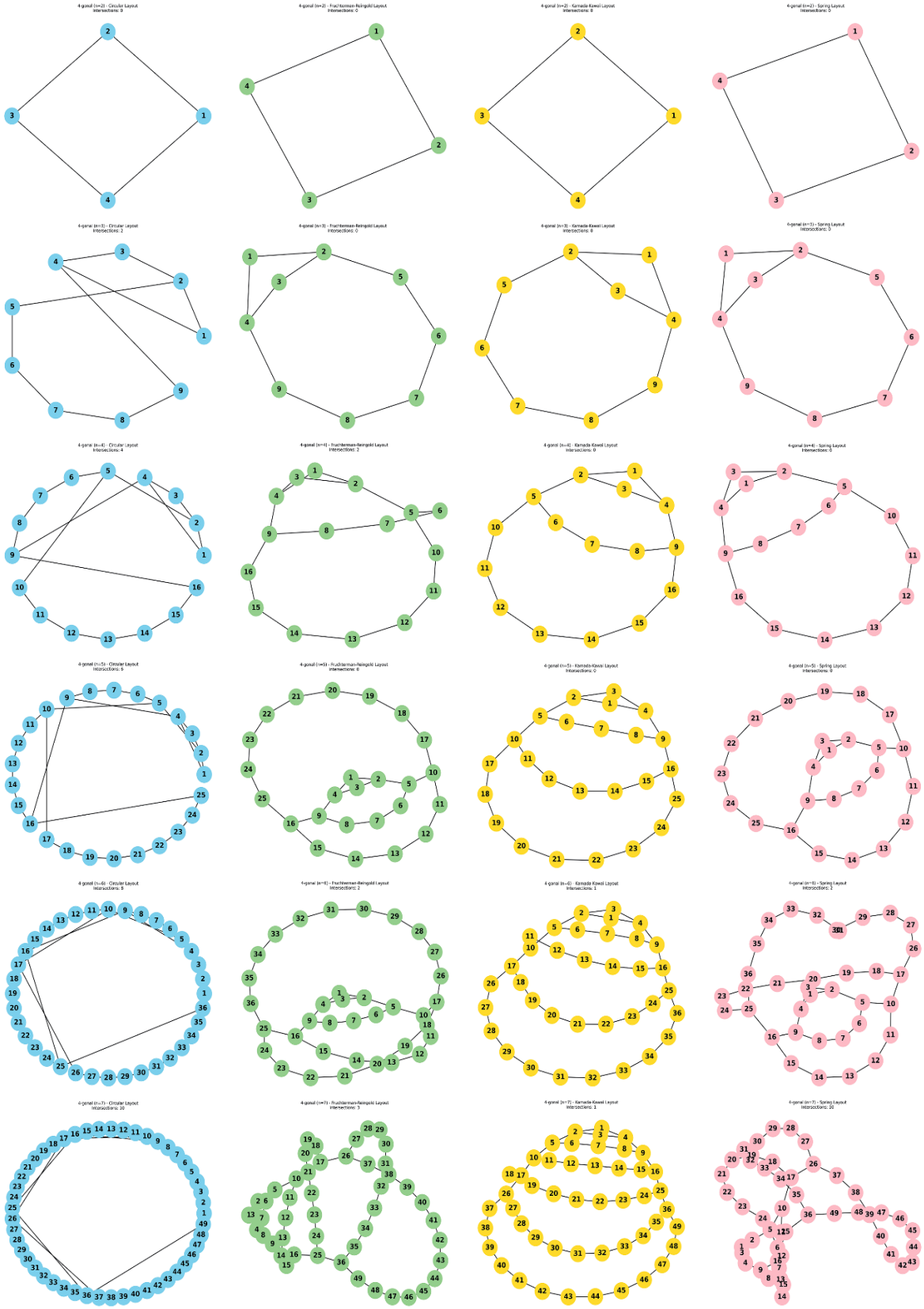


Figure 4: Layouts for $PC_4(n)$ with $n = 2,3,\ldots,7$

Algorithm

Algorithm: Polygonal Cycle Graph Construction & Visualization

1. Input: m (polygon sides), $n_{\max}$ (max layers)
2. for $n \leftarrow 1$ to $n_{\max}$ do
3. Compute $P_m(n) = \frac{(m-2)n^2 - (m-4)n}{2}$
4. Initialize graph G with $P_m(n)$ nodes
5. Add linear edges between consecutive non-polygonal nodes
6. Connect polygonal number nodes between successive layers
7. Create radial links between layer starting points layout $\in$ [Circular, FR, KK, Spring]
8. Compute node positions using specified algorithm
9. Detect edge intersections via CCW method
10. Generate visualization with intersection count
11. end for

Results and Analysis

This section presents the quantitative results obtained from applying the four selected graph layout algorithms Fruchterman-Reingold (FR), Spring, Kamada-Kawai (KK), and Circular to the constructed polygonal cycle graphs $PC_m(n)$, specifically for the triangular ($m = 3$) and square ($m = 4$) cases. We focus on the number of edge intersections as the primary metric for comparing the effectiveness of these algorithms in producing readable visualizations. The intersection counts, determined using the CCW method for increasing layer numbers ($n = 2$ to $n = 7$), are presented in the following tables (Table 1 and Table 2) The subsequent analysis highlights the performance differences and trends observed across the layouts.

Table 1: Edge intersection counts for triangular cycle graphs ($m = 3$) across layouts.

Layer (n)	Circular	FR	KK	Spring
2	0	0	0	0
3	2	0	0	0
4	4	0	0	0
5	6	0	0	0
6	8	0	0	1
7	10	1	0	4

Table 2: Edge intersection counts for square cycle graphs ($m = 4$) across layouts.

Layer (n)	Circular	FR	KK	Spring
2	0	0	0	0
3	2	0	0	0
4	4	2	0	0
5	6	0	0	0
6	8	2	1	2
7	10	3	1	10

Key findings:

•The FR layout significantly reduces edge intersections compared to the Circular layout, with observed reductions ranging from 50% to 100% for $n \geq 4$.

•The KK layout consistently results in fewer or equal edge intersections compared to the Spring layout, showing substantially better performance (50% - 100% fewer intersections when Spring layout had intersections in these tests).

•Edge intersection counts generally increase with the number of layers (n). This increase appears linear for the Circular layout, very slow for FR and KK layouts, and potentially faster than linear for the Spring layout, especially for $m = 4$.

Conclusion

This study introduced polygonal cycle graphs ($PC_m(n)$), a novel class of graphs derived from m-gonal numbers, bridging number theory and combinatorial geometry. Our primary objective was to evaluate the effectiveness of different graph layout algorithms in maximizing the readability of visualizations for this specific graph structure. In this context, we compared four common layout algorithms Circular, Fruchterman - Reingold(FR), Kamada-Kawai(KK), and Spring based on the number of edge intersections in the drawings they produce.

Using Python's NetworkX library, we systematically generated $PC_m(n)$ graphs (for $m = 3$ and $m = 4$, with layer numbers n ranging from 2 to 7), computed node positions by applying each algorithm, and counted the resulting edge intersections using the CCW method.

The obtained results (summarized in Table 1 and 2) demonstrated that force-directed algorithms (FR and KK) possess a distinct advantage over the Circular layout in visualizing $PC_m(n)$ graphs. Specifically, the Kamada-Kawai (KK) algorithm exhibited the best performance, producing the fewest edge intersections in most cases examined. The Fruchterman-Reingold (FR) algorithm also generally resulted in very few intersections and produced significantly better drawings than the Circular layout. While the Spring layout yielded good results in some instances, it could show noticeable increases in intersection counts, particularly for $m = 4$ and higher n values. The Circular layout, although attractive due to its simplicity and speed ($O(|V|)$ complexity), led to a high number of edge intersections due to the nested structure of $PC_m(n)$ graphs, producing unreadable drawings except in the simplest cases.

This work underscores the importance of layout algorithm selection when visualizing graphs with layered and cyclical substructures, such as $PC_m(n)$. Force-directed algorithms, despite their higher computational costs ($O(|V|^2)$ to $O(|V|^3)$), offer significant advantages in enhancing visual clarity. Our findings provide practical guidance for researchers and practitioners working in network visualization, mathematical modeling, and related fields.

Future work could involve investigating the performance of hybrid layout approaches that combine the strengths of different algorithms on $PC_m(n)$ graphs. Furthermore, extending the analysis to larger n values and different m values, as well as utilizing techniques like GPU acceleration to enhance algorithm performance, are potential research directions.

Declaration of Interest: The authors declare that there is no conflict of interest.

Acknowledgements: The authors acknowledge the support of İzmir Kâtip Çelebi University Continuing Education Application and Research Center.

Author Contributions: All authors contributed equally to the preparation of this article.

REFERENCES

[1] E. Deza and M. Deza, *Figurate Numbers*. World Scientific, 2012.

[2] U. Sarp, B. Demir, A. Emin, and İ. N. Cangül, "*Introducing a new family of graphs associated to the polygonal numbers: Polygonal cycle graphs*," to appear, 2025.

[3] I. Herman, G. Melançon, and M. S. Marshall, "*Graph visualization and navigation in information visualization: A survey*," IEEE Transactions on Visualization and Computer Graphics, vol. 6, no. 1, pp. 24–43, Jan.–Mar. 2000.

[4] T. M. Fruchterman and E. M. Reingold, "*Graph Drawing by Force-Directed Placement*," Software—Practice & Experience, vol. 21, no. 11, pp. 1129–1164, 1991.

[5] T. Kamada and S. Kawai, "An Algorithm for Drawing General Undirected Graphs," Information Processing Letters, vol. 31, no. 1, pp. 7–15, 1989.

[6] P. Eades, "*A heuristic for graph drawing*," Congressus Numerantium, vol. 42, pp. 149–160, 1984.

[7] J. O'Rourke, *Computational Geometry in C*. Cambridge, U.K.: Cambridge University Press, 1998.

Appendix

PyCode: Polygonal Cycle Graph Construction & Visualization

```
import networkx as nx
import matplotlib.pyplot as plt
from itertools import combinations
import os
def count_intersections(G, pos):
  """
  Counts the number of edge intersections in a graph drawing.
  """
  edges = list(G.edges())
  intersections = 0
  for (u1, v1), (u2, v2) in combinations(edges, 2):
    if len({u1, v1, u2, v2}) < 4:
      continue
    x1, y1 = pos[u1]
    x2, y2 = pos[v1]
    x3, y3 = pos[u2]
    x4, y4 = pos[v2]
    def ccw(A, B, C):
      return (C[1] - A[1]) * (B[0] - A[0]) > (B[1] - A[1]) * (C[0] - A[0])
    intersect1 = ccw((x1, y1), (x3, y3), (x4, y4)) != ccw((x2, y2), (x3, y3), (x4, y4))
    intersect2 = ccw((x1, y1), (x2, y2), (x3, y3)) != ccw((x1, y1), (x2, y2), (x4, y4))
    if intersect1 and intersect2:
      intersections += 1
  return intersections
def polygonal_cycle_graph(m, n):
  """
  Generates the polygonal cycle graph for the nth m-gonal number.
  """
  def P_m(n):
    return n * ((m - 2) * n - m + 4) // 2
  V = list(range(1, P_m(n) + 1))
  G = nx.Graph()
  G.add_nodes_from(V)
  if n == 1:
    return G
  for i in V:
    if i > 1 and i < P_m(n) and not any(P_m(j) == i for j in range(2, n + 1)):
      G.add_edge(i, i + 1)
  for i in range(1, n):
    G.add_edge(P_m(i), P_m(i + 1))
  for i in range(1, n):
    if P_m(i) + 1 < P_m(n):
      G.add_edge(P_m(i - 1) + 1, P_m(i) + 1)
  return G
def visualize_graphs(m, n_values):
  # Create a directory with dynamic name PC_{m}
  folder_name = f"PC_{m}"
  os.makedirs(folder_name, exist_ok=True)
  for n in n_values:
    G = polygonal_cycle_graph(m, n)
    if G.number_of_nodes() == 0:
      continue
    # Fruchterman-Reingold Layout
    plt.figure(figsize=(10, 8))
    pos_fr = nx.spring_layout(G, k=0.3, iterations=200, seed=42)
```

```
    intersections = count_intersections(G, pos_fr)
    nx.draw(G, pos_fr, with_labels=True, node_size=2200, node_color="lightgreen",
        font_size=20, font_color="black", font_weight="bold", width=2)
    title = f"{m}-gonal (n={n}) - Fruchterman-Reingold Layout\nIntersections: {intersections}"
    plt.title(title)
    plt.savefig(f"{folder_name}/{m}_gonal_n{n}_fruchterman_reingold.png", dpi=300, bbox_
inches='tight')
    plt.show()
    # Spring Layout
    plt.figure(figsize=(10, 8))
    pos_spring = nx.spring_layout(G, seed=42)
    intersections = count_intersections(G, pos_spring)
    nx.draw(G, pos_spring, with_labels=True, node_size=2200, node_color="pink",
        font_size=20, font_color="black", font_weight="bold", width=2)
    title = f"{m}-gonal (n={n}) - Spring Layout\nIntersections: {intersections}"
    plt.title(title)
    plt.savefig(f"{folder_name}/{m}_gonal_n{n}_spring.png", dpi=300, bbox_inches='tight')
    plt.show()
    # Kamada-Kawai Layout
    plt.figure(figsize=(10, 8))
    try:
      pos_kk = nx.kamada_kawai_layout(G)
      intersections = count_intersections(G, pos_kk)
      nx.draw(G, pos_kk, with_labels=True, node_size=2200, node_color="gold",
          font_size=20, font_color="black", font_weight="bold", width=2)
      title = f"{m}-gonal (n={n}) - Kamada-Kawai Layout\nIntersections: {intersections}" plt.title(title)
      plt.savefig(f"{folder_name}/{m}_gonal_n{n}_kamada_kawai.png", dpi=300, box_inches='tight')
      plt.show()
    except nx.NetworkXException:
      print(f"Kamada-Kawai unsuitable for n={n}")
    # Circular Layout
    plt.figure(figsize=(10, 8))
    pos_circular = nx.circular_layout(G)
    intersections = count_intersections(G, pos_circular)
    nx.draw(G, pos_circular, with_labels=True, node_size=2200, node_color="skyblue",
        font_size=20, font_color="black", font_weight="bold", width=2)
    title = f"{m}-gonal (n={n}) - Circular Layout\nIntersections: {intersections}"
    plt.title(title)
    plt.savefig(f"{folder_name}/{m}_gonal_n{n}_circular.png", dpi=300, bbox_inches='tight')
    plt.show()

# Execution (Change m value to test different polygonal cycles)
m = 3 # Changing this value updates folder name (e.g., m=5: PC_5)
n_values = range(1, 8)
visualize_graphs(m, n_values)
```

WILLIAM SHAKESBLAKE 2.0: FINE-TUNED LANGUAGE MODELS FOR HYBRID POETIC STYLE GENERATION

Doga Nalci*
Faculty of Engineering, Computer Engineering, Gebze Technical University, Kocaeli, Türkiye

Zehra Bilici
Faculty of Engineering, Computer Engineering, Gebze Technical University, Kocaeli, Türkiye

Saliha Yesilyurt
Faculty of Engineering, Computer Engineering, Gebze Technical University, Kocaeli, Turkey and Sivas Cumhuriyet University, Sivas, Türkiye
*(Corresponding author) doganalci@gtu.edu.tr

ABSTRACT

This study investigates the application of fine-tuned large language models for generating stylistically hybrid poetry inspired by William Shakespeare and William Blake. A dataset comprising over 400 sonnets and poems from both authors, retrieved via the PoetryDB API, was constructed for model training. Visual inputs were translated into descriptive prompts using the BLIP image captioning model. Two configurations were evaluated: a baseline GPT-3.5-Turbo model and a custom fine-tuned variant. Four distinct images served as prompts, generating eight poems in total. Outputs were assessed by a panel of English literature graduates, who rated each poem's stylistic alignment (on a Shakespeare–Blake continuum) and semantic relevance to the associated image. The fine-tuned model demonstrated superior performance, producing more semantically coherent and stylistically integrated poems. The frequent emergence of lexical markers such as "thy," "Jerusalem," "thee," and "Lamb of God" supported evidence of effective stylistic fusion. These results highlight that fine-tuning large language models can effectively balance classical poetic forms, thematic elements, and visual prompt relevance, enabling the generation of hybrid literary styles.

Keywords: Natural language processing, deep learning, generative artificial intelligence.

Introduction

The field of natural language generation (NLG) has seen transformative growth in recent years, particularly with the advent of large language models (LLMs) based on the transformer architecture [1,2]. These models have demonstrated an extraordinary capacity for producing fluent, syntactically sound, and semantically coherent text. Beyond traditional tasks such as summarization or translation, contemporary research has begun to explore more creative domains—most notably, poetry generation [3,4]. As machines are

increasingly tasked with emulating human artistic expression, the question arises: Can they meaningfully capture the style of canonical poets and respond to visual stimuli with literary nuance?

Prior work has taken an initial step toward this direction by comparing Markov Chain Models (MCM) and pre-trained LLMs in generating poetry using a custom dataset composed of poems by William Blake and William Shakespeare [5]. Although both approaches were able to reproduce stylistic tokens and vocabulary from the source texts, expert evaluation revealed that the generated poems—particularly those produced by MCM—tended to lack broader contextual and thematic coherence. On the other hand, the pre-trained LLM exhibited a more balanced distribution of stylistic markers and a deeper semantic consistency, yet still required further refinement to adequately blend the poetic styles of the two authors.

Building upon the foundational insights from previous work, this study introduces a fine-tuned multimodal language generation pipeline, William ShakesBlake 2.0, designed to synthesize a hybrid poetic voice in response to visual prompts. At its core, this system integrates a structured, curated poetic corpus, visual analysis via image captioning, and fine-tuning of the GPT-3.5-Turbo model to improve stylistic emulation. While the earlier study laid the foundation by demonstrating feasibility, the present work aims to enhance stylistic fidelity, visual relevance, and literary fluency through both architectural and procedural improvements.

The motivation behind this work stems from the unique literary contrast between Blake and Shakespeare. Blake's visionary symbolism and spiritual allegory stand in stark contrast to Shakespeare's rhetorical precision and dramatic structure. In this study, we place the works of Blake and Shakespeare side by side—first in the dataset, then in the generated poems—to examine whether a fine-tuned language model can learn their styles and combine them into a unified poetic voice. Furthermore, we investigate how visual cues, when paired with stylistically charged prompts, can guide the model toward generating content that is not only coherent but aesthetically situated within a specific literary tradition.

Several contributions distinguish this study from previous efforts:

- A restructured poetic corpus, programmatically compiled and manually verified, that deliberately alternates between works by Blake and Shakespeare to emphasize stylistic transitions.

- The incorporation of multimodal conditioning, using captions generated by the BLIP image-captioning model [6] to provide semantic grounding for the poems.

- Targeted fine-tuning on GPT-3.5-Turbo using OpenAI's fine-tuning framework, thereby producing a model that retains its foundational linguistic competence while adapting to a domain-specific aesthetic [7].

- A structured expert evaluation, engaging scholars of philology to assess generated poems in terms of stylistic attribution and image relevance.

Related work in the domain of literary generation has primarily focused on author

imitation using statistical methods [8] or transformer-based models fine-tuned on singular stylistic corpora [9,10]. Earlier efforts, such as those by Liu et al. (2018) and Yi et al. (2020), have highlighted the challenges of ensuring stylistic variety and visual alignment in hybrid or multimodal poetry generation [11,12]. More recent approaches have leveraged models like GPT-2 and GPT-3 in creative writing applications, yet often rely on prompt engineering alone without fine-tuning, which can limit the degree of stylistic internalization [13]. Our work extends this line of inquiry by applying both multimodal prompting and corpus-level fine-tuning, with a comparative evaluation across model generations.

In summary, this paper presents a comprehensive pipeline for multimodal poetic generation with a focus on stylistic fusion. By leveraging visual input, carefully structured training data, and fine-tuning of a powerful language model, this study seeks to advance the state-of-the-art in of what AI-generated literature can achieve—not merely reproducing poetic form, but synthesizing a new, hybrid literary voice inspired by the giants of English verse.

Material and Methods

Dataset Construction

The dataset used for fine-tuning was compiled by retrieving poetic works from the open-access database PoetryDB, which provides structured access to historical and contemporary poems through its public API [14]. Specifically, all available sonnets and free-form poems authored by William Shakespeare and William Blake were programmatically collected using automated scripts.

To ensure stylistic diversity and balance, the dataset was restructured following collection. All retrieved poems were sorted and manually curated to maintain authorial integrity and remove duplicates or inconsistencies. The final corpus was then ordered such that all of William Shakespeare's works appeared first, immediately followed by all of William Blake's. This specific ordering where a Shakespearean piece is followed by a Blakean counterpart throughout the dataset was implemented intentionally to expose the model to both poetic styles in sequential proximity, to help the model learn and internalize the stylistic distinctions and transitions between the two authors.

The alternation of contrasting literary voices within the dataset helps the model encode not only the thematic and lexical characteristics of each poet, but also their differences in rhythm, imagery, and philosophical tone. This deliberate sequence became the foundational structure upon which the fine-tuning process was built.

Model

The poetry generation pipeline involved two model configurations: GPT-3.5-Turbo as baseline without finetuning and GPT-3.5-Turbo with fine-tuned on our hybrid poetic corpus [7,16]. This caption was then embedded in a standardized prompt: "The image shows: '[caption]'. Each image was first processed using the "blip-image-captioning-base" model to generate a natural language caption describing its content [6,15]. Write a poem that blends the styles of William Blake and William Shakespeare." This prompt was submitted

to both models under identical conditions (temperature = 0.9, max_tokens = 300) via the OpenAI ChatCompletion API. Outputs from both models were stored with metadata including image ID, model version, and timestamp to ensure traceability during comparative analysis. This controlled setup enabled a direct comparison of stylistic fidelity between the base GPT-3.5-Turbo model and its fine-tuned counterpart, using only one changing variable: the model's internal weights [14]. While the GPT-2 model used in earlier experiments contained approximately 1.7 billion parameters, GPT-3.5-Turbo is significantly larger, with an estimated 175 billion parameters. This substantial increase in model capacity contributed to markedly improved semantic coherence and narrative flow in the generated poems, highlighting the strong correlation between parameter scale and the model's ability to produce contextually consistent and linguistically fluent outputs.

To improve stylistic coherence in generated poems, the GPT-3.5-Turbo model was fine-tuned using a custom dataset of 406 poems sourced from William Shakespeare and William Blake. The dataset was formatted according to OpenAI's chat-completion fine-tuning requirements, which included structuring each training sample with system, user, and assistant roles in JSONL format [7]. The fine-tuning process was initiated via authenticated API calls from a Google Colab environment using the official openai Python SDK [17]. Despite Colab being used for orchestration, all training occurred on OpenAI's secure cloud infrastructure. The total training time was approximately 1 hour. After completion, the model was registered on OpenAI's servers under a custom ID and became accessible through the same ChatCompletion endpoint as standard models. It retains full compatibility with existing APIs but outputs text that reflects the stylistic patterns encoded in the hybrid training data.

Experiments

Evaluation Method

To assess the poetic outputs generated by the two language models, a human evaluation was conducted using a structured online survey. The evaluation aimed to capture both stylistic attribution and semantic alignment of the generated poems with the original visual prompts.

Participants

A total of 20 participants completed the survey, each providing independent evaluations of the generated poems. Among them, 13 participants were specialists in the field of linguistics or literature—either academics or advanced-level students—while 7 participants came from other academic disciplines, contributing diverse perspectives on style and meaning.

Regarding educational background, the sample included 3 PhD holders, 2 master's degree holders, and 15 bachelor's degree. The inclusion of both expert and non-expert evaluators allowed for a more comprehensive assessment of the poems' literary qualities and accessibility.

Evaluation Materials

The survey included 8 poems generated by two models—GPT-3.5-Turbo (baseline) and the fine-tuned GPT-3.5-Turbo—using 4 different image prompts. The selected visual scenes included the following (survey_result). A waterfall in a forested area, People riding bicycles down a city Street, A wooden bridge among trees in a jungle, A boat floating on the sea during sunset. Each image was processed using the pretrained BLIP model to produce concise English-language captions that describe the visual content. The generated captions were as follows: "A river with a waterfall in the middle", "A group of people riding bikes down a street", "A wooden bridge in the middle of a jungle", "A boat is in the water at sunset".

Each question in the survey presented: The original image, The generated poem, The image caption used as a prompt, Two scoring criteria (on separate scales). Visual and poetic content were presented side by side to facilitate focused comparison and evaluation. Participants assessed each poem using two independent rating scales.

The evaluation process was conducted through a structured survey in which participants assessed the poems generated by three language models—GPT-2, GPT-3.5-Turbo, and fine-tuned GPT-3.5-Turbo—based on two criteria: stylistic attribution and image relevance. As illustrated in Table 1, stylistic attribution was rated on a 10-point Likert scale where 1 indicated strong resemblance to William Shakespeare, 10 to William Blake, and 5 represented a stylistic midpoint reflecting a balanced influence.

In parallel, participants evaluated the semantic alignment between each poem and the corresponding image prompt on a 3-point scale, where 1 denoted weak or unrelated content, 2 denoted moderate alignment, and 3 indicated strong relevance. These scores enabled a comparative analysis of both stylistic fidelity and prompt-based coherence across models, highlighting the effectiveness of fine-tuning in enhancing both literary style and visual contextualization.

Table 1: Survey Descriptions Used in the Evaluation of Generated Poems

Question	Explanation
Which poet's style does this poem most closely resemble? (Scale 1 to 10)	1: Strong resemblance to William Shakespeare 5: Stylistic midpoint / hybrid balance 10: Strong resemblance to William Blake
How well does this poem relate to the image prompt? (Scale 1 to 3)	1 = Weak or unrelated 2 = Moderately related 3 = Strongly aligned with the image content

Survey Platform

The survey was hosted on Microsoft Forms, which enabled secure online distribution and standardized response collection. Participants were informed that the poems were machine-generated and blinded to model source to avoid bias (Form_Link).

Results

Results and Interpretation

The o assess the evolution of stylistic and semantic quality across model generations, we compared output poems from three different models. GPT-2 is the previous baseline, GPT-3.5-Turbo zero-shot and Fine-tuned GPT-3.5-Turbo is custom-trained on a blended with our dataset.

The earliest experiments using GPT-2 produced outputs that, while stylistically suggestive of William Blake and William Shakespeare, suffered from limited coherence and narrative depth. The poems contained fragmented phrasing, vague metaphors, and lacked overall semantic flow. However, a close inspection revealed sporadic use of stylistic markers associated with both poets, such as Blakean imagery ("eternal fire," "invisible light") and Shakespearean diction ("thou," "thy," "alas"). Although readers might find these outputs confusing at first glance, the surface-level stylistic emulation was still detectable.

In contrast, the outputs generated by GPT-3.5-Turbo demonstrated improved semantic structure and fluency. The poems were logically cohesive and grammatically refined. However, the model exhibited a tendency to remain neutral in style and failed to adequately blend or emulate either Blake's or Shakespeare's voice. In some cases, the model even referenced the authors directly by name ("as Blake once wrote," "like Shakespeare's sonnets"), which suggested a lack of embedded stylistic ABSTRACTion and indicated that the model did not internalize their poetic characteristics as intended.

Figure 1: Image For Generating Poems with GPT Models

The most promising results came from the fine-tuned GPT-3.5-Turbo model. Poems generated by this version not only maintained semantic coherence but also exhibited clear stylistic features derived from both Blake and Shakespeare. Several outputs included lexical choices such as "Jerusalem," "Lamb of God," "king of England"—common motifs in Blake's prophetic works—alongside Shakespearean syntax and vocabulary including "wherefore," "alas," "thee," "thine," and "hath." These blended outputs reflected both form and substance, demonstrating that the fine-tuned model had internalized elements of both poets' voices and could creatively recombine them in response to visual prompts.

Table 2: Poem Generated from Figure-1 With GPT-2, GPT-3.5-Turbo and GPT-3.5-Turbo Finetuned

Model	Generated Poems
GPT-2	these the cold sorrowful days of sorrow, When, without your skill I might I in your care? Jill these days shoulhd have left you on You Mitgy ade? And you Through J have you Who do not destre my love, though thou hast thee in my love,
GPT-3.5-Turbo	Like a scene from Blake's divine art, A vision of joy, straight from the heart, Innocent souls in nature's embrace, A moment of peace, a sacred space. Shakespeare himself would surely smile, At this happy family, free from guile, For in their love and laughter sweet, They find true joy, their bond complete.
GPT-3.5-Turbo Finetuned (ours)	That you were once unkind befriends me now, And for that sorrow, which I then did feel, Needs must I under my transgression bow, Unless my nerves were brass or hammered steel. For if you were by my unkindness shaken, As I by yours, you've passed a hell of time; And I, a tyrant, have no leisure taken To weigh how once I suffered in your crime. O! that our night of woe might have remembered My deepest sense, how hard true sorrow hits, And soon to you, as you to me, then tendered The humble salve, which wounded bosoms fits! But that your trespass now becomes a fee; Mine ransoms yours, and yours must ransom me.

In summary, the transition from GPT-2 to GPT-3.5-Turbo to fine-tuned GPT-3.5-Turbo illustrates a clear progression in both stylistic fidelity and poetic structure. The fine-tuning process proved crucial in enabling the model to generate outputs that are not only readable but also stylistically nuanced and textually rich, aligning closely with the hybrid vision of "William ShakesBlake." This progression is exemplified through Figure 1 and Table 2: while Figure 1 displays the visual prompt used for poetic generation, Table 2 presents the corresponding poems produced by each model, enabling direct comparison of style, structure, and relevance.

A survey was hosted on Microsoft Forms, which enabled secure online distribution and standardized response collection. Participants were informed that the poems were machine-generated and blinded to model source to avoid bias [18].

Survey Results

As outlined in Table 1, participants evaluated each AIgenerated poem based on two main criteria: stylistic attribution (from Shakespeare to Blake) and image-poem coherence. The participants were primarily selected from academic backgrounds in English Language and Literature, Translation Studies, and English Language Teaching. In order to ensure a balanced evaluation, individuals without a literary background were also included in the survey. The results presented in Table 3 summarize the participants' evaluations regarding the coherence between the visual prompt and the generated poem. This allowed us to measure the effectiveness of the model's ability to not only imitate literary styles but also to generate contextually appropriate poetic responses based on visual prompts. The image-poem coherence scores provided further insight into how well the AI system could integrate multimodal stimuli into cohesive literary outputs.

Table 3: Visual-Poem Coherence Table (survey results from evaluators)

	GPT-3.5-Turbo			**GPT-3.5-Turbo Finetuned**		
Image #	(3)	(2)	(1)	(3)	(2)	(1)
1	7	3	10	12	3	5
2	3	7	10	12	2	6
3	8	4	8	14	3	3
4	9	2	9	15	1	4

The comparative results presented in Table 3 demonstrate a clear improvement in visual-poem coherence following the fine-tuning process. For each of the four evaluated images, the fine-tuned GPT-3.5-Turbo model consistently received a higher number of "Matches Well" (score = 3) ratings compared to the baseline GPT-3.5-Turbo model. Notably, Image 2 showed the most significant improvement, with the number of participants assigning a top score increasing from 3 to 12 following fine-tuning. This substantial gain suggests that the fine-tuned model developed a stronger capacity to interpret more ambiguous or contextually complex visual scenes and translate them into semantically aligned poetic responses.

Similarly, Image 1 exhibited a marked increase in high coherence ratings—from 7 to 12 participants—indicating enhanced alignment between visual content and poetic output. In addition to the rise in strong alignment scores, there was also a noticeable reduction in low coherence ratings (score = 1), particularly in Images 1 and 3, where the number of "poor fit" assessments dropped by 50% or more. These shifts imply that the fine-tuned model not only improved in capturing the core themes and tones of the images but also generated outputs that resonated more broadly with evaluators.

Images 3 and 4, which initially had moderate alignment scores under the baseline model, also benefited from the fine-tuning process. For Image 3, high alignment ratings rose from 8 to 14, while for Image 4, they increased from 9 to 15, indicating a near-con-

sensus among participants regarding the relevance of the generated poems. The strong performance across all four cases reinforces the conclusion that multimodal fine-tuning substantially enhances the model's ability to integrate visual stimuli into coherent poetic compositions.

Overall, these findings highlight the effectiveness of domain-specific fine-tuning in guiding large language models to produce not only stylistically rich but also contextually grounded literary texts. The fine-tuned model's improved consistency across different visual prompts supports the argument that alignment with visual semantics is a learnable skill within generative models, particularly when exposed to targeted training data. These results also underline the value of structured visual conditioning in creative AI applications, where interpretability and relevance are as crucial as stylistic fidelity.

The Table 4 presents the average scores of four poems, each generated in two versions: one by the baseline GPT-3.5 model and one by a fine-tuned version specifically trained on a curated dataset of works by William Shakespeare and William Blake. Participants were asked to evaluate each poem on a scale of 1 to 10, focusing on literary quality, stylistic authenticity, and emotional resonance.

Table 4: Average Ratings of Poems – GPT-3.5 vs Fine-Tuned Model

Poem	GPT-3.5	Fine-Tuned
1	5.8	5.2
2	5.85	4.05
3	5.45	4.45
4	5.6	4.45

When interpreting the results in Table 4, it is important to consider two different evaluation perspectives. The first is based on how close the scores are to 5, which represents an ideal blend of Shakespeare and Blake's styles. In this case, the smaller the deviation from 5, the better the hybrid style is captured. From this perspective, the baseline GPT-3.5 model generally performed better, producing poems that more closely matched the target hybrid style in three out of four cases.

The second perspective considers higher average scores as an indicator of overall literary quality or readability, regardless of their distance from the midpoint. These scores may reflect how fluent, emotionally resonant, or engaging the poem is for readers. In this regard, GPT-3.5 again received slightly higher ratings overall, suggesting that its outputs were not only stylistically balanced but also more broadly appealing.

Taken together, these results do not present a contradiction but rather illustrate a trade-off: the fine-tuned model may better emulate the original poets' styles, while the baseline model tends to create more accessible and stylistically balanced poems. Both contribute to the goal of stylistically rich and interpretable AI-generated poetry, each from a different strength.

Discussion

The results show that while GPT-3.5 slightly outperformed the fine-tuned model in terms of average ratings, the fine-tuned poems were still well received—particularly in their ability to capture the distinct styles of Shakespeare and Blake. Fine-tuning appears to improve stylistic precision, though occasionally at the expense of fluency or emotional accessibility.

One notable difference was observed in the formal structure of the outputs. GPT-2 frequently exhibited inconsistent capitalization and punctuation, often failing to produce grammatically coherent lines. In contrast, both GPT-3.5 and the fine-tuned version generated text with more consistent syntax and punctuation, indicating that larger model size and fine-tuning contribute positively to textual correctness.

As a shown Figure 2, participant evaluation data further supports these findings. Based on a 10-point scale, where 1 corresponds to Shakespeare, 10 to Blake, and 5 to a hybrid midpoint, the fine-tuned model received scores clustering around 5–6. This suggests that it best approximated the intended fusion of styles. In comparison, GPT-2's scores were concentrated around 4, reflecting a tendency toward Shakespearean elements but lacking hybrid integration. GPT-3.5, however, received scores leaning toward Blake (8–10), indicating a single-author bias.

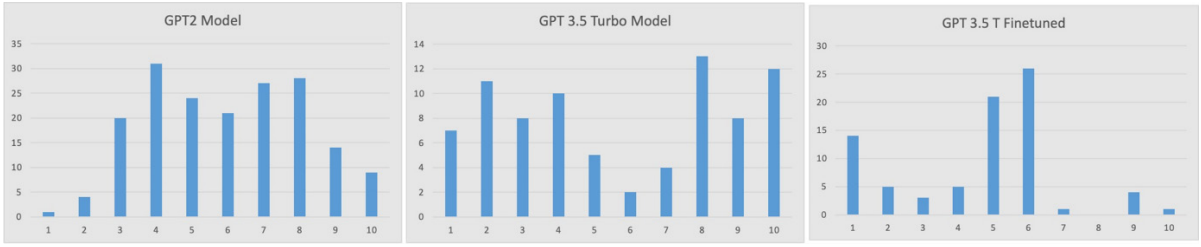


Figure 2. Survey Results of Visual-Poem Coherence
(compare of GPT 2, GPT 3.5 Turbo and GPT 3.5 Finetuned)

These results indicate that there may be a balance between stylistic accuracy and different forms of fluency in poetry. The baseline GPT-3.5 model, trained on a wide range of texts, tends to produce language that is easier to follow and more familiar to readers. However, the fine-tuned model shows stronger control over poetic form and topic continuity. Its outputs are often more consistent in style from one stanza to the next and better follow the structural patterns seen in the training data. This suggests that fluency can also mean maintaining clear structure and style—not just using simple or readable language. Therefore, instead of seeing fluency and style as opposites, the findings suggest that they appear in different ways: one as general readability, the other as organized and stylistically stable poetic flow.

While the results of this study highlight the technical capability of language models to emulate poetic style and structure, a broader philosophical question emerges: what does it truly mean for a machine to generate poetry? Unlike mathematical problem-solving or scientific modeling—where objective evaluation criteria exist—literature, poetry, and aesthetics are inherently subjective. In this light, generative AI may serve a different purpose: not necessarily to replace human artistic expression, but to provoke reflection, offer alternate perspectives, and expand creative possibilities.

For instance, generative systems may be valuable in hybrid thinking frameworks—asking, for example, how Oppenheimer or Feynman might have approached a contemporary scientific challenge, or how historical events might unfold differently under altered conditions. In such cases, AI acts as a tool to simulate thought patterns and generate speculative insight. However, applying this paradigm to poetry or literature raises epistemological and cultural concerns. When a poem's value is defined not by factual correctness but by emotional resonance, nuance, or ambiguity—qualities that differ across individuals and cultures—what is gained (or lost) when a machine becomes the poet?

Therefore, while this study affirms that fine-tuned language models can replicate stylistic elements and produce aesthetically coherent poetic forms, it remains essential to critically assess the purpose and implications of such outputs in artistic domains. AI-generated poetry may be best understood not as literature in itself, but as a reflection of our evolving relationship with language, authorship, and the creative process.

Conclusion

This study explored the generation of stylistically hybrid poetry using large language models fine-tuned on a curated corpus of works by William Shakespeare and William Blake. By integrating multimodal prompts, visual captioning, and domain-specific fine-tuning of GPT-3.5-Turbo, we aimed to produce poetic texts that blend distinct literary styles in response to visual stimuli.

Comparative evaluations demonstrated that the fine-tuned model was able to generate poems that reflected key stylistic features of both authors while maintaining semantic coherence and visual relevance. Participant assessments supported these findings, with the fine-tuned model receiving higher coherence scores across all images and exhibiting a stylistic balance closely aligned with the intended hybrid aesthetic. Additionally, the outputs showed improvements in structural consistency and poetic form.

Our results suggest that targeted fine-tuning can guide language models toward producing more stylistically grounded and context-aware literary content. However, findings also indicate a trade-off between stylistic fidelity and general readability, emphasizing the need to consider both technical and artistic goals in creative text generation.

Future work may expand upon this approach by incorporating additional authors, exploring multilingual stylistic blending, or experimenting with interactive generation frameworks where users can dynamically steer poetic direction. As generative AI continues to evolve, its application in literature offers not only new creative possibilities but also invites critical reflection on authorship, originality, and the role of machines in artistic expression.

REFERENCES

[1] Vaswani, N. Shazeer, N. Parmar, et al., "Attention is All You Need," in Proc. Advances in Neural Information Processing Systems (NeurIPS), 2017.

[2] T. Brown, B. Mann, N. Ryder, et al., "Language Models are Few-Shot Learners," arXiv preprint arXiv:2005.14165, 2020.

[3] H. Gonçalo Oliveira, "A Survey on Intelligent Poetry Generation: Languages, Features, Techniques, Reutilisation and Evaluation," in Proc. 10th Int. Conf. on Natural Language Generation (INLG), Santiago de Compostela, Spain, Sep. 2017, pp. 11–20. doi: 10.18653/v1/W17-3502

[4] J. Belouadi and S. Eger, "ByGPT5: End-to-End Style-conditioned Poetry Generation with Token-free Language Models," arXiv preprint arXiv:2212.10474, 2022. [Online]. Available: https://arxiv.org/abs/2212.10474

[5] D. Nalcı, S. Yeşilyurt, and B. Yılmaz, "William Shakesblake: Exploring Text Generation From Images Using N-grams and Large Language Models," in Proc. Karaelmas Int. Sci. and Eng. Symp. (KISES), Zonguldak, Türkiye, May 2024, pp. 65–66, Paper ID: 8997095.

[6] J. Li, D. Li, C. Xiong, and S. Hoi, "BLIP: Bootstrapping Language-Image Pre-training for Unified Vision-Language Understanding and Generation," arXiv.org, Jan. 28, 2022. https://arxiv.org/abs/2201.12086

[7] OpenAI, "Fine-Tuning GPT-3.5 Turbo," 2024. [Online]. Available: https://platform.openai.com/docs/guides/fine-tuning

[8] H. Thompson and G. Ritchie, "An evolutionary algorithm approach to poetry generation," Jul. 01, 2004. https://era.ed.ac.uk/handle/1842/314

[9] K. E. Gero and L. B. Chilton, "Metaphoria: An AI-powered creativity support tool for generating metaphors," in Proc. ACM CHI Conf. Human Factors in Computing Systems (CHI), Glasgow, UK, May 2019, pp. 1–13. doi: 10.1145/3290605.3300526

[10] Y. Yan, D. Wen, L. Yang, D. Zhang, and H. Lin, "Poetry generation combining poetry theme labels representations," in Proc. Recent Advances in Natural Language Processing (RANLP), Varna, Bulgaria, Sep. 2023, pp. 1162–1170. [Online]. Available: https://aclanthology.org/2023.ranlp-1.132/

[11] Liu, J. Fu, M. P. Kato, and M. Yoshikawa, "Beyond Narrative Description: Generating Poetry from Images by Multi-Adversarial Training," arXiv preprint arXiv:1804.08473, Apr. 2018. [Online]. Available: https://arxiv.org/abs/1804.08473

[12] X. Yi, R. Li, C. Yang, W. Li, and M. Sun, "MixPoet: Diverse Poetry Generation via Learning Controllable Mixed Latent Space," arXiv preprint arXiv:2003.06094, Mar. 2020. [Online]. Available: https://arxiv.org/abs/2003.06094

[13] T. Buz, B. Frost, N. Genchev, M. Schneider, L.-A. Kaffee, and G. de Melo, "Investigating Wit, Creativity, and Detectability of Large Language Models in Domain-Specific Writing Style Adaptation of Reddit's Showerthoughts," arXiv preprint arXiv:2405.01660, May 2024. [Online]. Available: https://arxiv.org/abs/2405.01660

[14] Poetry DB, "A JSON-based Poetry API," 2024. [Online]. Available: https://poetrydb.org/index.html

[15] Hugging Face, "Salesforce/blip-image-captioning-base," 2023. [Online]. Available: https://huggingface.co/Salesforce/blip-image-captioning-base

[16] OpenAI, "GPT Models Overview," 2024. [Online]. Available: https://platform.openai.com/docs/models/gpt-4-and-gpt-3

[17] Google, "Google Colab: Machine Learning in Your Browser," 2024. [Online]. Available: https://colab.research.google.com

[18] Google Drive, "Visual Prompts Used in Poetic Generation Experiment," 2024. [Online]. Available: https://drive.google.com/drive/folders/1iraaBrvgkLtyRoxwppqgfsrnpBxdOV1I

YAPAY ZEKA ŞİRKETLERİNİN SÖYLEMLERİ: KURUMSAL SLOGANLARDA ANLAM İNŞASI

İbrahim Etem Zinderen*
Atatürk Üniversitesi, İletişim Fakültesi, Erzurum, Türkiye

Aslıhan Zinderen
Atatürk Üniversitesi, İletişim Fakültesi, Erzurum, Türkiye
**(İletişimden sorumlu yazar) ibrahim.zinderen@atauni.edu.tr*

ÖZET

Bu çalışma, yapay zekâ alanında faaliyet gösteren küresel teknoloji şirketlerinin kurumsal sloganlarına odaklanmaktadır. Söz konusu sloganlar, birer tanıtım aracı olmanın yanı sıra şirketlerin etik, toplumsal ve ideolojik konumlanmalarını inşa eden söylemsel yapılardır. Bu bağlamda çalışmanın amacı, yapay zekâya ilişkin anlamın söylemsel düzeyde nasıl üretildiğini, küresel şirketlerin sloganları üzerinden analiz ederek ortaya koymaktır. Günümüzde teknoloji şirketleri ürün ve hizmet sunmanın ötesinde etik, sosyal ve politik düzeydeki konumlanmalarını da dilsel stratejiler aracılığıyla inşa etmektedir. Yapay zekâ alanında öne çıkan teknoloji şirketlerinin sloganları ve bu sloganlar üzerinden geliştirdikleri söylem, ilgili şirketlerin toplumsal rol ve işlevlerine yönelik önemli ipuçları sunmaktadır. Çalışma, sloganlar üzerinden şirketlerin kurumsal vizyonlarını nasıl anlamlandırdıklarını ve bu vizyonun ardında yatan etik, ideolojik ve toplumsal yönelimleri Eleştirel Söylem Analizi yöntemiyle ortaya koymayı amaçlamaktadır. Bu doğrultuda OpenAI, Microsoft, NVIDIA, Meta ve Google DeepMind olmak üzere beş küresel teknoloji şirketinin kamuya açık sloganları Norman Fairclough'un üç boyutlu Eleştirel Söylem Analizi modeli temel alınarak değerlendirilmiştir. Sloganların dilsel yapıları, üretim bağlamları ve toplumsal işlevleri çok katmanlı biçimde çözümlenmiştir. Analizlerde, kapsayıcılık, eşitlik ve sorumluluk gibi evrensel etik değerlerin öne çıkarıldığı; buna karşın güç ilişkileri, denetim sorunları ve potansiyel toplumsal risklerin söylem düzeyinde geri planda bırakıldığı görülmüştür. Sloganların marka imajının yanı sıra teknolojiye ilişkin belirli ideolojik yönelimleri ve kurumsal vizyonları da yansıttığı tespit edilmiştir. Yapay zekâ şirketlerinin söylemlerini eleştirel bir bakış açısıyla değerlendirerek teknolojiye dair anlam üretimini görünür kılmayı hedefleyen çalışma, bu yönüyle literatüre özgün bir katkı sunmaktadır.

Anahtar Kelimeler: Yapay zekâ, eleştirel söylem analizi, kurumsal söylem, slogan, teknoloji.

Giriş

Makinelere zekâ kazandırma etkinliğine bağlı olarak dijital devrimi hızlandıran ve yaşamın vazgeçilmez bir unsuru haline gelen yapay zekâ [1, 2; 3, s. 13], teknik bir yenilik olmanın yanı sıra avantajları ve dezavantajları ile birlikte toplumsal, etik ve ideolojik boyutlarıyla öne çıkan çok katmanlı bir tartışma alanı olma özelliği taşımaktadır [4, 5, 6, 1].

Çoğunlukla kısa cümlelerden oluşan; açıklayıcı ve ikna edici özelliklere sahip olan sloganlar, markaların kurumsal kimliğini ve imajını doğru bir biçimde oluşturmak açısından son derece önemli bir stratejik iletişim aracıdır [7, 8]. Bu çalışma, yapay zekâ alanında faaliyet gösteren küresel teknoloji şirketlerinin kurumsal sloganları üzerinden söylemsel anlam inşasını incelemeyi amaçlamaktadır. Söz konusu sloganlar, birer tanıtım ifadeleri olmalarının yanı sıra şirketlerin toplumsal rollerini, etik konumlanmalarını ve teknolojiye dair yaklaşımlarını yansıtan söylemsel yapılardır.

Çalışmanın temel amacı, yapay zekâya ilişkin anlamın bu sloganlar aracılığıyla söylemsel düzeyde nasıl üretildiğini Eleştirel Söylem Analizi yöntemiyle ortaya koymaktır. Bu doğrultuda, öncelikle yapay zekânın toplumsal ve kültürel boyutları ile kurumsal söylem arasındaki ilişki tartışılmış; ardından seçilen küresel şirketlerin sloganları, Norman Fairclough'un üç boyutlu Eleştirel Söylem Analizi modeli çerçevesinde çok katmanlı biçimde çözümlenmiştir.

Yapay Zekâ ve Kurumsal Söylem

Günümüzde teknoloji şirketleri ürün ve hizmet sunan aktörler olmanın yanında etik, sosyal ve politik düzeyde etkili söylemler üreten kurumlar hâline gelmiştir. Bu şirketler, çoğunlukla kendilerini tanımlayan kısa ve anlam yüklü sloganlar aracılığıyla kurumsal söylem inşa etmektedir. Kurumsal söylem, aynı zamanda bir ikna aracı olarak da işlev görmekte; teknolojinin meşrulaştırılması, toplumla ilişkisinin kurulması ve şirketlerin kendilerini nasıl konumlandırdıkları ile hedef kitlede nasıl bir algı yaratmak istediklerine dair stratejik bir yönelim sunmaktadır [9; 10, s. 1663]. Bu bağlamda kurumsal sloganlar, şirketlerin değer dünyasını kısa, etkili ve sembolik biçimde dışa vuran stratejik dilsel pratikler olarak değerlendirilebilir.

Yapay Zekâ Şirketlerinin Profili

Bu çalışmada analiz edilen beş küresel teknoloji şirketi (OpenAI, Microsoft, NVIDIA, Meta ve Google DeepMind), yapay zekâ alanındaki sektörel etkileri, söylemsel çeşitlilikleri ve kamuya açık slogan kullanımları dikkate alınarak seçilmiştir. Söz konusu şirketler gerek teknolojik üretimleriyle gerekse küresel ölçekteki söylem kurma güçleriyle de öne çıkmaktadır. Bu bağlamda, sloganlar üzerinden geliştirilen söylemler; etik sorumluluk, toplumsal kapsayıcılık, teknolojik liderlik, bilimsel ilerleme ve insanlık yararı gibi tematik eksenlerde farklılaşmakta ve şirketlerin teknolojiye dair anlam üretim biçimlerine yönelik karşılaştırmalı bir çözümleme imkânı sunmaktadır.

- OpenAI [11], açık ve güvenli yapay zekâ vizyonuyla küresel fayda odaklı bir söylem benimsemekte;

- Microsoft [12], kapsayıcı dil stratejisi ve sorumluluk vurgusuyla etik teknoloji üretimini öne çıkarmakta;

•NVIDIA [13], yapay zekâyı bilişim altyapısının temel unsuru olarak konumlandırarak teknik yetkinliğe dayalı bir liderlik söylemi üretmekte;

•Meta [14], yapay zekâyı bağlantılı bir dünya idealiyle ilişkilendirirken etik geliştirme taahhüdünü de dile getirmekte;

•Google DeepMind [15] ise bilimsel araştırma ve insanlık yararına katkı odaklı bir yapay zekâ anlayışıyla konumlanmaktadır.

Bu söylemsel farklılıklar, şirketlerin kurumsal kimliklerini ve yapay zekâya ilişkin toplumsal tahayyülleri nasıl biçimlendirdiklerini görünür kılmakta; böylece Eleştirel Söylem Analizi çerçevesinde anlamlı bir karşılaştırmaya olanak tanımaktadır.

Yöntem

Bu çalışma, yapay zekâ alanında faaliyet gösteren önde gelen teknoloji şirketlerinin kurumsal sloganlarında inşa edilen anlamları çözümlemek amacıyla, Norman Fairclough'un üç boyutlu Eleştirel Söylem Analizi modeli temelinde yapılandırılmıştır.

Eleştirel söylem analizi, metinlerin yüzeyinde doğrudan görünmeyen ideolojik ve güç ilişkilerini açığa çıkarmayı amaçlayan bir yaklaşım olarak toplumsal süreçlerle dilsel yapıların bağlantısını sistemli biçimde incelemeyi hedeflemektedir [16, s. 97]. Fairclough'a göre söylem, dilsel bir olay olduğu kadar, toplumla kurulan ilişkiyi, kimlikleri, kurumları ve iktidarı şekillendirme biçimidir [17, s. 258]. Dolayısıyla söylem, toplumsal gerçekliğin nasıl anlamlandırıldığını belirleyen ve egemen sosyal pratiklerle bağlantılı olarak şekillenen bir söylem düzeni olarak tanımlanmaktadır [16, s. 96].

Fairclough, iktidar ile toplumsal yapı arasındaki ilişkiyi çözümlemek amacıyla söylemin kurumsal âdetler ve toplumsal ilişkiler tarafından nasıl şekillendiğini inceleyen üç boyutlu bir analiz çerçevesi geliştirmiştir. Bu yaklaşım, metinlerin üretimi ve yorumlanış biçimleri ile toplumsal bağlam arasındaki ilişkiyi görünür kılmaktadır. Fairclough'un önerdiği üç boyutlu analiz modeli, her söylem olayını eşzamanlı olarak metin (text), söylem pratiği (discourse practice) ve sosyokültürel pratik (sociocultural practice) olmak üzere üç düzlemli bir çözümleme çerçevesi sunmaktadır [16, s. 98]. Bu doğrultuda çalışmada, şirket sloganları ilk olarak metin düzeyinde dilsel ve anlamsal özellikleri bakımından değerlendirilmiş; ardından söylemsel pratik düzeyinde, bu sloganların üretim bağlamları, kurumsal iletişim stratejileri ve ideolojik yönelimleri analiz edilmiştir. Son olarak, sloganların toplumsal pratikler bağlamında teknolojiye, topluma ve etik sorumluluğa dair hangi anlamları taşıdığı ve bu anlamların nasıl bir ideolojik konumlanma sunduğu incelenmiştir.

Veri seti, amaçsal örnekleme yöntemiyle oluşturulmuştur. Örnekleme dâhil edilen beş küresel teknoloji şirketi (OpenAI, Microsoft, NVIDIA, Meta ve Google DeepMind), yapay zekâ alanında söylem üretme kapasitesine sahip olmaları ve uluslararası düzeyde yüksek kurumsal görünürlüğe ulaşmaları nedeniyle seçilmiştir. Örneklem belirlenirken aşağıdaki ölçütler esas alınmıştır:

•**Sektörel Etki ve Tanınırlık:** Yapay zekâ alanında ürün geliştirme, bilimsel araştırma yatırımları ve teknolojik yenilik kapasitesi bakımından küresel ölçekte belirleyici olan şirketler tercih edilmiştir.

•**Söylemsel Çeşitlilik:** Seçilen şirketlerin sloganları, farklı söylem stratejileri ve ideolojik konumlanmalar ortaya koymaktadır. Bazı sloganlar; etik sorumluluk, bireysel güçlenme ve teknik yeterlilik temaları etrafında şekillenmektedir.

•**Kurumsallık ve Açıklık:** Yalnızca şirketlerin resmî web sitelerinde veya sürekli kullanılan kurumsal iletişim materyallerinde yer alan, kamuya açık ve sürdürülebilir nitelikteki sloganlar değerlendirmeye alınmıştır. Geçici kampanyalara ait söylemler ya da bağlamı belirsiz tanıtım ifadeleri dışarda bırakılmıştır.

Sloganlar, içeriklerine göre tematik kategorilere ayrılmış ve her biri Fairclough'un sunduğu çözümleyici çerçeve doğrultusunda analiz edilmiştir. Analiz sürecinde;

•Sloganların hem sözcük ve yapı düzeyindeki dilsel öğeleri hem de bağlamsal, temsili ve ideolojik yönleri ele alınmıştır.

•Her bir slogan, kurumsal imaj, teknolojiye ilişkin toplumsal tahayyül ve kamusal söylem üretimi açısından işlevsel olarak yorumlanmıştır.

•Söylemlerin, toplumsal düzlemde hegemonik anlam üretim süreçlerine katkısı ve bu süreçlerde üstlendikleri rol çözümlenmiştir.

Bu çalışmanın temel sınırlılığı, yalnızca beş şirketin sloganlarına dayalı, görece dar bir veri seti ile kısıtlı olmasıdır. Sloganlar her ne kadar kurumsal iletişimde resmî olarak kullanılan söylemler olsa da kamuoyundaki algılanış biçimleri ya da şirketlerin pratik uygulamalarıyla olan tutarlılıkları bu çalışmanın kapsamı dışında tutulmuştur. Ayrıca sloganlar, belirli bir zaman dilimindeki kurumsal söylemi yansıttığı için şirketlerin uzun vadeli söylemsel dönüşümleri, bu analiz kapsamında ele alınmamıştır. Bu bağlamda çalışma, sınırlı örneklem üzerinden gerçekleştirilen derinlemesine, niteliksel bir söylem çözümlemesi niteliği taşımaktadır.

Kurumsal Sloganlarda Söylem İnşası: Bulgular ve Değerlendirme

Norman Fairclough'un üç boyutlu eleştirel söylem analizine dayanan bu çalışmada, incelenen kurumsal sloganlar metin, söylemsel pratik ve toplumsal pratik düzeylerinde analiz edilmiştir. Bu çerçevede söylemlerin içerdiği temel değerler, temsil ettikleri kurumsal yönelimler ve taşıdıkları ideolojik yapı çözümlenmiştir. Bu çözümleme sonucunda dört temel söylem teması öne çıkmıştır: evrensel etik ve insanlık yararı, etik sorumluluk ve denetim, kapsayıcılık ve güçlendirme, teknoloji odaklılık ve inovasyon. Tablo 1'de analiz edilen beş şirketin sloganları, içerdiği temalar ve temsil ettikleri söylem yönelimleri ÖZETlenmiştir. Bu yapı, çözümlemenin temelini oluşturan tematik ayrımı yansıtmaktadır:

Tablo 1: Yapay Zekâ Şirketlerinin Sloganları ve Temsil Ettikleri Söylem Temaları

Şirket	Slogan	Temsil Ettiği Söylem
OpenAI	Tüm insanlığın yararı için yapay zekâ *AI for the benefit of all humanity*	Evrensel etik değerler odağında insanlık yararını önceleyen küresel vizyon
Microsoft AI	İnsanları güçlendirmek için sorumlu yapay zekâ üretiyoruz *We make responsible AI to empower people*	Bireysel güçlenme ve toplumsal kapsayıcılığı destekleyen sorumlu teknoloji söylemi
NVIDIA	Daha iyi bir dünya için güvenilir yapay zekâ *Trustworthy AI For A Better World*	Güvene dayalı teknik ilerleme ve gelecek vizyonu
Meta (AI)	Daha bağlantılı bir dünya için yapay zekâyı geliştiriyoruz *(We're advancing AI for a more connected world)*	Bağlantılılık ve bütünleşme üzerinden yapılandırılmış teknolojik gelişim söylemi
Google DeepMind	İnsanlık yararına, yapay zekâyı sorumlu bir şekilde geliştirmek *Build AI responsibly to benefit humanity*	Etik sorumlulukla çerçevelenmiş bilimsel ilerleme anlayışı

Evrensel etik ve insanlık yararı odaklı söylem

OpenAI ve Google DeepMind, sloganlarında yapay zekâyı evrensel etik ilkelerle uyumlu, insanlık yararını önceleyen bir teknolojik araç olarak sunmaktadır.

- Sloganlar metin düzeyinde kolektif fayda, insanlık ve evrensellik gibi değerler üzerine inşa edilmiştir.

- Söylemsel pratik düzeyinde şirketler kendilerini küresel kamu yararını gÖZETen sorumlu aktörler olarak tanıtmaktadır.

- Toplumsal pratik düzeyinde ise bu söylemler, teknolojik ilerlemenin etik sorumlulukla bütünleştiği ideolojik bir çerçeve oluşturmaktadır.

OpenAI'nin "AI for the benefit of all humanity" sloganı, yapay zekânın küresel eşitlik ilkesi doğrultusunda kullanılmasını savunurken; Google DeepMind'ın "Build AI responsibly to benefit humanity" sloganı, bilimsel bilgi üretimini toplumsal fayda ile ilişkilendirmektedir. Her iki söylemde de teknoloji olgusu etik ve insan merkezli bir yaklaşım ekseninde ele alınmaktadır.

Etik sorumluluk ve denetim vurgusu

Meta AI, yapay zekânın sorumlu bir şekilde geliştirilmesi gerektiğini vurgulayarak etik denetimi söylemin merkezine yerleştirmektedir.

- Metin düzeyinde, "responsibly" kavramı dikkat çekicidir. Bu kavram, teknolojinin hem gelişme hem de denetim gerektiren bir alan olduğuna gönderme yapmaktadır.

- Söylemsel pratik düzeyinde şirket, kamusal endişelere karşı duyarlı bir pozisyon almaktadır.

- Toplumsal düzeyde ise güvenlik, gizlilik ve sosyal etki gibi meselelerde sorumluluk üstlendiği yönünde bir kurumsal imaj oluşturmaktadır.

Meta'nın söylemi, kurumsal etik pozisyonunu güçlendirmenin yanı sıra teknolojinin öngörülemeyen sonuçları karşısında aktif bir denetim anlayışına sahip olduğunu vurgulamaktadır.

Kapsayıcılık ve güçlendirme söylemi

Microsoft AI, bireylerin yapay zekâ yoluyla güçlenmesini merkeze alan bir söylem üretmektedir.

- Metin düzeyinde, "empower people" ifadesi, bireysel gelişim ve dijital eşitlik gibi değerleri ön plana çıkarmaktadır.

- Söylemsel pratik düzeyinde şirket, teknolojiyi erişilebilir ve kapsayıcı bir araç olarak tanımlamaktadır.

- Toplumsal düzeyde ise söylem, dijital uçurumu azaltma ve bireysel potansiyeli destekleme yönünde ideolojik bir duruşu yansıtmaktadır.

Bu söylem, teknolojiyi verimlilik aracı ve toplumsal eşitliği teşvik eden temel unsur olarak ele alan bir bakış açısını ortaya koymaktadır.

Teknoloji odaklılık ve inovasyon temsili

NVIDIA, sloganında teknolojik kapasite ve mühendislik becerilerini öne çıkararak, yapay zekâyı teknik bir üstünlük alanı olarak konumlandırmaktadır.

- Metin düzeyinde, "AI computing company" ifadesi doğrudan teknik bir odak taşımaktadır.

- Söylemsel pratik düzeyinde şirket, inovasyon ve bilişim altyapısındaki liderliğini vurgulayan bir kurumsal kimlik inşa etmektedir.

- Toplumsal pratik düzeyinde ise bu söylem, etik ya da sosyal referanslardan ziyade, teknik yeterlilik ve rekabet üstünlüğü üzerinden bir meşruiyet arayışını ortaya koymaktadır.

Bu yaklaşım ile yapay zekâ olgusu dönüşümden çok, teknoloji endüstrisinin ilerleyişini temsil eden bir yapı olarak sunulmaktadır. Söz konusu söylem, etik ya da toplumsal referanslardan ziyade teknik yeterliliğe dayalı bir meşruiyet arayışına işaret etmektedir.

Söylemsel Temaların İdeolojik Yönelimleri

Kurumsal sloganlara dayalı söylem çözümlemesi, şirketlerin iletişim stratejilerini ve bu stratejiler aracılığıyla inşa ettikleri ideolojik çerçeveleri, meşruiyet söylemlerini ve toplumsal pozisyonlanmalarını görünür kılmaktadır.

- Etik temelli söylemler, şirketlerin kamu güveni oluşturma çabalarının bir yansıması olarak dikkat çekmektedir.

- Teknoloji merkezli söylemler, kurumsal liderlik ve rekabet gücü vurgularını öne çıkarmaktadır.

- Kapsayıcı söylemler ise toplumsal sorumluluk anlayışıyla bağlantılı olarak şirketlerin sosyal görünürlüğünü artırma işlevi taşımaktadır.

Yapay zekâ şirketlerinin sloganları, kurumsal vizyonu ÖZETlerken teknolojik gelişmenin toplumsal ve ideolojik anlamlarını yapılandıran stratejik söylemsel üretimler olarak da işlev taşımaktadır. Her bir söylem tipi, yapay zekâyı teknik bir olgu olmanın ötesinde; toplumsal anlam, etik sorumluluk ve kurumsal stratejiyle bütünleşen çok katmanlı bir alan olarak sunmaktadır.

Sonuç

Bu çalışma, yapay zekâ şirketlerinin kurumsal sloganlarını sadece tanıtım amaçlı ifadeler olarak değil, aynı zamanda şirketlerin teknolojiye, topluma ve etik sorumluluğa dair kurduğu ideolojik anlam evreninin bir parçası olarak değerlendirmiştir. Fairclough'un üç boyutlu Eleştirel Söylem Analizi modeli doğrultusunda gerçekleştirilen çözümleme, sloganların dilsel yapılarının ötesinde, bu söylemlerin üretim bağlamlarını ve toplumsal işlevlerini görünür kılmıştır.

Analizler; sloganların etik değerler, insanlık yararı, bireysel güçlenme ve teknik yeterlilik gibi kavramlar etrafında şekillendiğini ortaya koymuştur. Bu değerler aracılığıyla şirketler, kurumsal meşruiyetlerini güçlendirmeye, kamusal güven oluşturmaya ve teknolojik liderlik iddialarını pekiştirmeye çalışmaktadır. Ancak bu kapsayıcı görünümün ardında teknoloji şirketlerinin güç ilişkilerini, denetim eksikliklerini ve toplumsal eşitsizlikleri söylem düzeyinde görünmez kılan ya da geri plana iten bir ideolojik yönelimin inşa edildiği görülmektedir. Böylece kurumsal söylem, bir yandan meşruiyet üretirken diğer yandan eleştirel soruları görünmez kılan bir ideolojik filtreleme işlevi de üstlenmektedir.

Çalışma, teknoloji şirketlerinin kurumsal sloganlarının yalnızca içerik düzeyinde değil; söylemsel stratejiler, temsil pratikleri ve toplumsal etkiler düzeyinde de analiz edilmesi gerektiğini ortaya koymuştur. Gelecek araştırmalar, bu tür kurumsal söylemlerin medya, siyaset ve kullanıcı deneyimleri gibi farklı bağlamlar üzerinden incelenmesiyle daha kapsamlı analizler ortaya koyabilir. Bu tür karşılaştırmalı çalışmalar, teknoloji şirketlerinin dil aracılığıyla inşa ettikleri anlam evrenini daha geniş toplumsal ve kültürel bağlamlarda değerlendirme olanağı sağlayacaktır.

Çıkar Beyanı: Yazarlar, bu çalışmanın hazırlanması sürecinde herhangi bir çıkar çatışması bulunmadığını beyan etmektedir.

Yazar Katkıları: Her iki yazar da çalışmanın tüm aşamalarına katkı sağlamış ve nihai metni onaylamıştır.

REFERANSLAR

[1] M. Boni, "The ethical dimension of human–artificial intelligence collaboration," *European View*, vol. 20, no. 2, pp. 182–190, 2021. https://doi.org/10.1177/17816858211059249

[2] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education – where are the educators?," *Int. J. Educ. Technol. High. Educ.*, vol. 16, no. 1, 2019. https://doi.org/10.1186/s41239-019-0171-0

[3] N. J. Nilsson, *Yapay Zeka: Geçmişi ve Geleceği*, M. Doğan, Trans., İstanbul: Boğaziçi Üniversitesi Yayınevi, 2019. (Originally published 1998 as *Artificial Intelligence: A New Synthesis*)

[4] V. Dignum, "Relational artificial intelligence," *arXiv preprint*, arXiv:2202.07446, 2022. https://doi.org/10.48550/arXiv.2202.07446

[5] S. Mohamed, M. T. Png, and W. Isaac, "Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence," *Philos. Technol.*, vol. 33, pp. 659–684, 2020. https://doi.org/10.1007/s13347-020-00405-8

[6] C. Roche, P. J. Wall, and D. Lewis, "Ethics and diversity in artificial intelligence policies, strategies and initiatives," *AI Ethics*, vol. 3, pp. 1095–1115, 2023. https://doi.org/10.1007/s43681-022-00218-9

[7] M. Supphellen and I. Nygaardsvik, "Testing country brand slogans: Conceptual development and empirical illustration of a simple normative model," *J. Brand Manag.*, vol. 9, no. 4, pp. 385–395, 2002. https://doi.org/10.1057/palgrave.bm.2540085

[8] M. Y. Lee and D. W. So, "Corporate slogans of corporations operating in Greater China," *Corporate Commun.*, vol. 12, no. 1, pp. 58–74, 2007. https://doi.org/10.1108/13563280710723750

[9] J. A. Oravec, "Mottos and ethical statements of Internet-based organizations: Implications for corporate social responsibility," *Int. J. Civic Engagement Soc. Change*, vol. 1, no. 2, pp. 37–53, Apr.–Jun. 2014.

[10] S. E. Green and Y. Li, "Rhetorical institutionalism: Language, agency and structure in institutional theory since Alvesson 1993," *J. Manage. Stud.*, vol. 48, no. 7, pp. 1662–1697, 2011. https://doi.org/10.1111/j.1467-6486.2011.01022.x

[11] OpenAI, "About OpenAI," https://openai.com/about/, Erişim tarihi: Nisan 2025.

[12] Microsoft, "Microsoft AI," https://microsoft.ai/, Erişim tarihi: Nisan 2025.

[13] NVIDIA, "Trustworthy AI," https://www.nvidia.com/en-us/ai-data-science/trustworthy-ai/, Erişim tarihi: Nisan 2025.

[14] Meta AI, "About," https://ai.meta.com/about/, Erişim tarihi: Nisan 2025.

[15] Google DeepMind, "About," https://deepmind.google/about/, Erişim tarihi: Nisan 2025.

[16] N. Fairclough, *Discourse and Social Change*, Cambridge: Polity Press, 1992.

[17] N. Fairclough, *Critical Discourse Analysis: The Critical Study of Language*, London: Longman, 1995.

YAPAY ZEKÂ VE HABERCİLİK ÜZERİNE YAZILAN AKADEMİK ÇALIŞMALAR ÜZERİNE BİR İNCELEME

Alper Çolak*

Atatürk Üniversitesi, İletişim Fakültesi, Gazetecilik Bölümü, Erzurum, Türkiye

(İletişimden sorumlu yazar) alpercolak@atauni.edu.tr

ÖZET

Bu çalışma, bilgisayar teknolojisindeki gelişmelerle ortaya çıkan robot gazetecilik ve algoritmik gazetecilik kavramlarının gazetecilik mesleğindeki etkilerini bibliyometrik perspektiften incelemektedir. Yapay zekânın gazetecilik uygulamalarında kullanımı, haber toplama, üretim ve dağıtım süreçlerinde önemli değişimlere yol açmıştır. Araştırma, "Yapay Zekâ ve Gazetecilik" alanındaki akademik literatürün tarihsel gelişimini ve mevcut durumunu analiz etmeyi amaçlamaktadır. Çalışma kapsamında, Web of Science veri tabanında indekslenen "Yapay Zekâ ve Gazetecilik" konulu tüm yayınlar incelenenmiştir. Bibliyometrik analiz yöntemi kullanılarak, alandaki ilk çalışmaların ne zaman yapıldığı, yayın sayısındaki artış dönemleri, öne çıkan akademisyenler, çok yazarlılık oranları, en çok atıf alan araştırmacılar, yayınların ülkelere göre dağılımı ve bu araştırmaları destekleyen kurumlar haritalandırılmıştır. Bu analiz sonucunda, alandaki temel literatür belirlenecek, önemli çalışmalar tanımlanacak ve literatürdeki araştırma boşlukları tespit edilmiştir. Çalışma Web of Science veri tabanında 'Yapay Zekâ ve Gazetecilik' anahtar kelimesiyle arama sonucunda çıkan akademik çalışmalarla sınırlı olacaktır. Çalışmanın yöntemi bibliyometrik analiz tekniği olarak tespit edilmiştir. Bulgular, yapay zekâ ve gazetecilik alanındaki bilimsel bilgi üretiminin zamansal ve yapısal örüntülerini ortaya koyarak, gelecekteki araştırma yönelimlerine ışık tutacaktır. Yapılan incelemeler sonucunda 1992 yılında başlanılan çalışmalar 2021 yılı itibarıyla hız kazandığı görülürken ülke olarak Çin Halk Cumhuriyetleri, ABD ve İspanya öncülük etmektedir. 'Dijital Yerel Medya', 'Dijital Uçurum', 'Spor Gazeteciliği', Etik Riskler' gibi konular ise yeni araştırma alanlarını oluşturduğu görülmektedir.

Anahtar Kelimeler: Yapay zekâ, gazetecilik, bibliyometrik analiz.

Giriş

Teknoloji her alanda olduğu gibi gazetecilik alanında da değişim ve dönüşümlere: haber toplama, üretim ve dağıtım süreçlerinde değişikliklere sebep olmuştur. Bu durum haber odalarını, haber endüstrisini ve gazeteciler ile sermaye sahipleri arasındaki ilişkiyi dönüştürmüştür [1]. Lewis ve Guzman [2] 'Yapay Zekâ' teknolojisinin gazetecilik sektörünü büyük ölçüde dönüştürdüğünü, dönüşümün robotlar, algoritmalar, sosyal medya gibi dijitalleşme ile olduğunu ifade etmektedir. Yapay Zekâ teknolojisiyle üretilen haberler okuyucunun ilgi alakasına göre belirlenmekte, algoritmalar yardımıyla okuyucuya ulaştırılmaktadır. Bazı akademisyenlere göre bu süreç haberlerin sığ olarak kalmasına neden olmakta ve insan

diline yakın haberlerin üretilemeyeceği düşüncesine neden olmaktadır [3]. Bu düşünceler ve alanda gerçekleşen değişimlerle birlikte 'Yapay Zekâ ve Gazetecilik' konuş akademide popüler hale gelmiştir.

Stats Monkey ile başlayan yapay zekâ ile haber üretme uygulamaları daha sonradan çeşitlenerek: Narrative Science, Automated Insights, Wordsmith, Quakebot, Urbs Media gibi uygulamalar ve şirketlerle çoğalmıştır. Bu uygulamalar Associated Press, New York Times, Reuters, Los Angeles Times gibi ajans ve gazeteler tarafından ilk başlangıç aşamasından itibaren kullanıldığı görülmektedir. Carlson [4] robot gazeteciliği/algoritmik gazeteciliği: insan yardımı olmaksınız verilerin anlatısal bir haber metine dönüştürme süreci olarak tanımlamaktadır. Latar [5] ise robot gazeteciliği küçük veya büyük veri olduğuna bakmaksınız verilerin saniyeler içinde ve otomatik olarak yapay zekâ algoritmaları tarafından okunabileceğini, işlenerek anlam çıkarılarak bir rapora veya habere dönüştürülebileceğini ifade etmektedir. Bu bağlamda, yapay zekâ, herhangi bir platformdan anlık olarak veri çekebilecek kapasiteye sahip olup, aldığı verileri belirli bir algoritmaya göre analiz ederek, haber değeri taşıyıp taşımadığını değerlendirebilir. Ardından, bu verileri hızlı bir şekilde haber sitesine yükleyebilir ve yayımlanabilir hale getirebilir.

Bu bilgilere örnek olarak Los Angeles Times'ın kullandığı Quakebot verilebilir. Quakebot 17 Mart 2014 tarihinde meydana gelen depremi veri madenciliği sayesinde hemen çekerek haberleştirmiş ve bunu e-posta olarak Ken Schwencke'ye göndermiştir. Bu haberi diğer haber kuruluşlarına göre bir diğer ifadeyle insan muhabire göre çok daha hızlı yapmıştır [6]. Quakebot örneği dönüşümün ne düzeyde başladığını gösterir niteliktedir. Ancak şunu da gözden kaçırmamak gerekmektedir. 21 Haziran 2017 tarihinde ise Amerika Birleşik Devletleri Jeoloji Araştırmaları Kurumu (USGS) Santa Barbara merkezli 1925 yılında gerçekleşen 6.8 depremi bir depremin bilgi metni üzerinde güncelleme yapmış ve Quakebot bu depremin o an olduğunu sanarak haberleştirmiştir [7]. Yapay zekâ her zaman veriyi doğru işleyememekte ve doğru haber yapamamaktadır. Yapay zekâ, öğrenme yeteneğine sahip bir sistem olmasına rağmen, her zaman yanlış yorumlama veya yanıltıcı verilere dayanarak hatalı sonuçlar üretme riski taşımaktadır.

Bu bilgiler çerçevesinde yapay zekâ her an bağlı olduğu internette gerçekleşen olayları incelemekte ve bu verileri işleyerek normal bir insana göre mola vermeksizin ya da uyumaksınız birçok rapor ve haber metni üretebilme becerisine sahiptir. Çalışma bu çerçevede ilk olarak yapay zekâ ve gazetecilik üzerine bir literatür okuması sunacak daha sonra bibliyometrik analiz yöntemiyle 'Yapay Zekâ ve Gazetecilik' üzerine akademik çalışmalar haritalandıracaktır.

Yapay Zekâ ve Gazetecilik

Alan Turing 'Yapay Zekâ' kavramını ilk ortaya atan kişidir. 1948 yılında yazdığı makalede Turing insanın yaptığı belli işlevleri yapabilecek makinelerin geliştirileceğini ifade etmiştir [8]. Turing'in bu ifadesi ilk olarak makinelerin geliştirilmesi daha sonra ise bilgisayar teknolojileriyle birlikte yapay zekânın ortaya çıkmasıyla gerçekleşmiştir. Hatta günümüzde yapay zekâ ve robotların insanlık ölçütlerine ne kadar uygun olup olmadığı Turing Testi ile yapılmaktadır.

Bu bilgiler ışığında ilk yapay zekâ uygulamalarından birisi Google'ın DeepMind'ı olmuştur. DeepMind hep metin okuması yapabilmekte hem de Go oyunları oynayabilme özelliğine sahiptir [9]. Bunun yanında Google 'Google Assistan', Amazon 'Alexa', Apple 'Siri' Microsoft 'Cortona' gibi kişilerle konuşabilen ve bu konuşma çerçevesinde verilen talimatı yapabilen sistemler geliştirilmiştir [10]. Bu sistemler Turing'in ifade ettiği gibi insanların yapabileceği görevleri yapmaya başlamış ve bir kişinin asistanı gibi işlemler/görevler yapabilmektedir.

Yapay zekâda önemli olan unsurların başında dil öğrenme ve metin oluşturma becerisi gelmektedir. Guzman ve Lewis [11] makinelerin insani iletişim için; öğrenme, işleme ve insan dilinde mesaj üretebilme becerilerine sahip olmaları gerektiğinin önemi üzerde durmaktadır. Yapay zekâ bir metni okuyarak öğrenebilmeli, bunu işleyerek içindeki verileri anlamlı bir hale getirebilmeli ve insan dilinde anlamlı içeriklere dönüştürerek bunu sunabilmelidir. Bu özelliklerin ChatGPT gibi uygulamalarla mümkün olmasıyla birlikte gazeteciler sıradan haberler yerine daha kapsamlı haber hikayelerine odaklanma imkânına sahip olmuştur [12].

Yapay zekâda gerçekleşen bu teknolojik ilerlemeler birçok sektörde olduğu gibi gazetecilik alanını da etkilemektedir. Bu çerçevede Lewis ve Guzman [13] 'Yapay Zekâ'yı kapsamlı dijital dönüşümün son halkası olarak görmektedir. Bu teknolojik gelişimle birlikte gazetecilik: robot gazetecilik, algoritmik gazetecilik gibi isimlerle anılır olmaya başlamıştır [14]. Buradan hareketle gazetecilikte yeni bir dönemin açıldığı ve bu dönem içerisinde hem çalışanlar açısından hem de işverenler açısından değişimin olacağını görmek aşikardır. İlk robot gazeteci Çin tarafından 2018'da Xinhua Ajansında haber üretimi yapmak için kullanılan 'Jiajia'dır. 2019 yılında ise ilk yapay zekâ sunucusu 'Xin Xiaomeng' Çin'de ekranlara çıkmıştır [15]. Buradan hareketle yapay zekâ haber üretimi ve sunuculuk gibi haberciliğin birçok dalında görünürlüğünü arttırmaya başlamıştır.

Bu bilgiler ışığında yapay zekânın gazetecilik alanında kullanımı zaman ve maliyet tasarrufu sağlamasının yanında verimliliği de arttırdığı söylenebilir. Buna örnek olarak ise 'Stats Monkey' örneği verilebilir. Stats Monkey ilk başta oyun hikayeleri yazabilen bir yapay zekâ şirketiydi. Ancak daha sonra beyzbol maçları hakkında verileri kullanarak haber metinleri oluşturabilmekte, en önemli an ve durumları da ön plana çıkaracak bir sisteme evrilmiştir [16]. Bu özelliğiyle Stats Monkey 2011 yılında 400 bin rapor hazırlayabilmiştir.

Stats Monkey'in yanında Google News gibi sistemler de geliştirilerek algoritmalar aracılığıyla haber yapılmaktadır. The New York Times haberde hangi konunun ön plana çıkarılacağı, tık sayısının nasıl arttırılacağı, abone sayılarını arttırma yöntemleri için 'Google News'i kullanmıştır [17]. Moschovakis'e göre [18], yapay zekânın haber üretiminde tam olarak şu anda iyi olmasa da gelecekte insanların yerini alabileceğini düşünmektedir. Google News örneğinde ve Moschovakis'in düşünceleri çerçevesinde yapay zekâ gazetecilik sektöründe sıkça kullanılmaya başlandığı görülmekte ve teknolojinin yardımıyla daha iyi içerikler üretmesinin yanında sektörde büyük değişimlere neden olacağı görülmektedir.

Yöntem

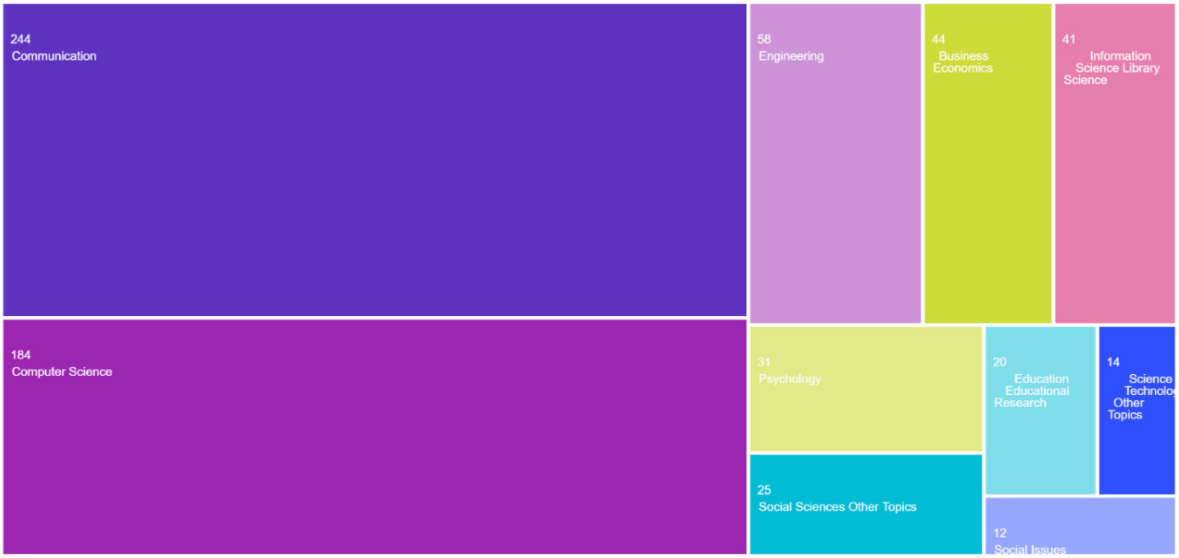
Bibliyometrik analiz, niceliksel olarak verilerin programlar veya uygulamalar aracılığıyla bibliyometrik verilere dayanarak görselleştirme süreci olarak tanımlanmaktadır [19]. Bu yöntem sayesinde literatürde en çok kullanılan kelimelerin tespiti, en fazla çalışma yapan

ülke veya ülkeler, en fazla çalışma yapan yazar veya yazarlar, hangi konu veya konularda çalışma yapıldığı, yazarların, nesnelerin veya çalışma türlerinin birbirleri ile olan bağları net olarak görselleştirilebilmektedir. Bu sayede literatürün süreç içerisinde geçirdiği gelişim rahat ve net olarak gözler önüne serilebilmektedir [20].

Bu çalışmanın amacı yapay zekâ ve gazetecilik alanını inceleyerek, literatürde ne zaman, kimler tarafından, ne konularda, hangi anahtar kelimeler ve hangi yayın türünde çalışma yapıldığını ortaya çıkarmaktır. Bu bilgiler çerçevesinde tanımlayıcı araştırma yöntemi olarak bibliyometri yöntemi kullanılmıştır. Bu yöntem sayesinde geçmişten 2024 yılına kadar olan literatür taranacaktır. Taranan veriler ışığında 2024 yılına kadar çalışılan konular ve hangi alanın bakir kaldığı tespit edilecektir. Bu çerçevede 5 Mart 2024 tarihi itibarıyla Web of Science'da (WOS) İngilizce olarak "Artificial Intelligence" ve "Journalism" kelimelerinin taraması yapılmıştır. Tarama sonucunda toplamda 615 akademik çalışma tespit edilmiştir. Verilerin yalnızca WOS veri tabanından alınması çalışmanın sınırlılığını oluşturmaktadır. Web of Science'dan elde edilen veriler ise VosViewer adlı program yardımıyla ve Web of Science'ın imkân tanıdığı görselleştirme yöntemleriyle görselleştirilmiştir.

Bulgular ve Analiz

Bu bölümde, Web of Science veri tabanından elde edilen verilerin analizi ve görselleştirilmesi sunulmaktadır. 5 Mart 2024 tarihinde gerçekleştirilen arama, "Artificial Intelligence" ve "Journalism" anahtar kelimeleri kullanılarak yapılmış ve toplamda 615 çalışma tespit edilmiştir. Çalışmalar, çeşitli disiplinlere dağılmakta olup, bunların 244'ü "İletişim", 184'ü "Bilgisayar Bilimleri" ve 58'i ise "Mühendislik" alanlarında gerçekleştirilmiştir. Geriye kalan çalışmalar ise, Şekil 1'de görsel olarak sunulmaktadır.



Şekil 1: Çalışmaların yoğunlaştığı alanlar.

WOS'tan elde edilen veriler doğrultusunda, toplam 615 çalışmanın 471'i makale, 209'u erken erişim ve 97'si bildiri türünde olduğu tespit edilmiştir. Doküman türlerine ilişkin detaylı bilgiler Tablo 1'de sunulmuştur.

Tablo 1: Çalışmaların doküman tipine göre dağılımı

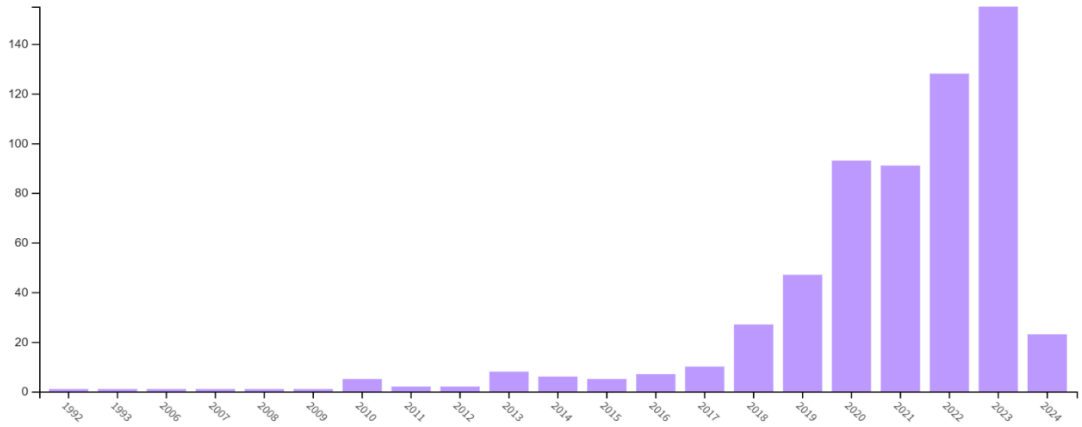
Select All ☐	Field: Document Types	Record Count	% of 615
☐	Article	471	76.585%
☐	Early Access	209	33.984%
☐	Proceeding Paper	97	15.772%
☐	Review Article	22	3.577%
☐	Book Chapters	20	3.252%
☐	Editorial Material	17	2.764%
☐	Book Review	5	0.813%
☐	Retracted Publication	4	0.650%
☐	Meeting Abstract	2	0.325%
☐	Art Exhibit Review	1	0.163%
☐	Book	1	0.163%

Analyze Data Table

Yapay zekâ ve gazetecilik alanındaki ilk çalışma, 1992 yılında T. Richards ile L. Richards tarafından yayımlanan "Database Organization for Qualitative Analysis: The NUDIST™ System" başlıklı çalışmadır. 2024 yılına doğru gelindikçe, bu alanda yapılan çalışmaların sayısında belirgin bir artış gözlemlenmektedir. Bu artışın nedenlerinden bazıları; 28 Ekim 2021 tarihinde Facebook'un şirket ismini "Meta" olarak değiştirmesi ve diğer teknoloji firmalarının da art arda metaverse ve yapay zekâya ilişkin çalışmalarını kamuoyuna duyurmasıdır [21]. Yapay zekâya yönelik ilginin artmasıyla birlikte, akademik alanda da bu konudaki yayınların sayısında yükseliş olduğu söylenebilir.

Bu doğrultuda, 5 Mart 2024 tarihi itibarıyla Web of Science veritabanında yer alan akademik çalışmalar incelendiğinde; özellikle 2021 yılından sonra yapay zekâ ve gazetecilik kesişiminde yapılan yayınların sayısının gözle görülür biçimde arttığı belirlenmiştir. Yıllara göre yayın sayıları incelendiğinde; 2020 yılında 93, 2021 yılında 91 olan yayın sayısı, 2022 yılında 128'e, 2023 yılında ise 155'e yükselmiştir. 2024 yılı itibarıyla, yalnızca 5 Mart'a kadar geçen sürede 23 akademik yayın yapıldığı tespit edilmiştir. Bu veriler, yapay zekânın gazetecilik alanında önemli bir yere sahip olduğunu ve hem akademik hem de sektörel düzeyde köklü değişim ve dönüşümlere yol açtığını ortaya koymaktadır.

Tablo 2: Çalışmaların yıllara göre dağılımı.



Web of Science veritabanında yer alan yayınların indeks türlerine göre dağılımı incelendiğinde; ilk sırada 261 kayıt ile Social Sciences Citation Index (SSCI), ikinci sırada 149 kayıt ile Emerging Sources Citation Index (ESCI), üçüncü sırada ise Science Citation Index Expanded (SCI-EXPANDED) yer almaktadır. Bu indeks türlerine ilişkin ayrıntılı bilgiler Tablo 2'de sunulmuştur.

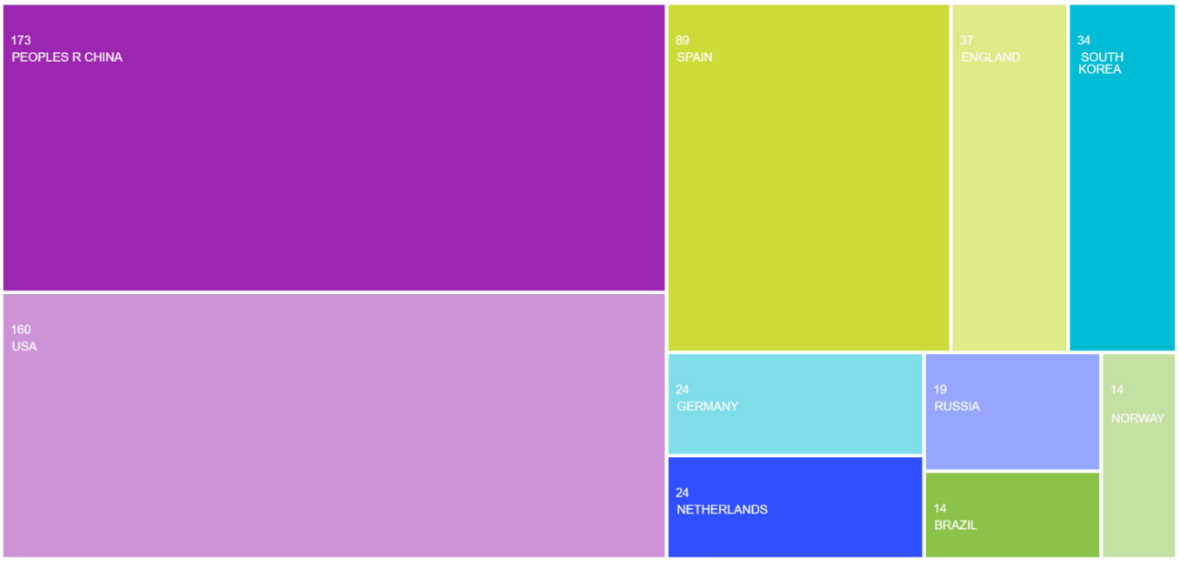
Tablo 3: çalışmaların Web of Science Index türünde yayın sayısı

Select All ☐	Field: Web of Science Index	Record Count	% of 615
☐	Social Sciences Citation Index (SSCI)	261	42.439%
☐	Emerging Sources Citation Index (ESCI)	149	24.228%
☐	Science Citation Index Expanded (SCI-EXPANDED)	131	21.301%
☐	Conference Proceedings Citation Index – Science (CPCI-S)	91	14.797%
☐	Book Citation Index – Social Sciences & Humanities (BKCI-SSH)	20	3.252%
☐	Conference Proceedings Citation Index – Social Science & Humanities (CPCI-SSH)	18	2.927%
☐	Book Citation Index – Science (BKCI-S)	14	2.276%
☐	Arts & Humanities Citation Index (A&HCI)	10	1.626%

Analyze Data Table

Akademik yayın sayısı bakımından en fazla katkı sağlayan üniversite, 24 yayınla State University System of Florida olmuştur. Onu, 22 yayın ile İspanya'dan Universidade de Santiago de Compostela ve üçüncü sırada ise University of Florida takip etmektedir. Toplam 615 akademik yayın ülke ve bölge bazında incelendiğinde; 173 yayının Çin Halk Cumhuriyeti, 160'ının Amerika Birleşik Devletleri ve 89'unun İspanya merkezli akademisyenler tarafından üretildiği belirlenmiştir. Bu veriler, Çin ve Amerika'nın bu alana önemli ölçüde yoğunlaştığını göstermektedir.

Üniversite ve ülke/bölge düzeyindeki dağılımlar karşılaştırıldığında; Şekil 3'te de görüldüğü üzere, birinci ve üçüncü sırada yer alan üniversitelerin Amerika Birleşik Devletleri'ne ait olduğu, ikinci sıradaki üniversitenin ise İspanya'dan olduğu görülmektedir. Öte yandan, Çin üniversiteleri ilk üç sırada yer almamasına rağmen, toplam yayın sayısı açısından Çinli akademisyenlerin bu alana yoğun biçimde katkı sunduğu anlaşılmaktadır.



Şekil 2: Çalışmaların Ülke ve Bölgelere göre dağılımı

Üniversite ve ülke/bölge bazlı analizlere ek olarak, bu yayınlara fon sağlayan kuruluşlar incelendiğinde konuya daha geniş bir perspektiften bakmak mümkün olmaktadır. Yapay zekâ ve gazetecilik alanındaki çalışmalara en fazla fon sağlayan kuruluşun, 31 yayın ile National Natural Science Foundation of China (NSFC) olduğu belirlenmiştir. Bu durum, Şekil 2'de sunulan verileri destekler niteliktedir. Fon sağlayan kuruluşlar arasında ikinci sırada European Union (EU), yani Avrupa Birliği yer almakta; üçüncü sırada ise İspanya Hükümeti bulunmaktadır.

Bu veriler doğrultusunda İspanya'nın, yayın üretiminde ikinci sırada yer almasının yanı sıra, fonlanan kuruluşlar arasında da üçüncü sırada olması dikkat çekicidir. Bu durum, İspanya'nın yapay zekâ ve gazetecilik alanına yoğun bir akademik ilgi gösterdiğini ortaya koymaktadır. Ayrıca, Avrupa Birliği her ne kadar yayın üretiminde ilk sıralarda yer almasa da bu alandaki çalışmaları fonlayarak önemli bir rol üstlenmektedir.

Elde edilen en önemli bulgulardan biri ise Çin'in bu alandaki baskın konumudur. Çin'in bu pozisyonu, Kothari ve Cruikshank'in [22] çalışmasında da vurgulandığı üzere, yapay zekâyı haber odalarında kullanmaya başlayan ilk ülkelerden biri olmasına dayanmaktadır. Çin'in Xinhua News Agency adlı haber ajansında yapay zekâ ile haber üretilip sunulması, bu konudaki öncülüğünü pekiştirmektedir. Bu durum, Şekil 3 ve Tablo 3'te sunulan Çin kaynaklı akademik yayınların sayısını açıklayıcı niteliktedir.

Tablo 4: Çalışmaları destekleyen kuruluşların dağılımı

Select All ☐	Field: Funding Agencies	Record Count	% of 615
☐	NATIONAL NATURAL SCIENCE FOUNDATION OF CHINA NSFC	31	5.041%
☐	EUROPEAN UNION EU	23	3.740%
☐	SPANISH GOVERNMENT	17	2.764%
☐	NATIONAL RESEARCH FOUNDATION OF KOREA	14	2.276%
☐	MINISTRY OF EDUCATION MOE REPUBLIC OF KOREA	11	1.789%
☐	MARIE CURIE ACTIONS	8	1.301%
☐	FUNDAMENTAL RESEARCH FUNDS FOR THE CENTRAL UNIVERSITIES	7	1.138%
☐	EUROPEAN COMMISSION JOINT RESEARCH CENTRE	5	0.813%
☐	EUROPEAN RESEARCH COUNCIL ERC	5	0.813%
☐	NATIONAL KEY R D PROGRAM OF CHINA	5	0.813%

5 Mart 2024 tarihi itibarıyla Web of Science (WOS) veritabanında, yapay zekâ ve gazetecilik alanında en fazla yayın yapan yazarlar incelendiğinde; ilk sırada 10 yayın ile Joung-Woo Hong, ikinci sırada 9 yayın ile Saro Lee ve üçüncü sırada 8 yayın ile Fabrizio Marconi yer almaktadır. En çok yayın yapan ilk 10 yazara bakıldığında, Çin Halk Cumhuriyeti kökenli birçok araştırmacının listede yer aldığı dikkat çekmektedir. Bu durum, Çin'in bu alana olan akademik ilgisini ve yoğunluğunu bir kez daha ortaya koymaktadır.

Şekil 3: En fazla yayın yapan yazar dağılımı

Bu veriler ışığında, Çinli akademisyenlerin toplumsal süreçlerde giderek daha fazla karşılaştıkları gazetecilikteki yapay zekâ uygulamalarını akademik düzeyde de yoğun biçimde araştırdıkları söylenebilir.

Ortak Yazar Analizi

Web of Science veritabanındaki akademik çalışmalar, VOSviewer yazılımı ile analiz edilmiş; analiz türü olarak "Co-Authorship" (ortak yazarlık), analiz birimi olarak ise "Authors" (yazarlar) seçilmiştir. Analiz kapsamında, en az bir yayını bulunan ve en az bir atıf almış yazarlar dâhil edilmiştir. Bu kriterler doğrultusunda 94 yazarın yer aldığı bir ortak yazarlık ağı oluşturulmuştur. Haritalandırma sonucunda yazarlar arasındaki iş birliği bağlarının güçlü

olduğu ve bu bağların kopmadığı bir yapı ortaya çıkmıştır. Bu yapıda, toplam 11 yazarın ağ içerisinde daha merkezi ve ön planda yer aldığı görülmektedir. Bu yazarlar arasında bir yayını birden fazla kişiyle çalışan akademisyenler şunlardır; 7 kişi ile: Duc Tien Dang-nguyen, Enrico Motta, Andreas I. Opdahl, Vinay Setty, Bjornar Tessem, Eivind Throndsen, Christoph Trattner, Are Tverberg'dir. 6 kişi ile çalışanlar: Ariadna, Angulo-Brunet, Mabel Machadom, Beatriz Elena Marin Ochoa, Maria Mendoza Michilot, Claudio Guillermo Rosell Arce, Maria T. Soto-Sanfiel ve Adriana Ibiti'dir. 4 kişi ile: Carlos Toural-Bran, Miguel Tunez-Lopez ve Maria Jose Ufarte Ruiz'dir. Şekil 4 incelendiğinde, yukarıda belirtilen veriler görsel olarak daha açık ve anlaşılır bir biçimde ortaya konmaktadır. Haritada kullanılan renkler, yazarların oluşturduğu kümeleri ve aralarındaki iş birliği düzeyini temsil etmektedir. Bu görselleştirme sayesinde, yazarlar arasındaki ortak çalışma ilişkileri net bir biçimde gözlemlenebilmektedir. Bu bağlamda, Ruiz ve Jose, Rubio ve Mauricio, Túñez-López ve Miguel ile Toural-Bran ve Carlos'un, diğer yazarlara kıyasla daha sık ortak çalışmalar yürüttükleri tespit edilmiştir.

Şekil 4: Çalışmaların ortak yazar analizi

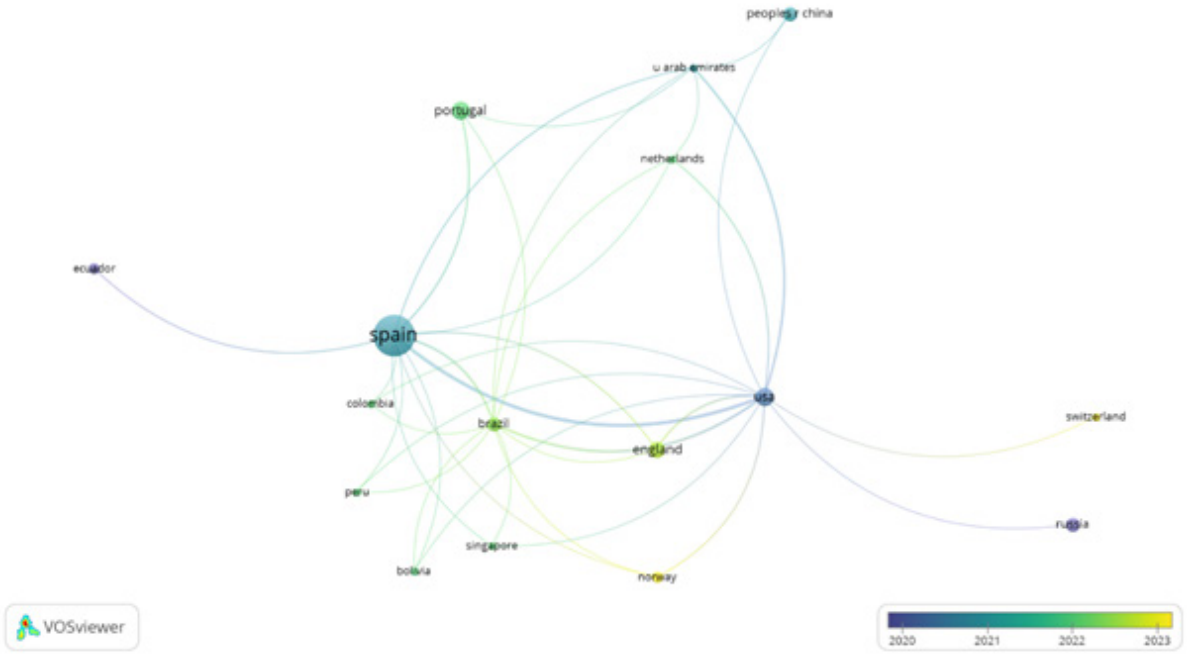
Yazar Atıf Analizi

Yapay zekâ ve gazetecilik alanındaki yazar etki ağını incelemek için Web of Science (WOS) veritabanındaki çalışmalar VOSviewer programında analiz edilmiştir. "Citation" (atıf) analiz tipi ve "Authors" (yazar) analiz birimi kullanılarak gerçekleştirilen değerlendirmede, en az 1 çalışma yayınlamış ve minimum 1 atıf almış yazarlar filtrelenmiştir. Bu kriterler doğrultusunda başlangıçta 94 yazarlık bir veri kümesi elde edilmiştir. Yazarlar arasındaki atıf ilişkilerinin bütünlüğünü koruyacak şekilde oluşturulan bilimsel haritalandırmada ise, birbiriyle anlamlı bağlantılar kuran 70 yazardan oluşan bir ağ yapısı tespit edilmiştir. Bu bağlantılı ağ, alandaki temel etki kümelerini ve iş birliği örüntülerini görselleştirmek açısından daha tutarlı bir temsil sunmaktadır. Bu yazarların atıf sayıları sıralandığında 84 atıfı olan yazarlar: Andrea I. Guzman, Seth C. Lewis ve Thomas R. Schmid'dir. İkinci olarak 54 atıfı olan yazarlar: John D. Banja, Carolyn C. Meltzer ve Nabile M. Safdar'dir. Üçüncü olarak 42 atıfı olan Jonathan Stray'dir. Bu atıf sayıları bu alanda çalışacak olan akademisyenlere bu kişilerin yayınlarının okunması gerektiğini ifade etmektedir.

Şekil 5: Çalışmaların yazar atıf analizi

Ülke Atıf Analizi

Yapay zekâ ve gazetecilik alanındaki uluslararası etki dinamiklerini incelemek için VOS-viewer yazılımında "Citation" (atıf) analiz tipi ve "Countries" (ülke) analiz birimi kullanılarak kapsamlı bir değerlendirme gerçekleştirilmiştir. Analize dahil edilmek için ülkelerin en az 1 çalışma yayınlamış ve minimum 1 atıf almış olma kriterleri belirlenmiştir. Bu parametrelerle yapılan ilk analizde 17 ülke tespit edilmiş, ancak ülkeler arası atıf bağlantılarının bütünlüğünü koruyacak şekilde oluşturulan haritalandırmada 16 ülkelik bir ağ yapısı daha net sonuçlar sağlamıştır. Analiz sonuçları incelendiğinde: İspanya, 24 farklı türde çalışma ve toplamda 226 atıf ile hem yayın çeşitliliği hem de bilimsel etki açısından lider konumda bulunmaktadır. Amerika Birleşik Devletleri, 5 türde yayın ve 196 atıf almasına karşın, 32 uluslararası bağlantı ile alandaki en kapsamlı iş birliği ağına sahip ülke olarak öne çıkmaktadır. Bu durum, ABD'nin küresel bilgi paylaşımı ve araştırma iş birliği noktasında merkezi bir rol üstlendiğini göstermektedir. Brezilya, 3 çalışma, 24 atıf ve 18 uluslararası bağlantı ile üçüncü sırada yer almakta ve alanda gelişmekte olan önemli bir aktör olduğunu kanıtlamaktadır. Ayrıca, İsviçre ve Norveç gibi ülkelerin bu araştırma alanında yakın dönemde yayın faaliyetlerine başladıkları ve gelecekte alan yazına katkıda bulunma potansiyeli taşıdıkları gözlemlenmiştir.



Şekil 6: Ülkelerin yıllara göre yayın dağılımı

Kurum Atıf Analizi

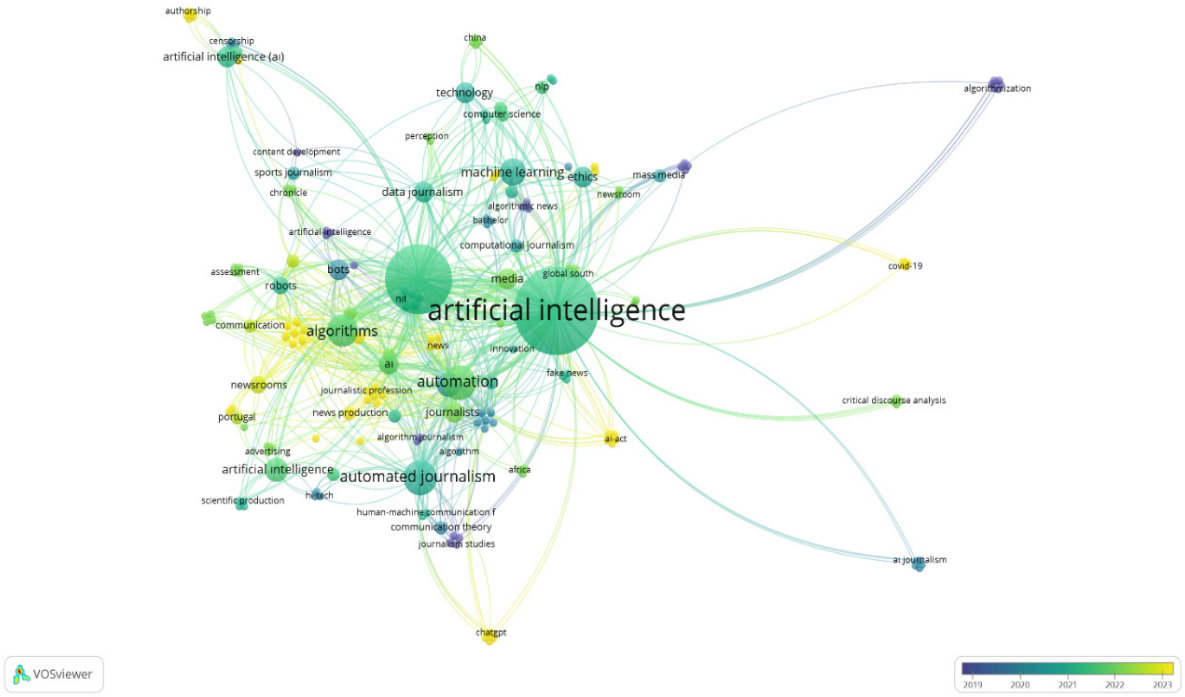
Yapay zekâ ve gazetecilik alanındaki kurumsal etki dinamiklerini belirlemek için VOS-viewer yazılımında "Citation" (atıf) analiz tipi ve "Organizations" (Organizasyon) analiz birimi kullanılarak detaylı bir analiz gerçekleştirilmiştir. En az 1 çalışma yayınlamış ve minimum 1 atıf almış olma kriterleriyle yapılan taramada ilk aşamada 54 organizasyon tespit edilmiştir. Kurumlar arası atıf bağlantılarının bütünlüğünü koruyacak şekilde oluşturulan haritalandırmada ise 45 organizasyondan oluşan daha tutarlı bir ağ yapısı ortaya çıkarılmıştır. Bu analizde özellikle üç kurum dikkat çekmektedir: Universidade De Santiago De Compostela (İspanya), University of Castilla-La Mancha (İspanya), University of Beira Interior (Portekiz).

Ortaya çıkan veriler, İspanya'daki akademik kurumların ve hükümet destekli araştırmaların yapay zekâ ve gazetecilik kesişiminde önemli bir yoğunlaşma gösterdiğini doğrulamaktadır. Bu durum, literatürdeki diğer çalışmalarla da tutarlılık göstermektedir. Albarran, Molina ve Gijon tarafından yayınlanan "Perception of Artificial Intelligence in Spain" başlıklı araştırmada [23], İspanya toplumunun %63'ünün yapay zekânın iş hayatında insanlara fayda sağlayacağına inandığı belirtilmektedir. Ayrıca, Perez-Seijo, Barbosa ve Vicente'nin "Artificial intelligence in journalism: case study of the Spanish, Portuguese and Brazilian news media systems" başlıklı çalışması [24], İspanya'da yapay zekâ ve gazetecilik entegrasyonunun pratik uygulamalarını ortaya koymaktadır. Bu çalışmaya göre, 2015 yılında otomatik içerik üretimi için özel bir şirket kurulmuş ve 2016 itibariyle yerel ve ulusal gazetelerin gazetecilik süreçlerinde yapay zekâ teknolojilerini aktif olarak kullanmaya başladıkları belgelenmiştir. Bu bulgular, İspanya'nın yapay zekâ ve gazetecilik alanındaki kurumsal liderliğinin hem akademik araştırma yoğunluğu hem de pratik uygulamalar açısından önemini vurgulamaktadır.

Şekil 7: Çalışmalarının organizasyonların destekleme oranı

Anahtar Sözcük Analizi

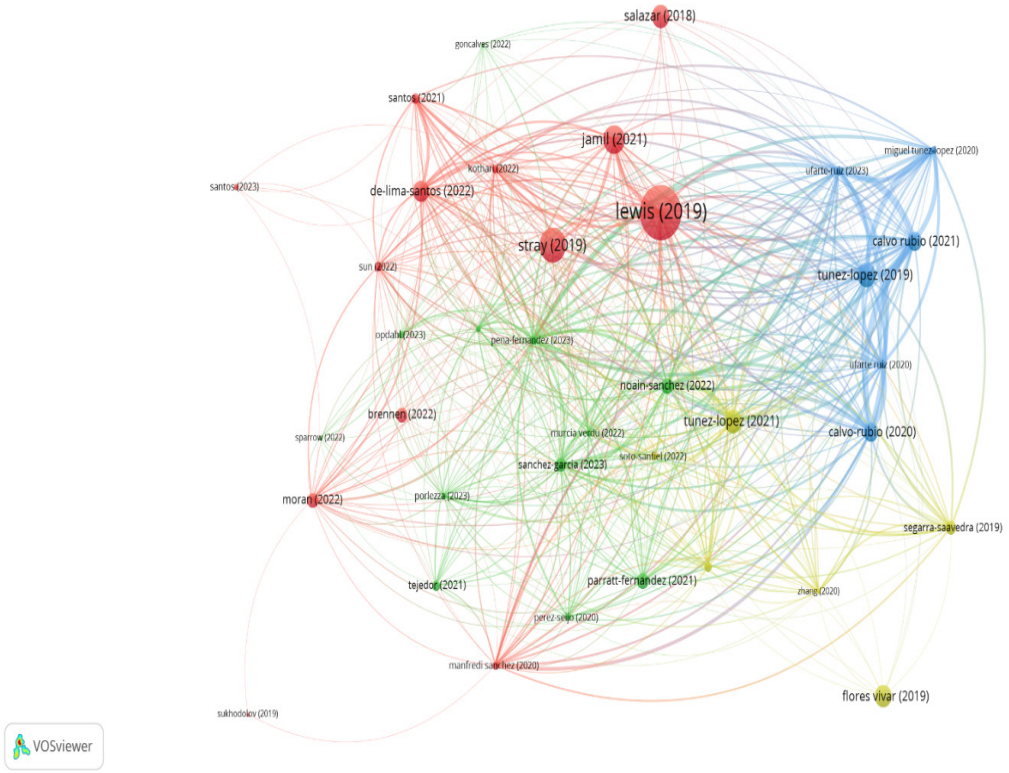
Yapay zekâ ve gazetecilik alanındaki kavramsal yapıyı ve tematik eğilimleri ortaya çıkarmak amacıyla VOSviewer yazılımında "Co-occurrence" (Eş-oluşum) analiz tipi ve "Author keywords" (yazar anahtar kelimeleri) analiz birimi kullanılarak kapsamlı bir inceleme gerçekleştirilmiştir. En az 1 kez kullanılmış olma kriteri ile yapılan taramada toplam 185 anahtar kelime tespit edilmiştir. Eş-oluşum analizine göre alandaki en baskın kavramlar şu şekilde sıralanmaktadır: Yapay Zekâ: 37 eş-oluşum sıklığı ve 217 bağlantı ile alanın merkezindeki temel kavram konumundadır. Gazetecilik: 26 eş-oluşum sıklığı ve 147 bağlantı ile ikinci ana tema olarak öne çıkmaktadır. Otomasyon: 9 eş-oluşum sıklığı ve 69 bağlantı ile üçüncü sırada yer almaktadır. Haritalandırmada (Şekil 6), literatürde yeni belirmeye başlayan ve henüz yoğun araştırma birikimi oluşmamış tematik alanlar da tespit edilmiştir. Bu yeni gelişen araştırma alanları şunları içermektedir: "Dijital Yerel Medya", "Haber Üretimi", "Dijital Uçurum", "Spor Gazeteciliği", "Veri Bilimi", "Üniversite Eğitimi", "Katılımcı Gazetecilik", "Etik Riskler", "ChatGPT", "Yeni Medya". Bu anahtar kelimelerin harita üzerindeki konumlandırılması ve renk kodlamaları, söz konusu kavramların yapay zekâ ve gazetecilik kesişimindeki araştırma literatüründe henüz yeni yeni yer almaya başladığını ve bu alanlarda gelecekte daha fazla araştırma potansiyeli bulunduğunu göstermektedir. Görsel temsilde bu anahtar kelimelerin kapladığı alanın daha küçük olması ve merkezi temalardan uzak konumlanmaları, bu konuların henüz gelişim aşamasında olduğunu açıkça ortaya koymaktadır.



Şekil 8: Çalışmalarında kullanılan anahtar kelime dağılımı

Metinlerin Bibliyografik Eşleşme Analizi of Documents

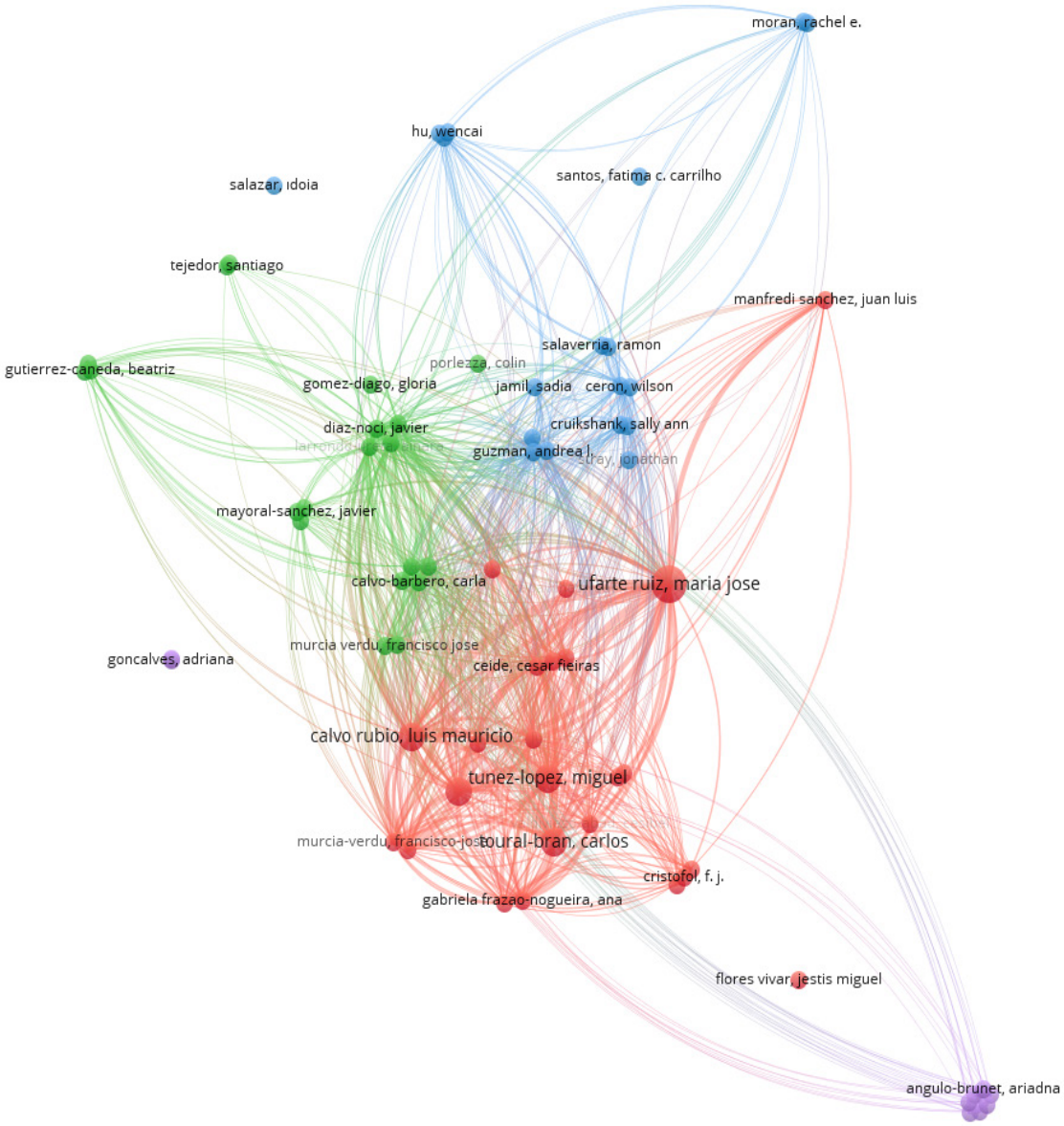
Metinlere bibliyografik eşleme analizi yapmak için VosWiever aracılığıyla analiz tipi olarak "Bibliographic Coupling" (Bibliyometrik bağlantı) analiz birimi ise "Documents" (doküman) olarak belirlenmiştir. Bu metin eşlemeleri arasında en az 1 doküman yapmış olma şartıyla bir analiz yapıldığında 41 sonuç çıkmaktadır. Bu metinlerin bibliyografik eşelemelerinin bağların kopmadığı bir haritalandırma yapıldığında 37 sonuç tespit edilmiştir. Buradan hareketle Şekil 7'de de görüldüğü üzere Seth Lewis, Nabile Safdar ve Jonathan Stray yayınlarına daha fazla atıf yapıldığı görülmektedir. Lewis'e 2019 "Automation, Journalism, and Human-Machine Communication: Rethinking Roles and Relationships of Humans and Machines in News" adlı yayınına 84, Safdar'ın 2020 "Ethical considerations in artificial intelligence" adlı yayınına 54, Stray'in 2019 "Making Artificial Intelligence Work for Investigative Journalism" adlı yayınına ise 42 alıntı yapılmıştır. Buradan hareketle Lewis'in yayını bu alanda okunması gereken bir yayın olduğu ortaya çıkmaktadır. Yazarların birbiriyle olan bağlantılarına bakıldığında Luis Mauricio Calvo Rubio'nun 2021 "Artificial intelligence and journalism: Systematic review of scientific production in Web of Science and Scopus (2008-2019)" adlı yayını 398 bağlantıya, María José Ufarte Ruiz'in 2023 yayını 346 bağlantıya ve Luis Mauricio Calvo Rubio 2020 "Perception of teachers, students, innovation managers and journalists about the use of artificial intelligence in journalism" adlı yayını 323 bağlantıya sahip olduğu tespit edilmiştir. Buradan hareketle Calvo Rubio'nun yayını diğer yazarlar tarafından ve kendi içeriğinde diğer yazarlardan bağlantılar barındırdığı ortaya çıkmıştır. Bu çerçevede bu harita yapay zekâ ve gazetecilik üzerine çalışma yapacak bir kişiye kimleri okunması gerektiğini anlatmaktadır.



Şekil 9: Çalışmaların metinlerinde bibliyometrik eşleşme analizi dağılımı

Yazarların Eş Atıf Analizi

Yazarların eş atıf (bibliyometrik bağlantı) analizi yapmak amacıyla VOSviewer yazılımı kullanılmış; analiz tipi olarak "Bibliographic Coupling" (Bibliyometrik Bağlantı) ve analiz birimi olarak ise "Authors" (Yazarlar) seçilmiştir. Bu analizde, yazarlar arasında en az bir doküman üretmiş ve en az bir diğer yazar tarafından atıf yapılmış olma şartı aranmıştır. Sonuç olarak, 94 yazarın yer aldığı bir analiz gerçekleştirilmiş ve bu yazarlar arasındaki bağların kopmadığı bir haritalandırma yapıldığında 86 yazar daha net bir şekilde belirlenmiştir. Analiz sonuçları incelendiğinde, Maria Jose Ufarte Ruiz ve Maria Jose'nin 3 yayınla bu alanda en fazla içerik üreten ve aynı zamanda diğer yazarlarla en fazla bağı olan yazarlardan oldukları tespit edilmiştir.



Şekil 10: Çalışmalarında yazarların eş atıf analiz dağılımı

Bu veriler ışığında, İspanya Hükümeti ve üniversitelerinin, 1992'den 2024 yılına kadar olan süreçte yapay zekâ ve gazetecilik alanında önemli sayıda akademik çalışma ürettikleri tespit edilmiştir. Bunun yanı sıra, çalışmanın başında da belirtildiği üzere, Çin'in de bu alandaki akademik katkıları yoğun olarak devam etmektedir. Web of Science verileri incelendiğinde, Çin'in 2021 yılı itibarıyla daha sık yayınlar yapmaya başladığı, ancak 2022 yılı itibarıyla diğer ülkelerle kıyaslandığında bu alandaki üretiminin azalmaya başladığı gözlemlenmektedir.

Sonuç

Bu çalışma, yapay zekâ teknolojilerinin gazetecilik alanında yarattığı dönüşümün akademik literatürde nasıl ele alındığını bibliyometrik analiz yöntemiyle incelemiştir. Web of Science veri tabanında "Artificial Intelligence" ve "Journalism" anahtar kelimeleriyle taranan yayınlar, VOSviewer yazılımı kullanılarak analiz edilmiş ve bu iki alan arasındaki kesişimin zaman içindeki gelişimi, yoğunlaştığı disiplinler, öne çıkan araştırmacılar ve ülkeler çerçeve-

sinde haritalandırılmıştır.

Analizler sonucunda, bu alandaki ilk çalışmaların 1992 yılında başladığı, ancak özellikle 2021 yılından itibaren dikkat çekici bir artış gösterdiği belirlenmiştir. 5 Mart 2024 itibarıyla toplam 615 akademik yayın tespit edilmiş, bunların 471'inin makale türünde olduğu görülmüştür. Çalışmaların ağırlıklı olarak "İletişim", "Bilgisayar Bilimleri" ve "Mühendislik" alanlarında yoğunlaştığı saptanmıştır.

Yapılan anahtar kelime analizinde, "Artificial Intelligence", "Journalism" ve "Automation" terimlerinin en sık kullanılan kavramlar olduğu ortaya çıkmıştır. Atıf sayıları dikkate alındığında Lewis, Safdar ve Stray gibi araştırmacıların alana önemli katkılar sunduğu anlaşılmaktadır. Ülke düzeyinde değerlendirildiğinde ise İspanya'nın bu alanda öncü bir konumda olduğu, Çin'in ise 2021 yılından itibaren önemli bir ivme kazandığı tespit edilmiştir. Çin'in robot gazeteci ve yapay zekâ destekli haber sunucuları gibi somut uygulamalarla alanda sadece akademik değil, sektörel anlamda da aktif rol aldığı görülmektedir. En fazla yayın yapan ülkeler sırasıyla Çin, Amerika Birleşik Devletleri ve İspanya olarak öne çıkmıştır.

Ayrıca, ortak yazarlıklı yayınlarda en yüksek atıf alan yazarlar arasında Guzman, Lewis, Schmid, Banha, Meltzer, Safdar ve Stray yer almaktadır. Bu bulgular, gelecekte bu alanda çalışacak araştırmacılar için etkili yazarlara, yaygın kullanılan anahtar kelimelere ve hangi ülkelerin yapay zekâ destekli gazetecilik araştırmalarına yatırım yaptığına dair önemli ipuçları sunmaktadır.

Sonuç olarak, yapay zekâ ile gazetecilik arasındaki etkileşimin giderek güçlendiği ve bu alandaki akademik üretimin çok disiplinli bir yapıya büründüğü görülmektedir. Bu durum hem akademik hem de sektörel düzeyde yeni araştırma alanlarının ve uygulama modellerinin ortaya çıkmasına zemin hazırlamaktadır.

REFERANSLAR

[1] Pawlik, J. (2000). The impact of technology on journalism. Journalism Studies, 1(2), 229-237.

[2] Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication: A human–machine communication research agenda. New media & society, 22(1), 70-86.

[3] Latar, N. L. (2015). The robot journalist in the age of social physics: The end of human journalism?. The new world of transitioned media: Digital realignment and industry transformation, 65-80.

[4] Carlson, M. (2015). "The Robotic Reporter". Digital Journalism, 3(3), 416-431. https://doi.org/10.1080/21670811.2014.976412.

[5] Latar, N. L. (2018). Robot Journalism: Can Human Journalism Survive?, New Jersey: World Scientific.

[6] Miroshnichenko, A. (2018). AI to Bypass Creativity. Will Robots Replace Journalists? (The Answer Is "Yes"). Information 9 (7), 183.

[7] Seager, S. (2017, Temmuz 14). No, It's Not Fake News, It's Robot-Written News.

[8] Copeland, B. J. (2004). Alan Turing 1912–1954. B. J. Copeland The Essential Turing: Seminal

[9] Whittaker, J. (2019). Tech Giants, Artificial Intelligence, and the Future of Journalism. Routledge.

[10] Castro, D., & New, J. (2016). https://datainnovation.org/2016/10/the-promise-of-artificialintelligence/. https://datainnovation.org/: https://datainnovation.org/2016/10/the-promise-of-artificialintelligence/, (Erişim tarihi: 7 Mart 2024)

[11] Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication: A human–machine communication research agenda. New media & society, 22(1), 70-86.

[12] Frackiewicz, M. (2023, 6 21). The Impact of AI on Journalism and Media Ethics. TS2:

[13] Guzman, A. L., & Lewis, S. C. (2020). Artificial intelligence and communication: A human–machine communication research agenda. New media & society, 22(1), 70-86.

[14] Dörr, K. N., & Hollnbuchner, K. (2016). Ethical Challenges of Algorithmic Journalism.

[15] Einav, G. (Ed.). (2014). The New World of Transitioned Media: Digital Realignment and Industry Transformation Springer.

[16] Wang, X., Liu, C., & Qi, Y. (2021). Research on new media content production based on artificial intelligence technology. In Journal of Physics: Conference Series (Vol. 1757, No. 1, p. 012062). IOP Publishing.

[17] Allen, N. D., Templon, J. R., McNally, P. S., Birnbaum, L., & Hammond, K. (2010). StatsMonkey: A Data-Driven Sports Narrative Writer. Computational Models of Narrative: Papers from the AAAI Fall Symposium. Evanston: Association for the Advancement of Artificial Intelligence, (F-S 10-04), (www.aaai.org).

[18] Stray, J. (2019). Making Artificial Intelligence Work for Investigative Journalism. Digital Journalism, 7(8), 1-22, https://doi.org/10.1080/21670811.2019.1630289

[19] Moschovakis, Y. N. (2001). What Is an Algorithm? B. Engquist, & W. Schmid, Mathematics Unlimited — 2001 and Beyond (s. 919-936). Springer.

[20] Van Eck, N. J. (2011). Methodological advances in bibliometric mapping of science. Doctoral dissertation, Erasmus Research Institute of Management (ERIM)).

[21] Garfield, E., Pudovkin, A. I. & Istomin, V. S. (2003). Why do we need algorithmic historiography?. Journal of the American Society for Information Science and Technology, 54(5), 400-412.

[22] Kocabay Şener, N. (2021). Facebook Nasıl "Meta" laştı?. Yeni Medya, 2021(11), 174-179.

[23] Kothari, A., & Cruikshank, S. A. (2022). Artificial intelligence and journalism: An Agenda for journalism research in Africa. African Journalism Studies, 43(1), 17-33.

[24] Albarrán, I., Molina, J. M., & Gijón, C. (2020). Perception of artificial intelligence in Spain.

[25] Pérez-Seijo, S., Barbosa, S., & Vicente, P. N. (2023). Artificial intelligence in journalism: case study of the Spanish, Portuguese and Brazilian news media systems. In Blurring Boundaries of Journalism in Digital Media: New Actors, Models and Practices (pp. 261-274). Cham: Springer International Publishing.

YAPAY ZEKÂNIN CEZAİ SORUMLULUĞU: SİSTEMATİK LİTERATÜR HARİTALAMASI ÇALIŞMASI

Cansel Arslan*
Jandarma ve Sahil Güvenlik Akademisi, Güvenlik Bilimler Bölümü, Ankara, Türkiye

Aslıhan Tüfekci
Gazi Üniversitesi, Bilişim Enstitüsü, Ankara, Türkiye
**(İletişimden sorumlu yazar) canselakyuzarslan@gmail.com*

ÖZET

Yapay zekâ teknolojisindeki gelişmeler ve yapay zekâ kullanımının yaygınlaşması hukuk sisteminde çözülmesi gereken yeni sorunlar ortaya çıkarmaktadır. Mevcut yasal düzenlemeler yapay zekâ konusunda yetersiz kalmaktadır. Hukuki açıdan yapay zekânın konumuna yönelik tartışmalar mevcuttur. Bu çalışmanın amacı yapay zekânın cezai sorumluluğu alan yazınını hukuki tartışmaları ve düzenlemeleri değerlendirmek, araştırma boşluklarını tespit etmek ve hukuki düzenlemelere yönelik öneriler sunmaktır. Çalışmada sistematik literatür haritalaması yöntemi kullanılmaktadır. 5 veri tabanından elde edilen 41 incelenerek alan yazınında yapay zekânın cezai sorumluluğu değerlendirmektedir. Çalışma yapay zekânın cezai sorumluluğu hakkındaki problemleri, çözüm önerilerini ortaya koymaktadır. Alan yazında yapay zekânın cezai sorumluluğu konusundaki benzer problemler yer almaktadır. Ancak yapay zekânın hukuki konumu açısından çeşitli görüşler bulunması ve çözüm önerilesilerinin farklılık göstermesi sebebiyle fikir birliğinin olmadığı tespit edilmiştir.

Anahtar Kelimeler: Yapay zekâ, cezai sorumluluk, hukuki statüsü.

Giriş

Yapay zekâ teknolojilerinde büyük gelişmeler yaşanmaktadır. Bu gelişmelerle birlikte yapay zekânın kullanımı hem endüstriyel hem de günlük yaşamda yaygınlaşmaktadır. Her geçen gün yapay zekâ sağlayıcılıkları ve kullanıcılarında artmaktadır. Bankacılık sisteminden bireysel araçlara kadar birçok alanda yapay zekâ kullanılmaktadır. Yaşamı kolaylaştıran, işgücü tasarrufu sağlayan yapay zekâ hayatın önemli bir parçası haline gelmiştir. Bu durum yapay zekânın yol açacağı sorunları ve bu konudaki endişeleri beraberinde getirmiştir. Yapay zekânın suçta kullanımı veya suç teşkil eden eylemlerde bulunması bu sorunlardan bazılardır.

Yapay zekânın suçla ilişkisi çeşitlilik göstermektedir. Yapay zekâ hem suç işlemede kullanılabilir hem de suç teşkil eden olaya sebebiyet verebilir. Sosyal mühendislik saldırılarından güvenlik açıklarının öğrenilmesine kadar birçok açıdan siber suçların gerçekleşmesinde fail yapay zekâyı kullanabilmektedir. Yapay zekâ insan müdahalesi olmadan da istenmeyen davranışalar, kazalar ve suça yol açabilmektedir. Otonom araç kazaları ve adli bilişimde

kullanılan yapay zekâ uygulamalarının önyargılı karar vermesi gibi eylemler yapay zekâyla ilgili suç teşkil eden eylemlere örnektir. Bu durum yapay zekânın hukuki açıdan nasıl ele alınacağına yönelik tartışmalar meydana getirmektedir.

Yapay zekânın neden olduğu suç olaylarında cezai sorumluluk konusunda da alan yazında farklı görüşler mevcuttur. Yapay zekânın öğrenme, karar verme becerilerine sahip olması, hukukta nasıl değerlendirilmesi gerektiğine ilişkin sorulara neden olmaktadır. Hukuk sitemleri işlenen suçlarda irade ve bilinci temel almaktadır. Yapay zekânın bilinçli sahip olması konusunda netlik bulunmadığı ve bu durumun kanıtlanamadığı için yapay zekâ konusundaki değerlendirmeler bilinçten bağımsız şekillenmelidir [1,2]. Konu hakkındaki tartışmalar yapay zekânın hukuki statüye ve cezai sorumluluğa sahip olup olmadığı üzerinedir. Konu hakkındaki görüşler farklılaşma ve çatışma göstermektedir. Bunların yanı sıra cezai sorumluluğu olmaması durumunda kimin sorumlu tutulacağı konusunda da tartışmalar mevcuttur.

Alan yazında yapay zekânın hukuki statüye sahip olması gerektiği tartışmalar bulunmaktadır. Hukuki statüye sahip olmaması gerektiği görüşü temel olarak yapay zekânın bilinçli olmaması ve eylemlerinin bu şekilde değerlendirilmesine dayanmaktadır. Yapay zekânın hukuki statüsü olmadığını savunan görüşler: eşya niteliği, köle konumu, kısıtlı ehliyet veya tam ehliyetsizlik cenin eşdeğerliği şeklindedir [1,3,4,5]. Yapay zekânın hukuki kişiliğe sahip olması gerektiği görüşleri ise yapay zekânın tüzel kişilik, vekil, elektronik kişilik ve euRobiotics görüşüdür [1,3,4,5]. Elektronik kişilik ve euRobiotics kavramları yasal olarak yenilik arz etmektedir [1]. Avrupa parlamentosu "Civil law rules on robotics" raporunda robotlara elektronik kişilik statüsü verilmesi atfedilmiştir [6]. EuRobiotics görüşü yapay zekâyı "yapay insan" olarak kabul etmektedir. Bu durumda insansı özellikleri olmasına rağmen insan olarak değerlendirilmemesi söz konusudur [1].

Bu çalışma yapay zekânın cezai sorumluluğu konusunda alan yazındaki durumu hakkında bilgi sunmak amacındadır. Çalışmada yapay zekânın cezai sorumluluğu hakkındaki problemler ve bu problemlere getirilen çözüm önerileri değerlendirilerek mevcut durum açıklanmaktadır.

Yöntem

Bu çalışma yapay zekânın cezai sorumluluğu konusunda alan yazındaki durumu hakkında bilgi sunmak amacındadır. Çalışmada yapay zekânın cezai sorumluluğu hakkındaki problemler ve bu problemlere getirilen çözüm önerileri değerlendirilerek mevcut durum açıklanmaktadır. Çalışmada sistematik literatür haritalaması tekniği kullanılmıştır. Sistematik literatür haritalaması alan yazında anahtar kelimelerle yapılan aramalarda elde dilen çalışmaların sınıflandırılması ve görselleştirilerek aktarılmasıdır [7,8]. Sistematik literatür haritalamasının belirli aşamaları bulunmaktadır [7]. Bunlar :

- Araştırma sorusunun belirlenmesi
- Aramanın yapılası
- Arama sonuçlarının incelenmesi
- Değerlendirme
- Analiz

Sistematik literatür resimleme tekniği aşamaların gerçekleştirilmesiyle birlikte çalışma konu ve çalışma sorusuna uygun olan kaynaklardan verilerin bir araya getirilmesi ile aktarılmasını sağlamaktadır [7]. Aramada elde edilen tüm çalışmalar araştırmanın kapsamı ve araştırma sorusuna uygun olamayabilir. Bu nedenle belirli kriterlere göre eleme işlemi uygulanmaktadır. Eleme sonucunda kalan çalışmalardan elde edilen bilgiler görselleştirilerek aktarılmaktadır.

Araştırma sorusunun belirlenmesi

Bu çalışmanın ana sorusu "Yapay zekâ cezai sorumluluk alınyazında nasıl yer almaktadır?" şeklindedir. Ana sorunun yanıtlamasına katkı sağlamak için çalışmaya dair bazı alt sorular oluşturulmuştur. Bu sorular:

S1. Yapay zekânın ceza sorumluluğu konusundaki makalelerde en çok yer alan anahtar kelimler nelerdir?

S2. Yapay zekânın ceza sorumluluğu konusundaki problemler nelerdir?

S3. Yapay zekâyla alakalı suç teşkil eden eylemlerde adaletin sağlanabilmesi için uygulanabilecek çözümler nelerdir?

Veri tabanlarında arama yapılması

Bu çalışmada 5 veri tabanında yapay zekânın cezai sorumluluğuyla ilgili arama yapılmıştır. Bu veri tabanları Science Direct, Sage, Springer Link, JSTOR ve TRdizin olarak belirlenmiştir. Veri tabanlarında yapılan aramalarda hem dünya hem de Türkiye'de yapılan çalışmaların elde edilmesi amaçlandığı için TR dizin'de veri tabanları arasında yer almaktadır. Yönteme uygun olarak veri tabanlarındaki aramalar için anahtar kelimeler belirlenmiştir. Yapay zekâ kısmı için "artificial intelligence" ve kısaltması olan "ai" kelimeleri seçilerek "artificial intelligence or ai" şeklinde birleştirilmiştir. Yapay zekânın sorumluluğu için "criminal liability" arama metnine eklenmiştir. Bu çalışmadaki arama kelimeleri "(artificial intelligence or ai) and criminal liability" şeklinde formalize edilmiştir. TRdizinde Türkçe aramalar yapılacağı için aramalar "yapay zekâ ve cezai sorumluluk" şeklinde gerçekleştirilmiştir. Tablo 1'de görüldüğü üzere yapılan aramalar sonucunda 5 veri tabanında toplam 3246 çalışmaya ulaşılmıştır. Aramalar sonucunda en fazla Sage veri tabanında konu hakkında çalışma bulunmuştur.

Tablo 1: Veri tabalarında arama sonuçları

Veri Tabanı	Yayın sayısı
Science Direct	322
Sage	1412
TR dizin	20
Springer Link	253
JSTOR	1239
Toplam	3246

Çalışmaların elenmesi ve incelenmesi

Veri tabanlarında arama sonucu ulaşılan çalışmalar bazı kriterlere göre değerlendirilir ve değerlendirmelere göre elemeler gerçekleştirilir. Bu araştırmada eleme işlemi iki aşama halinde gerçekleştirilmiştir. İlk aşamada çalışmanın türü, çalışmaya erişim durumu ve çalışmanın farklı veri tabanlarında tekrarlamasına bağlı elemeler gerçekleştirilmiştir.

Veri tabanlarından elde edilen araştırmaların eleme kriterleri ve elenen çalışmaların sayısı Tablo 2'de verilmiştir. Veri tabanlarından aramalar sonucunda ulaşılan 3246 çalışma makale türündeki yayınlar filtrelenmiş, diğer türlerdeki yayınlar (kitap, bildiri vb.) bu çalışmadan çıkarılmıştır. Filtreleme sonucunda ulaşılan 806 makale erişim durumuna göre değerlendirilerek, erişim sınırı bulunan makaleler elenmiştir. Bu eleme sonucunda tam erişim sağlanan 510 makale kalmıştır. Kalan makalelerin 2 tanesi çalışmaya dahil edilen veri tabanlarında tekrar edilmesi sebebiyle elenmiştir. Çalışmada ilk eleme aşaması tamamlanarak 508 makale ile ikinci aşama gerçekleştirilmiştir.

Tablo 2: Eleme kriterlerine bağlı elenen yayın sayısı

Eleme Kriterleri	Yayın Sayısı
Konuyla tam ilgisi bulunmayan	467
Tam erişim olmayan	296
Makale türünde olmayan	2440
Tekrar eden	2
Toplam	3205

Çalışmanın ikinci aşamasında 508 makalenin ÖZETi, anahtar kelimeleri ve içeriği incelenmiş ve yapay zekânın cezai sorumluluğuyla ilgili olmayan 296 makale elenmiştir. Eleme aşamaları tamamlanmış ve 41 makale çalışmaya dâhil edilmiştir. Çalışmaya veri tabanlarından dâhil edilen makalelerin sayısı Tablo 3'te yer almaktadır.

Tablo 3: Elemeler sonucunda çalışmada kullanılan makalelerin veri tabanlarına dağılımı

Veri Tabanı	Yayın sayısı
Science Direct	11
Sage	4
TR dizin	11
Springer Link	8
JSTOR	7
Toplam	41

Çalışmaların anahtar kelimelerinin ve ÖZETlerin değerlendirilmesi

Bu çalışmaya dahil edilen 41 makalelerin anahtar kelimeleri bir araya getirilmiştir. Türkçe ve İngilizce makalelerin yer alması nedeniyle tüm anahtar kelimeler Türkçeleştirilmiştir. Tablo IV'te birden fazla geçen anahtar kelimler ve tekrar sayısı yer almaktadır.

Tablo 4: Tekrar eden anahtar kelimeler ve tekrar sayısı

Anahtar Kelime	Tekrar sayısı
Yapay zekâ	22
Robotik	9
Otonom araç	5
Cezai sorumluluk	5
AI etiği	4
Ürün sorumluluğu	4
Hesap verebilirlik	3
Hukuki sorumluluğu	3
Yükümlülük	3
İhmal	3
Yasa uygulama sorunları	2

Makalelerin hepsinde anahtar kelimler bulunmamaktadır, 37 makalede yer alan anahtar kelimeler bir araya getirilerek frekans analizi gerçekleştirilmiştir. Makalelerde yer alan toplam anahtar kelime sayısı 190'dır. Çalışmalarda 138 farklı anahtar kelime yer almaktadır. En fazla geçen "yapay zekâ" anahtar kelimesidir.

Bulgular

Bu çalışmada elemeler sonucunda kalan makalelerin tam metin incelemeleri gerçekleştirilmiştir. Dâhil edilen en eski makalenin 2013 yılında yayınlanmıştır. Makalelerin dili Türkçe (n=11) ve İngilizce (n=30) şeklindedir. Makalelerin %54'ü tek yazarlıdır (n=22). En fazla yayın Birleşik Krallık'ta yayınlanan dergilerde olmuştur (n=14). Makalelerin yayımladıkları dergilerin %17'si Türkiye menşeilidir (n=11). İncelemeler sonrasında araştırma sorularına cevap verebilecek bilgiler tablo ve grafiklerle aktarılmıştır.

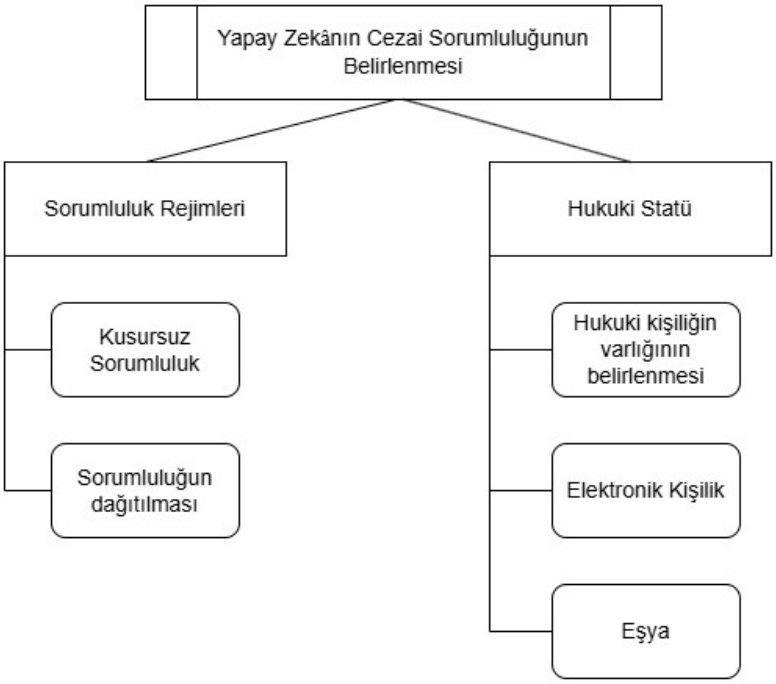
Makalelerde yapay zekânın cezai sorumluluk belirsizliği, yasal düzenlemelerin eksikliği, etik problemler, hukuki statü belirsizliği ve niyet belirsizliği sorunlarına değinilmiştir. Şekil 1'de makalelerde yer verilen yapay zekânın cezai sorumluluğu konusundaki problemlerin dağılım grafiği bulunmaktadır. Makalelerin %80'ninde yapay zekânın cezai sorumluluk belirsizliği sorunu yer almaktadır (n=33). Makalelerde yapay zekâ uygulamaların yarattığı etik problemlerde hesap verilebilirlik ve şeffaflık eksikliği üzerinde durulmuştur. Yapay zekânın niyetinin bilinmemesi, olayda kasıt olup olmadığının net olmaması gibi durumların hukuki problemler taşıdığı bildirilmiştir.

Şekil 1: Makalelerde yer alan yapay zekânın cezai sorumluluğuyla ilgili problemler

Makalelerde yer alan sorunlara çözüm önerileri de yer almaktadır. Çözüm önerileri yapay zekânın cezai sorumluluğunun nasıl değerlendiğine bağlı olarak farklılaşmaktadır. Örneğin yapay zekânın hukuki statüsü olmayıp eşya olarak değerlendirilmesi ve cezai sorumluluğun sahibi/kullanıcısında olduğu kanısında olan çalışmalarla birlikte, yapay zekânın hukuki statüsü olduğu ve elektronik kişi olarak değerlendirilmesi gerektiği önerisini taşıyan çalışmalar da bulunmaktadır. Çözüm önerileri genel olarak yasal çerçevede olmasına rağmen teknik öneriler de bulunmaktadır. Makalelerde yer alan çözüm önerileri:

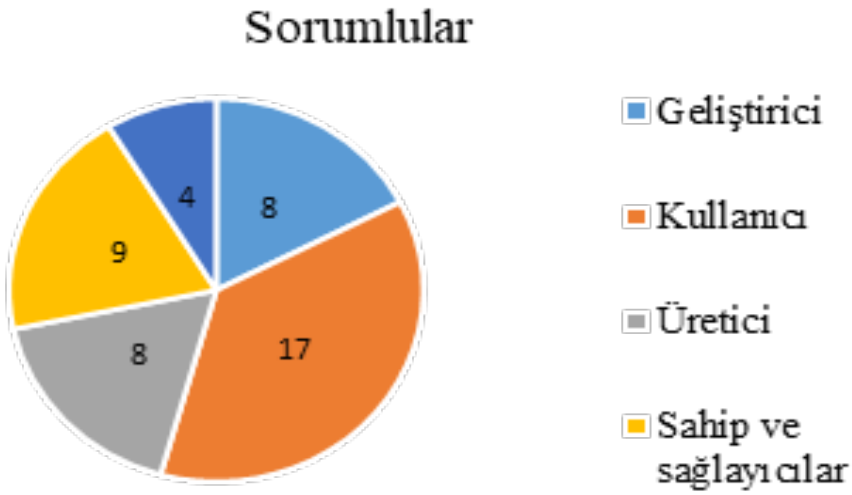
- Yapay zekânın cezai sorumluluğunun netleştirilmesi
- Yasal düzenlemelerin yapılması
- Uluslararası standartların belirlenmesi ve yapay zekâların bu standartlara uygun olarak geliştirilmesi
- Nihai karar mekanizmasının insan olması
- Yapay zekâ üzerinde insan kontrolünün ve sorumluluğunun artırılması
- Yapay zekâ yazılımlarının denetlenmesi

Makalelerde yapay zekâyla ilgili cezai sorumluluk problemlerine getirilen çözüm önerileri Şekil 2'de görülmektedir. Yapay zekânın cezai sorumluluğu sorumluluk rejimlerine yönelik öneriler bulunmaktadır. Kusursuz sorumluluk verilmesi (n= 6) ve sorumluluğun dağıtılması (n= 5) sorumluluk rejimi önerileridir. Yapay zekânın hukuki statüsüne yönelik öneriler de bulunmaktadır. Hukuki statüsü konusunda statünün olup olmaması durumuna göre öneriler ikiye ayrılmaktadır. Hukuki statüsü olmadığında eşya değerlendirilmesi (n=1) ve hukuki statüsü bulunduğunda elektronik kişi (n=4) olarak değerlendirilmesi şeklinde farklılık göstermektedir.



Şekil 2: Yapay zekânın cezai sorumluluğu konusundaki sorunlara çözüm önerileri

Makalelerde yapay zekânın işlediği suçlarda cezai sorumluğunun kim olduğuna yönelik değerlendirmeler mevcuttur. Şekil III'te görüldüğü üzere en fazla sorumluluk atfedildiği kişiler yapay zekâ kullanıcıları olmuştur. Bu sorumluluk atfedilmesinin olaya özgü değişkenlik göstereceği, tek bir sorumlu yerine birden fazla sorumlunun olabileceğine değinilmektedir.



Şekil 3: Yapay zekânın işledikleri suçlarda cezai sorumluluk atfedilen kişilerin dağılımı

Tartışma ve Sonuç

Alan yazında yapay zekânın cezai sorumluğunun mevcut durumunu ortaya koyan bu çalışmada sistematik literatür haritalaması yöntemi kullanılmıştır. Çalışmada 5 veri tabanında yapılan elemeler sonucunda ulaşılan 41 makale detaylı incelenmiş, araştırma sorularına tablo ve grafiklerle yanıt verilmiştir.

"Yapay zekânın ceza sorumluluğu konusundaki makalelerde en çok yer alan anahtar kelimler nelerdir?" sorusunun cevabı ise sırasıyla yapay zekâ, robotik ve otonom araçlardır. Çalışmayı yapay zekânın cezai sorumluluğunu konu alan makaleler oluşturduğu için durumun öznesi (yapay zekâ) ve uygulama alanlarının (robotik ve otonom araçlar) en çok kullanılan anahtar kelimeleri sonuçları öngörülebilirdir. Diğer tekrar eden anahtar kelimelerin hepsi hukuk ile ilgili kelimelerdir. Bu açıdan multidisipliner olan bu çalışmada ulaşılan anahtar kelimeler hukuk ve teknolojiyle ilgilidir. Tekrar eden anahtar kelimeler dahil edilen makalelerin çalışma konularıyla tutarlıdır, dolaysıyla eleme kriterlerin doğru uygulandığı sonucuna ulaşılabilir.

"Yapay zekânın ceza sorumluluğu konusundaki problemler nelerdir?" sorusunun cevabı cezai sorumsuzluğun belirsiz olması yasal düzenlemelerdeki yetersizlikler, etik problemler, hukuki statüsü ve niyet durumunu şeklindedir. Esasen problem olarak cezai sorumluluğunun nasıl değerlendirildiği belirsizliği ve bu belirsizliklerin nedeni olan yasal düzenlemelerdeki yetersizliklerdir. Yapay zekânın hukuki statüsü ve suç durumunda kimin sorumlu tutulacağı literatürde yer alan tartışılmaktadır [1]. Bu açıdan çalışma sonucu literatür ile tutarlıdır.

"Yapay zekâyla alakalı suç teşkil eden eylemlerde adaletin sağlanabilmesi için uygulanabilecek çözümler nelerdir?" alt sorusuna çalışmada yanıt sağlanmıştır. Yapay zekânın suç teşkil ettiği olaylarda suçlarda sorumluluğun en fazla atfedildiği kişiler yapay zekâ kullanıcıları, sahipleri ve sağlayıcılarıdır. Bu durum yapay zekâ uygulamasının sorunsuz olduğunu, geliştirilme aşamasında sorun teşkil etmediği ve yapay zekânın bilinçli olmaması varsayımının olduğu durumlar için geçerlidir. Geliştirici ve üreticilerin cezai olarak sorumlu tutulası ise kusursuz sorumluluğu ile tutarlıdır. Kusursuz sorumluluk kişinin zararı vermesi sorgulanmadan olaydan sorumlu tutulmasıdır [2]. Yapay zekâ uygulaması "anormal dereceli tehlikeli" olarak kabul edildiğinde kusursuz sorumluluk yapay zekânın faaliyet yürüttüğü kişiye uygulanabilir [10].

Çalışmalarda her yapay zekânın suç teşkil ettiği olaylarında aynı kişinin sorumlu olamayacağını, olayın özelliklerine bağlı olarak sorumluların değerlendirilmesi gerektiği önerileri bulunmaktadır. Suç olgununa özneldir, her olgu farklı faktörlere bağlı gerçekleşmiş olabilir. Bu durumda olgunun incelenmesi ve olaya yönelik değerlendirme doğru olacaktır [11]. Bu durum sorumluluğun dağıtılması olarak değerlendirilebilir. Yapay zekânın işlediği suçlarda da cezai sorumluluğun olayın incelenerek belirlenmesi gerektiği görüşleri mevcuttur [11,12]. Yapay zekânın işlediği suçlarda da olaya bağlı olarak sorumlu üretici, kullanıcı veya başka biri olabilir. Yapay zekânın cezai sorumlu tutulması görüşünde ise hukuki kişiliğe sahip olması gerekliliği sonucuna ulaşılmaktadır. Bu durumda yapay zekânın hukuki kişiliğin nasıl değerlendirilmesi gerektiği tartışmasının temelini oluşturmaktadır [1].

Yapay zekânın cezai sorumluluğuyla ilgili problemlere getirilen çözüm önerileri yasal düzenlemelerin çerçevesinde oluşturulmuştur. Yapay zekânın cezai sorumluluğunda kusursuz sorumluluk değerlendirilmesi veya sorumluluğun dağıtmasının hukuki yapıda yer alması önerilmiştir. Bunun yanı sıra hukuki statüsü olması görüşünde yeni bir hukuki kişiliğin tanımlanması, yapay zekânın elektronik kişilik olarak kabul edilmesi önerileri bulunmaktadır. Bu sonuç literatürü desteklemektedir. Yapay zekânın elektronik kişilik olarak değerlendirilmesini gerektiği yönde çalışmalar mevcuttur [13]. Yapay zekânın hukuki statüsü olamayacağı gö-

rüşü olan çalışmada ise yapay zekânın eşya olarak konumlandırılması önerilmiştir. Yapay zekâ eşya statüsünde kabulünde cezai sorumluluk eşyanın üretici veya kullanıcısına atfedilecektir [14]. Yapay zekânın kontrolünün insan tarafından sağlanması, yapay zekâ uygulamalarında insan kontrolünün artırılması ve nihai karar vericinin insan olması önerilmiştir. Yapay zekânın suça neden olacağı endişesi bu öneri şekillenmiş olabilir. Yapay zekâ uygulamalarında insan kontrolünün artmasının suç oranlarında azaltacağı veya cezai sorumluluğun yapay zekâyı kontrol eden kişi olarak değerlendirmesi çözümün önerilmesinde etkisi olabilir. Teknik çözüm önerileri ise yapay zekâ uygulamalarının standartlara uygun hale getirilmesi ve düzenli denetimlerden geçmesidir.

Alt soruların cevapları beraber değerlendirildiğine "Yapay zekânın cezai sorumluluk alın yazınında nasıl yer almaktadır?" ana sorusuna yanıt verilmektedir. Yapay zekânın cezai sorumluluğu konusunda birçok problem ve çözüm önerisi mevcuttur. Makalelerde yapay zekânın suç teşkil ettiği olayların değerlendirilmesindeki yasal problemler ortak görüşlerdir. Bu sorunun temelinde yapay zekânın bilinçli olması konusunun netliğe sahip olmaması ve hukuk yapısının bilinç çerçevesinde değerlendirme yapmasıdır. Yapay zekânın hukuki statüye sahip olup olmadığı, hukuki kişiliğinin nasıl değerlendirilmesi gerektiği ve sorumluluk rejimleri konusunda çalışmalarda farklı görüşler bulunmaktadır. Çalışmalarda yapay zekânın cezai sorumluluğu konusunda ortak problemlere odaklanmalarına rağmen getirilen çözüm önerileri farklılıklar göstermektedir. Çözüm önerilerindeki farklılar yasal düzenlemelerin yapılmasında geciktirici etkiye sahip olabilir.

Bu çalışma yapay zekâ alanındaki gelişmelerin yaşamasına rağmen hukuk açısından gelişmelerin yetersiz olmadığını ortaya koymaktadır. Bu anlamda yasal düzenlemelerin teknolojik gelişme hızına uyum sağlaması gerekmektedir. Cezai sorumlulukta İngilizce ve Türkçe makaleleri bir arada değerlendirmiş ve çalışmaları sentezlemişidir. Bu çalışma yapay zekanın cezai sorumluluğu konusundaki problemleri ve çözüm önerilerini bir araya getirerek durumu ortaya koymaktadır. Yasal düzenlemelerin olmadı gerektiği çözüm önerisi bütün çalışmalarda yer alamsına rağmen bu düzenlemeler yapay zekanın hukuki kişiliği, cezai sorumluluğu konusunda fikir birliği bulunmamaktadır. Bu nedenle literatürde problemin varlığı kabul edilmekte ancak çözümler çeşitlenmektedir. Farklı bakış açıları da yasal düzenlemelerin yapılmasında geciktirici unsur olabilir. Çalışmada yapay zekânın varlığının anlamlandırılması ve yapay zekâ hakkında bilgi sahibi olunmasıyla cezai sorumluluğun şekillendirilmesine katkı sağlayacağı öngörülmektedir.

Yapay zekânın teşkil ettiği suçlarda ne gibi cezaların verileceğinin bulunmaması bu çalışmanın sınırlılığıdır. Gelecekte yapılacak çalışmalara hukuk uzmanlarıyla gerçekleştirilen odak grup çalışmalarıyla görüşlerin nasıl şekilleneceğinin araştırılması ve uzman görüşleriyle birlikte ortak çözümlerin geliştirilmesi üzerine araştırmalar yapılması önerilmektedir.

Çıkar Beyanı**:** Yazarlar arasında çıkar çatışması bulunmamaktadır.

Yazar Katkıları: Cansel Arslan yazı taslağı hazırlama, veri analizi ve yorumlamada katkı sağlamıştır. Aslıhan Tüfekci çalışma taşasımı belirleme, veri yorumlama ve içeriğin eleştirel incelmesinde katkı sağlamıştır.

REFERANSLAR

[1] P. Bacaksız ve S.Y. Sümer, "Robotlar, Yapay Zeka ve Ceza Hukuku, Ankara: Adalet Yayınevi, 2021.

[2] S.S. Akkurt, "Yapay Zekânın Otonom Davranışlarından Kaynaklanan Hukukî Sorumluluk", Uyuşmazlık Mahkemesi Dergisi, c. 33, s. 39-59.

[3] M.H. Casın, D. Al ve N.D. Başkır, "Yapay Zekâ ve Robotların Eylemlerinden Kaynaklanan Cezai Sorumluluk Sorunu", Ankara Barosu Dergisi, 79, 1, 1-73, 2021.

[4] P. Čerka, J. Grigienė, and G Sirbikytė, "Is it possible to grant legal personality to artificial intelligence software systems?", Computer Law & Security Review, 33, 5:,685-699, 2017.

[5] I.A. Filipova and V.D. Koroteev, "Future of the Artificial Intelligence: Object of Law or Legal Personality?", Journal of Digital Technologies and Law, 1, 2, 359-386 2013.

[6] European Parliament, "Civil Law Rules on Robotics", European Parliament resolution 2015/2103(INL), Feb. 2017. [Online]. Available: https://www.europarl.europa.eu/doceo/document/TA-8-2017-0051_EN.html

[7] K. Petersen , R. Feldt, S. Mujtaba and M. Mattsson, "Systematic Mapping Studies in Software Engineering", in *Proc. 12th Int. Conf. Evaluation and Assessment in Software Engineering (EASE)*, BCS Learning & Development, June 2008.

[8] M.J. Grant and A. Booth, "A Typology Of Reviews: An Analysis Of 14 Review Types And Associated Methodologies", *Health Information & Libraries Journal*, vol. 26, no. 2, pp. 91-108, 2009.

[9] B. Başoğlu, "Sözleşme Dışı Kusursuz Sorumluluk Hukuku ve Özellikle Tehlike Sorumluluğuna İlişkin Değerlendirmeler", İnönü Üniversitesi Hukuk Fakültesi Dergisi, c. 6, 2, s.29-56, 2015

[10] S. Basu, A. Omotubora, M. Beeson and C. Fox, "Legal Framework for Small Autonomous Agricultural Robots", Ai & Society, vol.3, pp.113-134, 2020.

[11] M. Bashayreh,, F.N. Sibai and A. Tabbara, "Artificial Intelligence and Legal Liability: Towards an International Approach of Proportional Liability Based on Risk Sharing", Information & Communications Technology Law, vol. 30, 2, pp. 169-192, 2021.

[12] G. Şahin Özkan ve Ç. Şahin, "Yapay Zekalı Varlıklara Elektronik Kişilik Modeli Tanınmasına İlişkin Eurobotıcs Raporu ve Fikri Mülkiyet Sorunu Bağlamında Meseleye Yaklaşım", İnönü Üniversitesi Hukuk Fakültesi Dergisi, c. 13, 1, s. 110-138, 2022.

[13] N. Çetingül, "Ceza Sorumluluğu Bakımından Yapay Zekânın Hukuki Statüsünün Tartışılması", İstanbul Ticaret Üniversitesi Sosyal Bilimler Dergisi, c.20, 41, s.1015-1042.

YAPAY ZEKÂ, SANAT, MÜZİK ALANINDAKİ GELİŞMELERE GENEL BAKIŞ

Mehmet Şahin Akıncı*
Zonguldak Bülent Ecevit Üniversitesi, Devlet Konservatuvarı, Zonguldak, Türkiye
(İletişimden sorumlu yazar) mehmetsahinakinci@hotmail.com

ÖZET

Bilimsel çalışmaların çıktıları olan teknolojik ürünler, sanal ve fiziki ortamlarda her geçen gün etkisini artırmakta ve bu ürünlere her yaş seviyesinden bireyin ilgisi ciddi bir artışla fark edilir olmaktadır. Teknolojinin ve teknolojik ürünlerin sunduğu imkanlar ekonomi, tıp, çeşitli mühendislik alanları, mimari faaliyetler, sosyal hayat ve daha da artırılabilecek birçok alanda kullanılmaktadır. Son on yıldır adını sıklıkla duyduğumuz yapay zekâ algoritma uygulamaları da bilim ve teknolojinin sunmuş olduğu bir ürün olarak sosyal hayatta etkisini birçok alanda göstermektedir. Peki sanat alanında durum nedir?. Yukarıda sayılmış olan tüm alanlarda kullanılabilirliği artık normal bir durum olarak kabul edilebilecek yapay zekâ uygulamalarının sanat ve sanatın farklı alanlarında kullanılabilirlik durumları hakkında neler konuşulabilir?. Örneğin insanoğlunun duygularıyla işleyip müzikal düşünceleri ile tasarladığı ve bestelediği müzik eserlerinin üretim yetkinliğini yapay zekâ insanın elinden alabilir mi?. Resim sanatı alanında insanlara göre yapay zekanın oluşturduğu eserler daha etkili ve başarılı olabilir mi?. Yapay zekanın sanat üretimine ilişkin düşünceleri nelerdir?. Araştırmada bu tür sorulara cevaplar sunulması, güncel hayatta yapay zekâ, sanat ve müzik alanında mevcut olan gelişmelerin incelenerek izlenimlerin paylaşılması ve bir durum değerlendirmesi yapılması amaçlanmıştır. Yapay zekâ-sanat-müzik ilişkisinde hazırlanmış sorulara Chat-GPT'nin verdiği cevaplar ve ulusal-uluslararası statüde geçerli bilgiler sunan güncel haberler araştırmanın verileri olarak alınmış ve analizler yapılmıştır.

Anahtar Kelimeler: Yapay zekâ, sanat, müzik.

Giriş

Bilim, Teknoloji ve Yapay Zekâ Algoritması

İnsanlık tarihinin çok yönlü gelişimi hakkında sayılamayacak kadar fazla tarihi kaynaktan bilgiler edinilmiştir. Bu kaynaklarda ilk insanlardan başlayarak günümüze uzanan süreçte gelişen sosyal yaşamlar, toprak kazanımı ve ilişkili olarak yeni ekonomiler edinme amacıyla sürdürülen mücadeleler, insanların iyiliğine ve kötülüğüne yapılmış olduğu öne sürülen birçok plan ve sonuçları mevcuttur. Bir mücadele bitti derken diğeri başlamıştır. Bu bilgiler çoğu kaynak tarafından sıklıkla tekrarlanmıştır.

Milenyum yılı olarak bilinen 2000'li yıllar ile birlikte yıllardan beri var olan sosyal-ekonomik-coğrafi mücadeleler çeşitli farklılıklara uğramıştır. Bilim ve bilimin en önemli ürünlerinden birisi olan teknolojik gelişmelerle birlikte mücadelelerin yönü teknolojik üstünlük kurma yarışlarına evrilmiştir.

Ülkeler tarafından uzay atmosferine çok defa uydular gönderilmiş, internetin gelişimi ve yayılması için uyduların sağladığı veri-bağlantı altyapıları sürekli güncellenmiştir. Bir süre sonra internetin ağı-gücünün gelişmesi, teknoloji ve teknolojik ürünlerin çeşitli denetim sistemlerini gerektirmesi ve ülkelerin istedikleri bilgilere/verilere en kısa zamanda kolaylıkla ulaşabilme ihtiyacı nedenleriyle hayatımızın tüm alanlarındaki bilgiler sanal ortamlara aktarılmıştır. Fakat bu bilgilerin Google, Yandex ve diğerleri gibi çok büyük bilgi paylaşımı şirketlerinde toplanması da gün gelip yetmez hale gelmiştir. İyinin daha iyisine ulaşabilme ve bundan faydalanabilme ihtiyacı nedeniyle hem en kısa zamanda hem de tüm bilgilere aynı anda en faydalı-kullanışlı hali ile ulaşma gerekliliği doğmuştur. Bu ihtiyacın karşılığının da internet arama motorlarından veriler edinilmesi veya insanlar denetimindeki bir sistemle her istenilen anda kontrollü ve hatasız veriler edinilmesi sayesinde yeterli biçimde alınamayacağı anlaşılmıştır. Artık bir üst denetim, işletim ve kontrol elemanına-ögesine-otomasyonuna ihtiyaç doğmuştur. Bu ihtiyaç da bilim adamlarını ve teknoloji uzmanlarını otonom anlayışa, otomasyona, robotik kodlamalara, robotlara, yani yapay zekâ algoritmalarını yaratmaya ve bu ürünü tüm sosyal hayata adapte etmeye teşvik etmiştir.

Sosyal hayatta tüm dünyada neredeyse son 5 yıldır çok ciddi bir yükselişle yapay zekâ algoritmalarından bahsedilmektedir. Sosyal hayatta görsel ve işitsel kaynaklar, askeri kurumlar, ticari işletmeler ve üretim yerleri, ekonomi, bilim, eğitim, sağlık, mühendislik alanları, kültür, sanat ve belki de aklımıza gelmeyen daha bir sürü alanda uzun süredir kullanılmakta olan yapay zekâ algoritmaları ile planlamalar yapılmaktadır. Yapay zekâ ve benzeri algoritmaların insanların kısa zamanda yapamayacağı işleri, çeşitli bilgilerin derlenerek sunumunu ve bu bilgilerin kullanıma hazırlanmasını dakikalar hatta saniyeler içinde gerçekleştirebildiği görülebilmektedir.

Bütün bu yapay zekâ algoritmaları ile birlikte ülkelerin hem kamu kuruluşları hem de özel işletme-kuruluşları tarafından yapay zekâ uygulamalarının geliştirilmesi ve kullanımı teşvik edilmekte, hatta uzun süredir çoğu kurumun kurumsal iç işleyişlerinde yapay zekâ algoritmaları ile işler sürdürülmektedir. Bu gelişmelerle birlikte hali hazırda çalışmaları devam etmekte olan robot üretimi ve robotların sosyal hayatın birçok alanına adaptasyonu da desteklenmektedir.

Robot, Türk Dil Kurumu sözlüğüne göre otomatik araç demektir. Bu araçlar kendisi için özel tasarlanmış bir manyetizma (elektrik veya buna benzer bir güç-kaynak ile oluşan alan) ile belirli işleri gerçekleştirmektedir [1]."Humanoid" ismi ile tanıtılan insansı robotlar, insanların birçok görevini yerine gerçekleştirmektedir. Bir ürünün imalatından hizmet olarak sunulmasına kadar görevler almakta ve bireylere asistanlık da yapmaktadır. Kullanım alanı bilimden ve sinemaya kadar uzanmakta olan bu robotlar Android ve Gynoids sistemlerle üretilmektedir. İnsana görünüş ve fiziksel olarak da benzeyen robotlar yapay zekâ sistemi ile bilinmektedir. İlk robot dünya vatandaşı olan "Sophia", fiziksel özellikleri açısından güçlü olan ve takla dahi atabilen robotu "Atlas", televizyon kanallarında sunucu olan "Kodomo-

roid" ve görünüşü ile çok gerçekçi olan "Jia Jia" isimli robotlar farklı özellikleri ve donanımları ile tanınmış robotlardır [2].

Birbirleri ile sarmal bir şekilde ilerleyen bilim-teknoloji-yapay zekâ kavramlarının ve alanlarının yapabileceği çok iş olduğu ve üretebileceği çok ürün olacağı açık ve kesindir. "Yapay zekâ-sanat-müzik alanındaki güncel gelişmeler nelerdir? ana araştırma sorusundan yola çıkılarak araştırmada aşağıdaki alt problemler ve sorulara da cevaplar bulunması amaçlanmıştır. Yapay zekâ algoritmaları ve robotlar şimdiye kadar yüzyıllardır insanlar tarafından yapılagelmiş olan sanat eserlerini üretebilir mi?. Yapay zekâ edebiyat alanında eserler üretebilir mi?, Yapay zekâ veya robotlar farklı sanat akımlarına uyarak tablolar yapabilir mi? Yapay zekâ uygulamalarının sanat ve sanatın farklı alanlarında kullanılabilirlik durumları hakkında neler düşünülebilir?. Ülkemizde ve diğer ülkelerde yapay zekâ ve robotların ürettiği sanat eserleri ve sanatsal faaliyetleri varsa nelerdir?.

Bu araştırmada yapay zekâ-sanat-müzik ilişkisinde var olan durumun çeşitli alanlarda incelenmesi, kendisi de bir yapay zekâ algoritması olan ChatGPT'nin sanat ve yapay zekâ üzerine düşüncelerine ulaşılması, yapay zekâ-sanat-müzik alanında güncel veriler olarak sunulmuş haber-bilgilendirmelerin tespit edilerek sunulması ve üzerinde değerlendirmeler yapılması amaçlanmıştır.

Yöntem

Bu araştırmada nitel araştırma yöntemlerinden betimsel inceleme-analiz modeli kullanılmıştır. Araştırmanın amacı ve alan yönelimi doğrultusunda var olan durum çeşitli açılardan incelenerek açıklamalar sunulmuştur.

Araştırmanın verileri iki şekilde elde edinilmiştir. Birinci türde veri edinilmesine ilişkin adımlar şu şekildedir. Araştırmacı tarafından mobil olarak kullanılan Chat GPT (ios modelleri için tasarlanan model) uygulamasına sorulmak üzere araştırmanın amacı doğrultusunda veri edinmeye yönelik sorular hazırlanmıştır. Bu 7 adet soru ChatGPT tarafından cevaplanmış ve alınan cevaplar araştırmacı tarafından bir değişikliğe uğratılmadan bulgular bölümünde yer verilmiştir. ChatGPT'den veri toplama süreci 10 günlük bir sürece yayılarak edinilmiştir.

ChatGPT'ye yöneltilmiş olan sorularda herhangi bir hedeflenen cevaba doğru yönlendirme yapılmamıştır. ChatGPT'nin verdiği cevaplar aynen muhafaza edilerek bulgular bölümünde sunulmuştur. ChatGPT tarafından sorulara verilen cevaplar uygulama içeriği ve özellikleri nedeniyle uygulamanın geçmiş belleğinde tutulmaktadır.

İkinci veri elde edimi ise alana yönelik literatür taraması yolu ile edinilmiştir. İnternet arama motoru olan Google sayfasında arama boşluğuna sırası ile "yapay zekâ ve sanat", "yapay zekâ ve resim sanatı", "yapay zekâ ve müzik", "yapay zekâ ve edebiyat" başlıkları yazılarak verdiği bilgiler açısından kaynakları bilinen gelişim haberleri araştırılmıştır. Bu şekilde edinilmiş olan veriler, sundukları bilgiler bakımından araştırma içinde yer alması adına gerekli ve önemli olup olmaması açılarından incelenmiştir. Bu veriler bulgular bölümünde yer alan başlıklarda metinler halinde sunulmuştur.

Bulgular

Yapay zekâ uygulaması ChatGPT'ye sorulan sorular sonrasında edinilen cevaplar aşağıda verilmiştir.

Yapay zekâ algoritmasından edinilen veriler

Araştırmacı: Sanat nedir?

ChatGPT: Sanat, insanın duygu, düşünce ve hayal gücünü estetik bir biçimde ifade etme biçimidir. Resim, müzik, edebiyat, tiyatro, dans, heykel gibi birçok farklı dalı kapsar. Sanat, sadece güzellik yaratma amacı taşımaz; aynı zamanda toplumsal mesajlar verebilir, bireysel ve kolektif deneyimleri yansıtabilir, duyguları harekete geçirebilir ve insanları düşündürebilir. Sanatın tanımı tarih boyunca değişmiş ve farklı kültürlerde farklı anlamlar kazanmıştır. Aristoteles'e göre sanat, doğanın taklididir (mimesis); Romantik dönemde ise bireyin iç dünyasını ve duygularını ifade etme aracı olarak görülmüştür. Günümüzde sanat, kavramsal ve deneyimsel yönleriyle de değerlendiriliyor; yani bir nesne ya da eser kadar, yaratım süreci ve izleyicide uyandırdığı etkiler de sanatın bir parçası olabilir

Araştırmacı: Sanat eserlerini kimler üretebilir?

ChatGPT: Sanat eserlerini herkes üretebilir, ancak sanatçının kim olduğu ve sanatın ne olduğu konusunda tarih boyunca farklı yaklaşımlar olmuştur. Genel olarak sanat eserleri şu kişiler tarafından üretilebilir:

1-Sanatçılar: Resmi bir sanat eğitimi almış ya da kendini geliştirmiş ressamlar, heykeltraşlar, müzisyenler, yazarlar, yönetmenler, fotoğrafçılar gibi yaratıcı kişiler.

2-Halk Sanatçıları ve Zanaatkarlar: Geleneksel sanat dallarında (örneğin el dokuması, çömlekçilik, minyatür) eserler üreten kişiler.

3-Yapay Zekâ ve Algoritmalar: Günümüzde algoritmalar ile sanat eserleri giderek yaygınlaşıyor. Ancak burada sanatçının kim olduğu tartışmalı bir konu.

4-Deneysel Sanatçılar ve Performansçılar: Kavramsal sanat, enstalasyon ve performans sanatı gibi türlerde eser üreten kişiler.

5-Amatörler ve Hobi Sanatçıları: Profesyonel olmadan sanat üreten kişiler de sanata katkı sağlar.

Araştırmacı: Yapay zekanın ürettiği sanat ürünleri sanat eseri sayılır mı?

ChatGPT: Bu, sanatın nasıl tanımlandığına ve yaratıcılığın ne olduğuna dair felsefi bir tartışmaya dayanıyor. Bir yapay zekâ modelinin ürettiği bir resim, şiir, müzik veya başka bir eser sanat eseri sayılabilir mi?

Bazılarına göre sanat, insanın duygu, düşünce ve bilinç ile ürettiği bir ifade biçimidir. Bu bakış açısına göre yapay zekâ, bilinç sahibi olmadığı ve kendi deneyimlerinden yola çıkarak bir şey yaratmadığı için gerçek sanat eseri üretemez. Yapay zekânın oluşturduğu eserler, daha çok veri analizi ve algoritmaların sonucudur ve bu yüzden mekanik bir üretim olarak görülebilir.

Öte yandan, sanatın yalnızca insan yaratıcılığına dayanmadığını savunanlar da var. Bu görüşe göre sanatın en önemli kriteri, onun izleyici üzerinde estetik veya duygusal bir etki yaratmasıdır. Yapay zekâ tarafından üretilen bir şiir ya da tablo, eğer insanlarda bir duygu uyandırabiliyorsa, neden sanat sayılmasın? Sonuçta, sanatın anlamı ve değeri, onu üretenin niyeti kadar, onu deneyimleyenin algısına da bağlıdır.

Ayrıca, sanat tarihinde birçok sanatçının eserlerini üretirken rastlantısallık, teknik ve dış araçlardan faydalandığını da hatırlamak gerekir. Örneğin, Duchamp'ın hazır nesneleri veya John Cage'in rastgelelik içeren müziği de başlangıçta "gerçek sanat mı?" tartışmalarına yol açmıştı.

Sonuç olarak, yapay zekânın ürettiği eserleri sanat eseri olarak kabul edip etmemek, sanatın ne olduğuna dair kişisel ve toplumsal algıya bağlıdır. Ancak şurası kesin ki, yapay zekâ ile sanat arasındaki ilişki, gelecekte sanat dünyasında önemli bir tartışma konusu olmaya devam edecek.

Araştırmacı: Yapay zekâ çocuklar için ders veren, öğütleyen bir hikâye yazabilir mi?. Bana en fazla 200 kelimeden oluşacak ve çocuklara öğretici bilgiler sunabilecek bir hikâye yazabilir misin?. 9-10-11 yaşlarına uygun olabilir bu hikâye seviye olarak.

ChatGPT: Tabii! İşte 9-11 yaş aralığındaki çocuklar için öğretici bir hikâye:

Hikâye başlığı "Kaybolan Zaman": Can, her sabah geç kalkar, ödevlerini son ana bırakırdı. Annesi ona sürekli "Zamanını iyi kullanmalısın" dese de Can bunu önemsemezdi. Bir gün dedesi ona eski bir saat verdi. "Bu saat sihirli" dedi. "Eğer zamanını boşa harcarsan hızla ileri sarar." Can inanmadı ama saati bileğine taktı. O gün video izlerken saate baktı: 10 dakika geçmişti. Gözlerini ekrandan ayırmadan devam etti. Bir süre sonra tekrar baktığında ise saat birden akşam olmuştu!. Panikle defterlerini açtı ama yazılar bulanık görünüyordu. "Dede, saatim bozuldu!" diye şikâyet etti. Dedesi gülümseyerek, "Saat değil, senin zamanın boşa gitti," dedi. O günden sonra Can, zamanını daha iyi kullanmaya başladı. Ödevlerini zamanında yaptı, oyun oynayacağı saatleri planladı. Artık zaman kaybolmuyor, verimli geçiyordu. Hikâyeden çıkarılacak ders: Zamanı doğru kullanmak, gelecekte pişman olmamak için önemlidir.

Araştırmacı: Doğanın güzellikleri üzerine en fazla üç kıtadan oluşacak kafiyeli bir şiir yazabilir misin?

ChatGPT: Tabii! işte doğanın güzelliklerini anlatan, kafiyeli ve üç kıtalık bir şiir:

Doğanın Ezgisi
Rüzgâr eser yavaş yavaş,
Dağlar yükselir maviye.
Nehrin sesi tatlı bir nağme,
Yeşil uzanır sonsuz diye (Orman eşlik eder ahenge)

Güneş gülümser çiçeğe,
Sabah ışıkla dans eder.
Gökyüzü bir düş ülkesinde,
Kuşlar şarkılar besteler.

Deniz köpürür sevdasıyla,
Dalgalar vurur sahile.
Toprak uyanır baharla,
Hayat doğan bin renkle.

Bu şiir ChatGPT tarafından yazıldıktan sonra ChatGPT'ye ilk kıtanın son mısrasını değiştirebilir misin şeklinde yazıldı. ChatGPT hemen değişikliği parantez ile belirtilmiş şekli ile yaptı.

Araştırmacı: Yapay zekâ yetişkin yaştaki bireyler için bir aşk romanı yazıyor olsa, bu romanın önemli bir bölümü olan şu kısmı en fazla iki paragraf ve 15 cümlede, okuyanları duyguya boğacak ve ağlatacak şekilde nasıl yazardı ? Birbirlerine deli gibi âşık olan iki sevgiliden kadın olan sevdiği adamın kollarında ölme anındadır. Bu sahneyi okuyucuyu yoğun duygusallığa iterek ve çaresizlik hissettirecek bir şekilde nasıl yazardın ?

ChatGPT: Adam, titreyen elleriyle sevdiği kadının yüzünü avuçlarının arasına aldı, gözlerinden akan yaşlar onun solgun yanaklarına düşerken hıçkırıklarını bastırmaya çalıştı. Kadının nefesi kesik kesik çıkıyor, her an kayıp gidecekmiş gibi hafifliyordu. "Gitme," diye fısıldadı adam, sesi çaresizlikle kırılmıştı. Kadın yorgun göz kapaklarını aralayıp gülümsedi; zayıf, kırılgan, ama hâlâ içinde aşk taşıyan bir gülümsemeydi bu. "Senin gözlerinde ölmek... en güzel rüya gibi..." dedi, sesi rüzgârın taşıdığı bir melodi kadar hafifti. Adamın göğsü daraldı, parmaklarını onun parmaklarına kenetledi ama sıcaklık giderek kayboluyordu.

Kadının gözleri gökyüzüne dönerken, son bir kez onun adını fısıldadı ve nefesi sessizce havaya karıştı. Adam, sanki zaman durmuş gibi nefes almayı unuttu, kalbi boşluğa düştü. "Hayır... hayır, ne olur, beni sensiz bırakma!" diye haykırdı, onu sıkıca göğsüne bastırırken. Ama kadının bedeni artık yanıt vermiyordu, sonsuz bir huzurun içine çekilmişti. Adam, onun saçlarına yüzünü gömüp ağlarken, dünya sessizleşti. Bir an, sanki tüm evren yalnızca onun kaybına yas tutuyordu—rüzgâr esmeyi, kuşlar ötmeyi, zaman akmayı bırakmıştı. Sevgisinin sıcaklığı hâlâ avuçlarının içinde ama sevdiği kadın artık orada değildi.

Araştırmacı: "İnsanın çaresizliği" ana temasından yola çıkarak bana empresyonizm, realizm, romantizm ve sürrealizm sanat akımlarının anlayışı ile yapılmış 4 tablo yapabilir misin?. Tablolarda insan olacaksa cinsiyet belli olmasın lütfen.

ChatGPT:

Şekil 1: İnsanın Çaresizliği temalı ChatGPT'nin farklı sanat akımlarında oluşturduğu tablolar

Yapay Zekâ Algoritmaları ile İlgili Güncel Kaynaklardan Edinilen Veriler

Bilim, Teknoloji, Yapay Zekâ ve Sanat

Benzer gerçekleri yansıtmakta olan bilim ve sanat kavramlarına ilişkin sanatın, bilimin oluşturduğu somut ürünlerden ve çıktılardan yararlandığı söylenebilir. Farkında olmadan kendisi için gerekli olanları bilimden alan sanatçı, bilimsel çalışmaları takip eder. Bilimsel alanda meydana gelen yenilikleri adım adım izler ve kendi bünyesinde ve çalışmalarında bunları uygular. Fakat bu faaliyetler sırasında asıl amacı bilimsel çalışmalar yapmak değil sanatını karşılıksız bir beklenti ile üretmektir. Bilimin duygularla işlenmiş biçimi sanattır denilebilir [3].

Yapay zekanın sanatsal hayata olan olumlu-olumsuz olası etkileri hakkında ne ülkemizde ne de dünya ülkelerinde herhangi bir resmi-net basın açıklaması, kamusal bilgilendirme raporu, varsa olumsuz yanlarına ilişkin uyarı açıklamaları henüz yapılmamıştır. Bu alanda açıklama, rapor yazma, bildiriler sunma, önlemler alma, yasalarla sınırlıklar çizme gibi yazılı-sözlü faaliyet ve girişimlerin, ülkelerdeki sanat alanında etkisi olan dernekler, vakıflar, bakanlıklar, ulusal-uluslararası teknoloji ofisleri, sanat alanında yükseköğretim eğitimi kurumlarında çalışan yetkili akademik personeller, sanatçılar ve devlet yönetimleri tarafından gerçekleştirilebilecek görevler olduğu söylenebilir.

Sosyal hayatın her alanında etkisini hissettirmekte olan yapay zekâ alanında tüm ülkeler yatırımlar yapmaktadır. Yapay zekâ kullanımı ve geliştirimi hakkında planlamaların, insanların faydasına olduğu veya olmadığı sıklıkla devam eden tartışmalar ve girişimlerle sürdürüldüğü söylenebilir. Hatta ülkemiz dışında sanatın farklı alanlarında da yapay zekâ ve algoritmaların etkilerinin olduğu görülebilmektedir. Araştırmanın bu bölümünden itibaren henüz ülkemizde etkisi yeterince hissedilmeyen sanat ve yapay zekâ alanındaki gelişmelerden bilgiler sunulmaya çalışılmıştır.

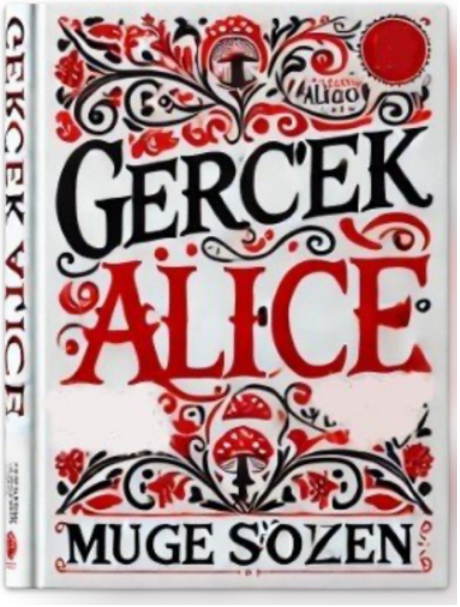
Yapay Zekanın ve Robotların Sanatsal Yaşama Etkileri

Sanat: öz'dür, birdir yani biricik'tir, estetiktir ve yaratıcısının duygu ve düşünceleriyle sanatsal ortamlara kendisini sunar. Yüzyıllar boyunca en ilkel düşünsel oluşumlarla başlayıp gelişen sanat: güzellik, bireysellik, görelilik gibi kavramlarla ilişkiler kurarak ilkel özelliklerinden modern biçime evrilmiş ve günümüz dönemine kadar ulaşmıştır. Güncel hayatımızda sanatsal ürünlerin amaç ve sonuçlarının değişiklikler göstermesi nedeniyle yaratılan sanatsal ürünlerin de kimi çevreler tarafından sanat eseri, kimileri için ise sanat eseri olduğunun kabul edilmemesi gibi karmaşık süreçler içinde kalınabilmektedir. Sanat eserlerinin öznel (subjektif) veya nesnel (objektif) yargılarla sanat eseri sayılıp sayılamayacağı ile ilgili bu tür karmaşık karar verme süreçlerine son yıllarda "yapay zekanın sanatı ve yarattığı sanat eserleri" tartışmalarının da eklendiği rahatlıkla görülebilmektedir.

2024 yılının Kasım ayında duyurulan bir habere göre Ai-Da isimli robotun (Şekil-2) çizdiği "Alan Turing" portresi 1,8 milyon dolara satıldı. İngiliz matematikçi ve bilgisayar bilimcisi olan Alan Turing isimli bilim adamının portresi bir robot tarafından yapılmış olan ilk sanat eseri sayılmaktadır. Açık artırmada satışa sunulmuş olan bu portrenin en fazla 180 bin dolara satılması beklenirken bu tablo çok daha pahalı satışı elde etmiştir. Alan Turing'in 1950'lili yıllarda yapay zekâ üzerine temel düşünceleri şekillendirmiş bir bilimci olduğu bilinmekteydi [4].

Şekil 2: Robotik kolları ve yapay zekâ işletim sistemi ile çalışan Ai-Da isimli robot ve arkasında çizdiği tablolar

Edebiyat alanında da yapay zekâ destekli çeşitli sanat eserleri üretiminin olduğu görülebilmektedir. Müge Sözen isimli yazarın yapay zekâ destekli sunduğu bu e-kitabın ismi (hikâye) **"**Alice Harikalar Diyarında: Gerçek Alice" olmuştur. Orijinal içeriği ile gerçek sahibi Lewis Carroll tarafından yaratılmış olan bu hikâye, Viktorya döneminde geçmektedir. Yapay zekâ desteğiyle hazırlanan bu kitap, Müge Sözen'in düşlem gücüyle Oxford'un canlı sokaklarından Christ Church Koleji'nin tarihi koridorlarına ve büyüleyici katedrallerin bahçelerine kadar uzanmaktadır. Yapay zekanın desteği ve yaratıcılık girişimleri kitabın birçok bölümünde anlaşılabilmektedir.

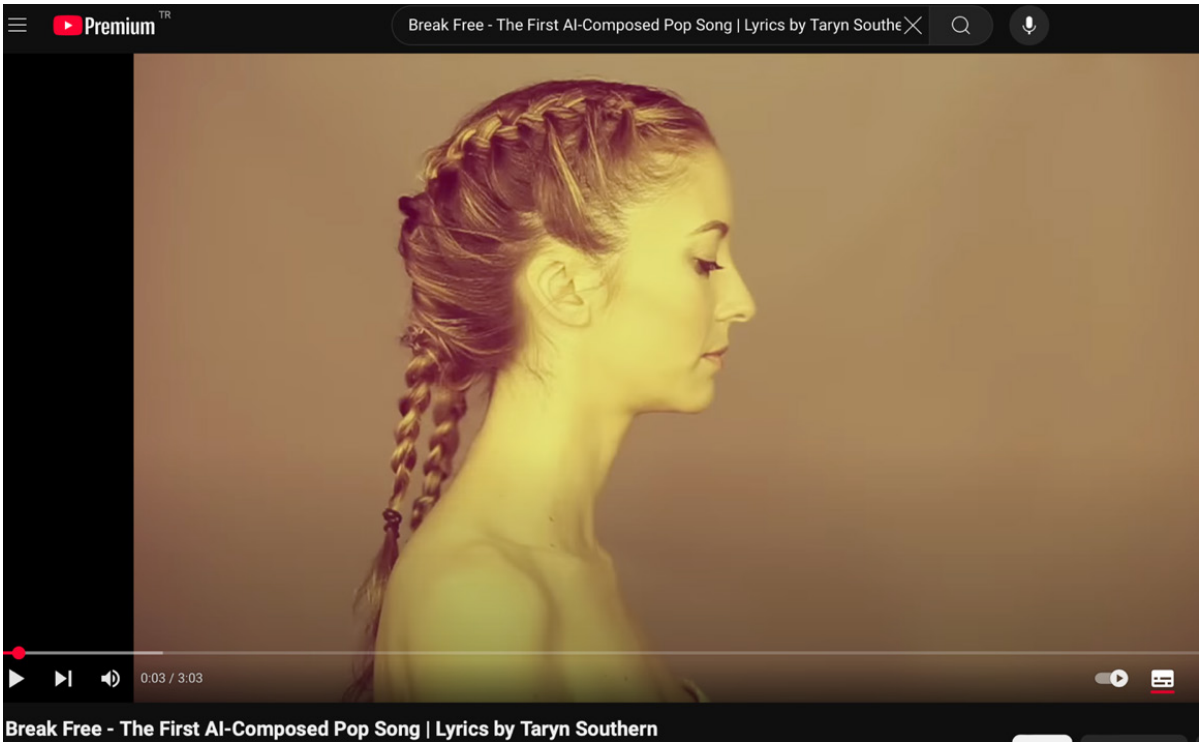


Şekil 3: Müge Sözen'in yapay zekâ desteği ile yazdığı **"**Alice Harikalar Diyarında: Gerçek Alice" isimli kitap kapağı (play.google.com). [5]

Yapay Zekanın Müziğe ve Müzikal Yaşama Etkisi

Yüzyıllar boyunca müzikal birikimler geleneksel anlayışlarla fiziki olarak saklanmış, akustik yani doğal performanslarla seslendirilmiş, bu sayede natural-doğal tınıların sürekliliği sağlanmaya çalışılmıştır. Fakat bilim ve teknolojnin hızlı şekilde ilerlemesiyle sanal-online sunumlara önem verilmiş, müzikal ses-görüntü kayıtları ve diğer içerikler teknoloji içerik özelliklerine kavuşturulmuştur. Cep telefonu ve bilgisayar uygulamaları ile güncel hayatta gelişimlerini hızlandırmış olan müzik teknolojisi, yapay zekâ algoritmaları ile iş birliği yaparak çok daha faydalı, zengin, hızlı ve erişilmesi kolay niteliklerini iyice artırmıştır.

2016'da Sony firmasının "Flow Machine" isimli yazılımı, Beatles grubunun müzik tarzında ilk pop müzik şarkısını üretmiştir. Bu bestenin sözleri bir insan söz yazarına aittir. Bestelenmesinde ise yapay zekanın ürettiği müzikal bölümlere de yer verilmiştir. 2017 yılında ise Amper isimli şirket ile YouTube sanatçılarından Taryn Southern'ın anlaşma yapması önemli bir adım olmuştur. Amper isimli şirket çok kısa süre içinde profesyonel müzik üretebilen bir yapay zekâ geliştirmiştir. Bu anlaşmadan şirket, insan-yapay zekâ iş birliği ile ortaya çıkan ilk tekli (single) albümü oluşmuştur. Bu eser sıfırdan üretilmiş bir beste olmuştur. Bestenin yapılması için yapay zekaya Amper tarafından şarkının ruh hali, müzik tarzı ve saniyede kaç vuruş olacağının (bpm) değerleri ve tanımlamaları verilmiş ve besteyi yapması beklenmiştir (trtmuzik.net.tr). [6].



Şekil 4: 2017 yılında Amper şirketi ve Taryn Southern işbirliğinde bestelenen yapay zekâ-insan besteci ortaklığında ilk şarkı (https://www.youtube.com/watch?v=XUs6CznN8pw).

Üç ayrı robotik kolun oluşturduğu teknolojik düzenek olan orkestra şefi "MAiRA Pro S'" (Şekil 5), 2024 yılı ekim ayı başında Almanya'nın Dresden şehrinde sahneye çıkan orkestrayı yönetmiştir. MAiRA Pro S' isimli robot, Dresden Senfoni Orkestrası'nın 25. yılı kutlamalarında önemli bir müzik olayı olması düşüncesiyle orkestranın sanat yönetmeni tarafından sahneye çıkarılmıştır. Orkestranın sanat yönetmeni Markus Rindt, bu farklı düşüncenin robot şeflerin insan orkestra şeflerinin önüne geçmesini sağlamak olarak algılanmasını istememiştir. Bu konserin gerçekleştirilmesinden önce Rindt, üç koldan oluşacak şekilde dizayn edilmiş olan robota yönetim hareketlerini öğretmiştir [7].

Şekil 5. MAiRA Pro S' isimli robot ve yönettiği orkestra

Güney Kore'nin Seul kentinde gerçekleştirilen bir konserde ise, sesleri dinleme yeteneği olmayan Android işletim sistemli insansı görünümlü robot EveR 6 (Şekil 6 ve 7), toplam beş eserden oluşturulmuş olan konserin üç eserini yönetmiştir. Orkestranın şefi Choi, bu faaliyetteki amacın, robotların insanların yerini alabileceğini göstermek değil bir arada çalışabileceklerini gösterebilmektir şeklinde açıklamada bulunmuştur [8].

Şekil 6: "EveR 6" isimli robot ve Güney Kore'de yönettiği orkestra

Şekil 7: "EveR 6" isimli robot Güney Kore'de yönettiği orkestra ile konser öncesi provada

Sonuçlar ve Öneriler

Yapay zekanın toplumun gelecek yaşamında birçok şeyi değiştirecek güçte olması çok açıktır. Burada önemli olan telaşlanmak ve sonun geldiğini düşünmek değildir. Olması gereken bu teknolojik geçişe hazırlıklı olmaktır. En kısa zamanda robotbilim ile başlayan araştırma ve geliştirme faaliyetlerine başlanmalıdır. Bunun için gerekli ise yasal maddeler de gözden geçirilmeli ve insanları her açıdan olumsuz etkileyebilecek durumlara karşı etik özelliklere de dikkat edilmelidir [9].

Güncel hayatta görülebilmektedir ki yapay zekâ ChatGPT'yi ve algoritmaları herhangi bir şey için kullanmayan yok denecek kadar azdır. Bu çalışmanın içeriğinde de belirtildiği üzere ChatGPT, algoritmalar ve robotlar, sanat eseri yaratımında da oldukça etkileyici ürünler oluşturabilmektedir. Yapay zekâ ile çalışan makine görünümlü robotlarla birlikte son zamanlarda insan ve hayvan görünümlü yapay zekâ algoritmaları ile çalışan robotların da varlığına şahit olunmaktadır. Bu makinalar ve robotların insanların üretebildiği düşünsel ürünleri ve üretebildiği fiziki ürünleri de tasarlayarak sunabildiği görülebilmektedir.

Bu araştırmanın oluşturulmasında en önemli nedenlerden birisi de yapay zekâ yani yapay bir oluşum olan ChatGPT'nin, bir sanat üreticisi olan insana göre çok kısa sürede, sanatsal açıdan değerli olup olmayacağı belki de uzun zaman tartışılacak olan yaratıcılık-üreticilik gücüne sahip olmasının incelenmesi ve bu sıra dışı duruma dikkat çekmektir. Sosyal, akademik ve diğer üretim-inşaa alanlarının çoğunda sıklıkla kullanılmakta olan yapay zekâ ve algoritmaların, sanat alanında da ciddi bir güç-etki alanı yaratıp, insan düşüncesi ve eli ile üretilmiş olan sanat eserlerinin öneminin eksilmesine neden olabilme tereddütlerinin ortaya bir ihtimal olarak koyulmasıdır.

Elbette her ülkede sanat hem devlet-kamu hem de özel kurumlardaki sanat savunucuları ve sanatçılar tarafından takip edilmekte, desteklenmekte ve sanatsal gündemdeki çeşitli konular tartışılmaktadır. Sanatın halkın içine doğru ilerleyen, işleyen ve çevresini etkileyebilen bir kavram olduğu farz edilirse ve halkın da yapay zekâ sanatına bireysel ilgi ve meraktan dolayı önem vererek yapay zekâ sanatına rağbet etme eğiliminde olmasının normal bir durum olabileceği kabul edilebilir. Bu sakin bekleyişin insan eli ile üretilen sanat eserlerinin değeri ve sanatçılarının sanatsal gelişimi açısından olumlu olumsuz çıktılar yaratabileceği de düşünülmektedir.

Geçmişte olduğu gibi güncel hayatta da sanat eserleri hem insan eli hem de yapay zekâ ve çeşitli algoritmalarca üretilmeye devam edecektir. Yeni bir üretim alanı olan yapay zekâ sanatının ileriki süreçlerde durumunun ne olacağı, kabul edilip edilmeyeceği hakkında net yargılarda bulunmak güç bir durumdur. Yapay zekâ üretimi olan sanat eserlerine olan rağbetin artması ve insan eli ile üretilmekte olan sanat eserlerinin önem, değer ve ilgisinin kaybolma ihtimaline karşın devletler tarafından yönetilmekte olan sanat kurumları, sanat otoriteleri, sanat dernekleri, sivil toplum kuruluşları, sanatın tüm alanlarında sanatçılar ve sanat eğitimi alanında önder akademisyenlerin bu problematiğe öncülük ederek olası olumsuz deneyimlerin yaşanmasına imkân verilmemesi için önemlerin alınması gerekmektedir. Bu önlemler çerçevesinde çeşitli sanatsal ve akademik buluşmalar, çalıştaylar ve paneller düzenlenerek konu ayrıntılı biçimde irdelenebilir.

Müzik alanında eğitim vermekte olan yükseköğretim kurumlarındaki müzik teknolojileri bölümlerinde, teknik donanım olarak yeterlik gösteren müzik bölümlerinde ve ilgili diğer teknolojik bölümlerde "yapay zekâ, algoritmalar-robotlar ve müzik" içerikli derslerin öğretim programlarına dahil edilmesinde faydalar sağlanabilir. Daha da profesyonel kapsamda düşünülerek ilgili yükseköğretim okullarında bu alana özel anabilim-anasanat dallarının açılması sayesinde yapay zekâ-sanat alanlarında olası olumsuz gelişme ve benzeri durumların akademik ve bilinçli denetimlerle engellenebileceği de düşünülmektedir.

Çıkar Beyanı: Araştırmada herhangi bir neden, kişi veya kuruma ilişkin çıkar çatışması yoktur.

REFERANSLAR

[1] Türk Dil Kurumu, https://tdk.gov.tr/

[2] Sayar, B. (2021). Yapay Zekâ ve Robot Teknolojisinin Gazetecilik Pratikleri Üzerindeki Etkileri [Yayınlanmamış Doktora Tezi]. Marmara Üniversitesi.

[3] Şen, Ü. S., (2005). Sanat Eğitiminde Bilimsel Araştırma Yöntemlerinin Kullanılması. *Atatürk Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 5 (1), 343-360.

[4] Ege Telgraf haber kaynak: (https://www.egetelgraf.com/ai-da-robotunun-cizdigi-resim-tam-18-milyona-satildi-iste-o-tablo)

[5] Alice Harikalar Diyarında: Gerçek Alice-Google Play: https://play.google.com/store/books/details?id=a2pNEQAAQBAJ&rdid=book-a2pNEQAAQBAJ&rdot=1&source=gbs_vpt_read&pcampaignid=books_booksearch_viewport

[6] trtmuzik.net.tr:https://trtmuzik.net.tr/haber/iste-yapay-zeka-tarafindan-yazip-bestelenen-ilk-sarki_189

[7] TRT Haber kaynak:https://www.trthaber.com/haber/dunya/uc-kollu-robot-orkestra-sefi-dresden-de-ilk-kez-sahneye-cikti-882243.html

[8] ntv.com.tr haber kaynak: https://www.ntv.com.tr/galeri/n-life/kultur-ve-sanat/ilk-kez-bir-robot-or-kestra-sefligi-yapti,MpUA9bz9v0GC732gksLndQ/UmeNufU7UkmnF0BEzMuB3A

[9] Oygür, A.V. (2020). Yapay Zekanın Meslekleri Tehdidi Karşısında Jeolojinin Durumu. Madencilik ve Yer Bilimleri Dergisi. Haziran Sayısı.

5G ÇEKİRDEK AĞINDA YAPAY SİNİR AĞI MODELİ İLE SALDIRI TESPİTİ

Harun Şeker*
Bilecik Şeyh Edebali Üniversitesi, Lisansüstü Eğitim Enstitüsü, Bilgisayar Mühendisliği,, Bilecik, Türkiye

Hüseyin Parmaksız,
Bilecik Şeyh Edebali Üniversitesi, İktisadi ve İdari Bilimler Fakültesi, Yönetim Bilişim Sistemleri, Bilecik, Türkiye
**(İletişimden sorumlu yazar) harun@harunseker.org*

ÖZET

Beşinci nesil (5G) mobil iletişim sistemleri, yüksek hız, düşük gecikme süresi ve daha fazla cihaz bağlantısı gibi avantajlarla birçok sektörde dönüşüm yaratmaktadır [1]. Ancak bu gelişmiş özellikler, beraberinde yeni ve daha karmaşık güvenlik tehditlerini de getirmektedir [2]. Bu çalışmada, özellikle 5G çekirdek ağına yönelik Komuta ve Kontrol (Command and Control - C2) saldırılarının tespiti amacıyla bir yapay sinir ağı (Artificial Neural Network - ANN) tabanlı derin öğrenme modeli geliştirilmiştir. Geliştirilen ANN modeli, 10 farklı girdi özelliğiyle (örneğin, protokol, kaynak/hedef IP ve port numaraları, TTL, bayraklar vb.) eğitilmiş olup dört gizli katmandan oluşmaktadır. Model, %88.54 doğruluk, %99.25 kesinlik, %78.75 duyarlılık ve %87.82 F1-skoru elde ederek saldırı tespitinde yüksek bir performans sergilemiştir. Özellikle düşük false positive oranı sayesinde modelin gereksiz uyarılar üretmediği gözlemlenmiştir. Bu çalışma, geliştirilen derin öğrenme modelinin 5G çekirdek ağ trafiğinde gerçek zamanlı saldırı tespiti için umut verici bir potansiyele sahip olduğunu ortaya koymaktadır. Bununla birlikte, modelin duyarlılık oranının artırılması gerektiği ve daha geniş veri kümeleri ile yeniden eğitilmesinin model performansını iyileştirebileceği önerilmektedir. Dijitalleşme sürecinin hızla ilerlediği günümüzde, bu tür güvenlik önlemlerinin ve gelişmiş tespit sistemlerinin, gelecekteki ağ altyapılarının daha güvenli ve verimli olmasına önemli katkılarda bulunacağı düşünülmektedir.

Anahtar Kelimeler: ANN, 5G çekirdek ağı, siber güvenlik, ağ trafiği analizi, saldırı tespiti.

Giriş

Beşinci nesil (5G) mobil iletişim sistemleri; yüksek veri aktarım hızları, düşük gecikme süresi ve geniş bağlantı kapasitesi sayesinde, bilgi ve iletişim teknolojilerinde önemli bir dönüm noktası olmuştur [1]. Ancak, artan bağlantı hızları ve mobil ağlar aracılığıyla internete bağlanan cihaz sayısındaki artış, beraberinde çeşitli güvenlik risklerini de getirmiştir [2]. Bu risklerden biri de komuta ve kontrol (Command and Control – C2) saldırılarıdır [3]. Bu tür

saldırılarda, zararlı yazılımlar tarafından enfekte edilen bilgisayar sistemleri, saldırganlarla bağlantı kurarak onların verdiği komutları yerine getirir. Saldırganlar bu yöntemle hedef ağda kalıcılık sağlayarak; veri hırsızlığı, hizmet dışı bırakma (DoS) saldırıları ve benzeri kötü niyetli faaliyetler için ele geçirilmiş sistemleri kullanabilmektedir [4]. 5G teknolojileri ile artan hız ve giderek karmaşıklaşan saldırı vektörleri karşısında geleneksel saldırı tespit ve engelleme yöntemlerinin yetersiz kalması olasıdır. Bu nedenle 5G Sistemleri hedef alan bu tür saldırıları tespit etmek için yapay zeka teknolojilerini kullanan çeşitli çalışmalar literatürde yer almaktadır.

Prasad ve arkadaşları (2020), ağ saldırılarının tespiti amacıyla Bayes teoremi ve yaklaşık küme teorisine (Rough Set Theory - RST) dayalı yeni bir özellik seçimi ve sınıflandırma yöntemi önermiştir. CICIDS2017 veri seti üzerinde gerçekleştirilen deneylerde, önerilen yaklaşım yüksek doğruluk ve düşük yanlış alarm oranı sağlamıştır. Özellik seçiminde, önemli nitelikler olasılık tahminine dayanarak belirlenmiş; kesin ve belirsiz bilgiler RST ile ayrıştırılmış ve ardından Bayes teoremi kullanılarak sınıflandırma gerçekleştirilmiştir. Bu yöntem, saldırı tespit sistemlerinde eğitim süresi ile hesaplama karmaşıklığını önemli ölçüde azaltmıştır [5].

Ferrag ve arkadaşları (2020), siber güvenlikte saldırı tespiti için kullanılan derin öğrenme yöntemlerini, kullanılan veri setlerini ve bu yaklaşımlar arasındaki karşılaştırmaları kapsamlı bir şekilde incelemiştir. Çalışmada 35 farklı siber güvenlik veri seti kategorize edilerek tanıtılmış ve yedi temel derin öğrenme modeli (DNN, RNN, CNN, DBN, RBM, DBM ve Deep Autoencoder) analiz edilmiştir. Ayrıca, CSE-CIC-IDS2018 ve Bot-IoT veri setleri üzerinde yapılan deneylerde, derin öğrenme modellerinin doğruluk, hata oranı ve tespit oranı gibi performans kriterleri kullanılarak etkinlikleri değerlendirilmiştir [6].

Khan ve arkadaşları (2020), 5G ağlarının güvenlik ve gizlilik gereksinimlerini kapsamlı bir şekilde ele alarak, ağ yazılımlaştırma (SDN, NFV, MEC, NS) gibi temel teknolojilere özgü tehditleri ve çözüm önerilerini değerlendirmiştir. Çalışma, 5G'nin fiziksel katman güvenliği, ağ izleme, yönetim, kullanıcı gizliliği, standartlaştırma çalışmaları ve güvenlik projelerine ilişkin ayrıntılı analizler sunmaktadır. Ayrıca, mevcut güvenlik tehditlerine karşı alınabilecek önlemleri, standartlaştırma kuruluşlarının rolünü ve 5G ağları için gelecekteki araştırma yönelimlerini tartışarak literatüre bütüncül bir bakış açısı kazandırmıştır [7].

Gupta ve arkadaşları (2018), çok aşamalı 5G kablosuz iletişim ağlarında (WCN) güvenlik tehditlerini analiz ederek özellikle bant genişliği sahteciliği (bandwidth spoofing) saldırısına odaklanmıştır. Çalışmada oyun teorisi kullanılarak saldırıların matematiksel modeli geliştirilmiş ve ardından röle, küçük hücre erişim noktası (SCA) ve baz istasyonu (BS) gibi ağ bileşenlerinde saldırıları tespit etmek için gizli Markov modeli (HMM) tabanlı uyarlamalı bir saldırı tespit sistemi (IDS) önerilmiştir. Simülasyon sonuçları, önerilen yöntemin bant genişliği sahteciliği saldırılarını etkili şekilde tespit ederek 5G ağlarının güvenlik seviyesini artırabileceğini göstermektedir [8].

Vinayakumar ve arkadaşları (2019), hem ağ tabanlı hem de host tabanlı saldırıların tespiti için derin öğrenmeye dayalı yeni nesil bir saldırı tespit sistemi geliştirmiştir. Çalışmada, farklı ağ ve sistem etkinlikleri üzerinde Derin Sinir Ağları (DNN) kullanılarak kapsamlı bir deneysel analiz gerçekleştirilmiş ve klasik makine öğrenmesi yöntemleriyle karşılaştırmalı sonuçlar sunulmuştur. Yazarlar ayrıca büyük veri ölçeğinde çalışan, gerçek zamanlı saldırı tes-

piti yapabilen ve ölçeklenebilir bir hibrit framework olan Scale-Hybrid-IDS-AlertNet (SHIA) sistemini önermiştir. KDDCup99, NSL-KDD, UNSW-NB15, Kyoto, WSN-DS ve CICIDS2017 gibi farklı veri kümeleri kullanılarak yöntemlerin etkinliği test edilmiştir [9].

Fonyi (2020), 5G ağlarının güvenlik mimarisi ve potansiyel zafiyetleri üzerine kapsamlı bir inceleme sunmuştur. Çalışmada 5G'nin getirdiği yeni teknolojiler (örneğin, 5G NR, SDN, D2D iletişim, Massive MIMO) ve bu teknolojilerin güvenlik açıklarına nasıl kapı aralayabileceği tartışılmıştır. Özellikle, güvenlik zafiyetleri gizlilik, bütünlük ve erişilebilirlik ekseninde sınıflandırılmış; 5G-AKA protokolündeki açıklardan, MITM saldırılarına, DDoS tehditlerine ve IoT cihazlarının 5G ağına entegrasyonu ile artabilecek risklere kadar geniş bir değerlendirme yapılmıştır. Yazar, standartların henüz tamamlanmamış olması nedeniyle, 5G ağlarının güvenlik açısından dikkatle izlenmesi gerektiğini vurgulamaktadır [10].

Zhang, Li ve Wang (2019), nesnelerin interneti (IoT) ortamlarında saldırı tespiti için geliştirilmiş bir genetik algoritma (GA) ve derin inanç ağı (DBN) tabanlı bir model önermiştir. Çalışmada, farklı saldırı türlerine uyum sağlayabilen bir ağ yapısı elde etmek amacıyla GA kullanılarak DBN›nin katman sayısı ve her katmandaki nöron sayısı optimize edilmiştir. NSL-KDD veri seti üzerinde gerçekleştirilen deneyler, önerilen GA-DBN modelinin hem saldırı tespit oranını artırdığını hem de ağ yapısının karmaşıklığını azalttığını göstermektedir. Bu yöntem, düşük frekanslı saldırıların tespitinde dahi yüksek doğruluk ve düşük eğitim maliyeti sağlamıştır [11].

Liu ve Lang (2019), makine öğrenimi ve derin öğrenme tabanlı saldırı tespit sistemlerine (IDS) yönelik kapsamlı bir literatür taraması sunmuştur [12]. Çalışmada, veri nesneleri merkezli bir taksonomi önerilerek IDS araştırmaları sınıflandırılmış; sık kullanılan makine öğrenimi algoritmaları, değerlendirme metrikleri ve benchmark veri kümeleri tanıtılmıştır. Ayrıca, makine öğrenimi ve derin öğrenme teknikleriyle mevcut IDS problemlerinin nasıl çözülebileceği açıklanmış ve alanın güncel zorlukları ile gelecekteki araştırma yönelimleri tartışılmıştır. Bu derleme, siber güvenlik araştırmacılarına hem teorik bir çerçeve hem de uygulamalı yönlendirmeler sunmayı amaçlamaktadır.

Gai, Qiu, Tao ve Zhu (2015), heterojen 5G ortamlarında mobil bulut bilişim güvenliğini sağlamak amacıyla saldırı tespit sistemlerine (IDS) dayalı yeni bir güvenlik çerçevesi önermiştir [13]. Çalışmada, MCC (Mobile Cloud Computing) mimarisinin temel teknolojileri tanıtılmış, çeşitli IDS teknikleri (imza tabanlı, anomali tabanlı, özellik tabanlı, protokol analizi tabanlı ve hibrit yaklaşımlar) detaylı şekilde analiz edilmiştir. Yazarlar, geniş bant ve heterojen bağlantı imkanları sayesinde, bulut tabanlı IDS'lerin mobil cihazlar için etkili bir güvenlik mekanizması sunabileceğini vurgulamış ve enerji verimliliği, gizlilik koruması gibi gelecekteki araştırma yönelimlerini tartışmışlardır.

Samarakoon ve arkadaşları (2022), gerçek bir 5G test ağı üzerinde oluşturulan veri akışlarını ve saldırıları etiketleyerek 5G-NIDD isimli saldırı tespit veri kümesi geliştirmiştir. 5G ağlarına özgü port taraması ve DoS saldırıları gibi saldırı trafikleri içeren bu veri seti aynı zamanda 5G cihazlar tarafından üretilen normal trafik verilerini de barındırmaktadır. Çalışma mevcut saldırı tespit sistemlerine (Intrusion Detection Systems – IDS) ait veri kümelerinin 5G ağları için yetersiz kaldığını vurgulamaktadır. Bu veri seti üzerinde çeşitli makine öğrenmesi modelleriyle gerçekleştirdikleri sınıflandırma deneyleri ile veri setinin doğruluğunu ve

saldırı tespit sistemleri konusunda yapılacak araştırmalarda kullanılabilirliğini ortaya koyarak yapay zeka temelli saldırı tespit sistemlerinin geliştirilmesine katkı sağlamaktadır [14].

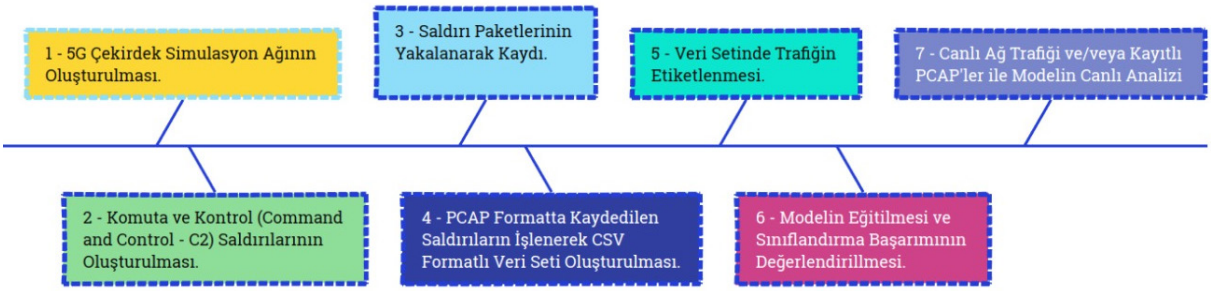
Gupta ve arkadaşları (2017), 5G ağlarındaki tehditleri analiz ederek, özellikle Küçük Hücre Erişim Noktaları (Small Cell Access – SCA) üzerindeki bant genişliği saldırılarını (Bandwidth Spoofing) ele almışlardır. Bu saldırıların etkisi oyun kuramı (Game Theory) çerçevesinde ele alınmış ve tutarsız istemcilerin ağ kaynaklarını nasıl tüketebilecekleri matematiksel olarak modellenmiştir. Saldırgan davranışlarını tespit edebilmek için Gizli Markov Modeli (Hidden Markov Model) tabanlı bir saldırı tespit sistemi önerilmiştir. Sunulan bu Gizli Markov Modeli, sahte istemcilerin davranışlarını işaret eden veriler ile eğilmiş ve anomalileri tespit ederek saldırgan sistemleri yakalamıştır [15].

Hamroun ve arkadaşları (2025), 5G ve Wi-Fi ağlarının bütünleşik kullanımının ortaya çıkardığı güvenlik risklerine odaklanarak, bu ortamlara özgü saldırı tespit sistemlerine yönelik kapsamlı bir literatür taraması sunmuştur. Çalışma, tehdit yüzeylerini, IDS mimarilerini ve makine öğrenimi temelli tespit stratejilerini sistematik biçimde incelemekte; federated learning ve anomali tespiti gibi güncel yaklaşımların potansiyel katkılarını vurgulamaktadır [16].

Imanbayev ve arkadaşları (2022), 5G ve sonrası mobil ağlarda makine öğrenmesi ve derin öğrenme tabanlı saldırı tespit sistemlerini incelemiş, CICIDS2017 ve CSE-CIC-IDS2018 veri setlerini kullanarak lojistik regresyon, rastgele orman, gradyan artırma ve çeşitli derin öğrenme modellerinin performanslarını karşılaştırmıştır. Gradyan artırma yöntemi yüksek doğrulukla öne çıkarken, farklı ağ verilerinde model genellenebilirliğinin düştüğü görülmüştür. Çalışma, denetimli ve denetimsiz öğrenme yöntemlerinin avantajlarını değerlendirerek 5G için etkili IDS tasarımına yönelik öneriler sumuştur [17].

Yadav ve arkadaşları (2022), 5G tabanlı IoT ortamlarında siber saldırıların tespiti için derin öğrenme temelli bir saldırı tespit sistemi modeli önermiştir. UNSW-NB15 veri kümesi üzerinde yapılan çalışmada, Random Forest, XGBoost, Autoencoder tabanlı DNN gibi modeller karşılaştırılmıştır. Autoencoder destekli DNN modeli %99.76 doğruluk oranı ile öne çıkmıştır. Çalışma, 5G-IoT yapısında gerçek zamanlı çalışan ve yüksek doğrulukta tespit sağlayan derin öğrenme tabanlı çözümlerin etkinliğini ortaya koymaktadır [18].

Bu çalışmada 5G Çekirdek ağında meydana gelebilecek Komuta ve Kontrol (C2) saldırılarını tespit etmeyi hedefleyen bir yapay zeka modeli oluşturulmuştur. Açık kaynak yazılımlar kullanılarak bir 5G çekirdek ağı kurulmuş ve Komuta-Kontrol yazılımları kullanılarak çeşitli saldırılar gerçekleştirilmiştir. Gerçekleştirilen saldırılara ait ağ trafiği kayıt edilerek bir veri setine dönüştürülmüş ve uygun biçimde etiketlenmiştir. Elde edilen veri seti ile bir Yapay Sinir Ağı (YSA) modeli eğitilmiş; bu model, hem canlı ağ trafiği üzerinde hem de önceden kayıt edilmiş ağ trafiği verileri üzerinde saldırı tespiti amacıyla kullanılmıştır. Çalışma, yedi aşamada gerçekleştirilmiş olup, bu aşamalara ilişkin işlem adımlarının akış diyagramı Şekil 1'de sunulmuştur.



Şekil 1: Çalışmanın akış diyagramı

Metot ve Materyal

5G Çekirdek ağında meydana gelebilecek Komuta ve Kontrol (Command and Control – C2) saldırılarını tespit etmek amacıyla, Yapay Sinir Ağları'na (Artificial Neural Networks – ANN) dayalı bir derin öğrenme modeli geliştirilmiştir. Önerilen model, ağ trafiğinden elde edilen veri setleriyle eğitilerek, normal ağ etkinliklerine ait trafik ile saldırı kaynaklı trafiği yüksek doğrulukla ayırt edebilmektedir. Elde edilen sonuçlar, 5G çekirdek ağı güvenliğinin sağlanmasında derin öğrenme tabanlı yaklaşımların taşıdığı potansiyeli ortaya koymaktadır.

Çalışmada kullanılan 5G çekirdek ağ mimarisi; temel 5G ağ fonksiyonlarını barındıran Çekirdek (Core) ağı, dışarıya açık 5G ağ fonksiyonlarını barındıran Yalıtılmış Ağ Bölgesi (Demilitarized Zone - DMZ) ve 5G kullanıcılarına ait ekipmanların bulunduğu Radyo Erişim Ağı (Radio Access Network – RAN) olmak üzere üç kısımdan oluşmaktadır. Kurulmuş olan deney ağı; 12 adet 5G ağ fonksiyonu yürüten sanal sunucu, 1 adet baz istasyonu görevi yürüten sanal gNB sunucusu, 3 adet Kullanıcı Ekipmanı görevi yapan sanal bilgisayar ve 1 adet güvenlik duvarı sistemi ve 1 adet saldırgan sistem olmak üzere toplamda 18 sanal sistemden oluşmaktadır. Bu çalışmada kullanılan test ortamına ait ağ envanteri Tablo 1'de gösterilmiştir.

Tablo 1: 5G Çekirdek ağı envanteri

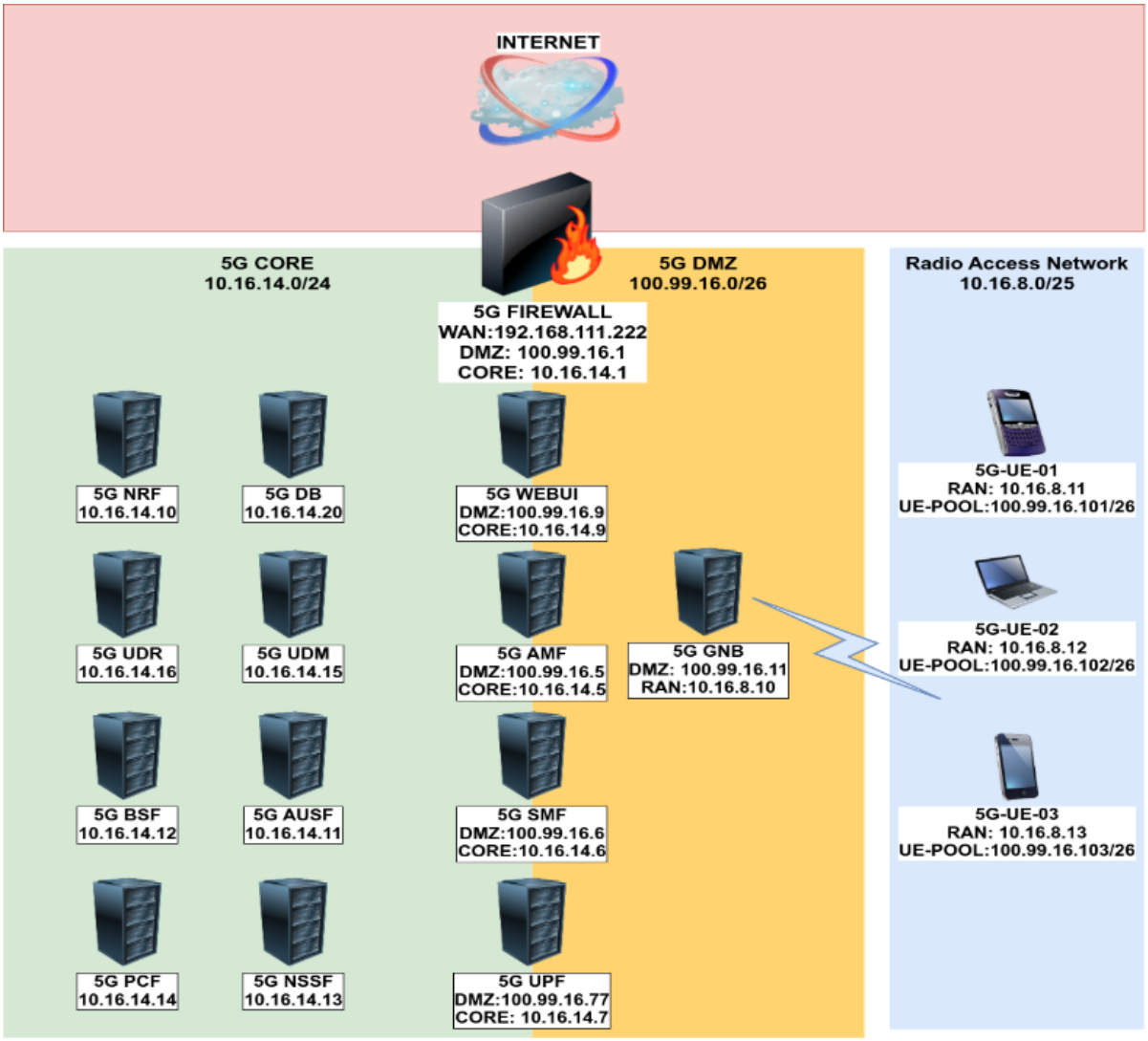
Cihaz	Yazılım	Adet
5G Ağ Fonksiyonları	Open5GS	12
gNB Baz İstasyonu	UERANSIM	1
Kullanıcı Cihazları	UERANSIM	3
Güvenlik Duvarı	pfSense	1
Saldırgan Sistem	Kali Linux	1

Deneysel Ortam

Linux Mint 21.2 Victoria işletim sistemi ile çalışan Intel i7-1255U işlemci ve 40GB belleğe sahip bir dizüstü bilgisayar sistemi kullanılmıştır. Çalışmada kullanılan bilgisayar sistemi her hangi bir GPU aygıtına sahip olmadığı için model eğitimi ve uygulama süreçlerinde yalnızca CPU kullanılmıştır. Bu sistem üzerinde sanallaştırma amacıyla Vmware Workstation 17.6.3 yazılımı kullanılmıştır.

Open5GS, 5G ve 4G LTE (4G Long Term Evolution) çekirdek ağı standartlarını destekleyen bir açık kaynak yazılım paketidir [19]. Bu çalışmada Open5GS v2.7.0 yazılım paketi

tercih edilmiş ve 5G ağ fonksiyonlarının her biri bir sanal bilgisayarda çalışacak şekilde Ubuntu Linux 22.04 üzerine kurulmuş ve 5G Çekirdek ağı ayrık mimari olarak hazırlanmıştır. Kurulan 5G ağ topolojisi Şekil 2'de verilmiştir.



Şekil 2: 5G Çekirdek ağ topolojisi

Her bir sanal bilgisayar, 5G ağ fonksiyonlarından birisini yerine getirecek şekilde ayrı ayrı kurulmuş ve yapılandırılmıştır. Application Function Manager (AFM), Uygulama fonksiyonlarının yönetiminden sorumlu olan bu bileşen, ağ politikalarıyla etkileşim kurarak uçtan uca hizmetlerin kontrolünü sağlar. Özellikle ağ dilimleme ve servis kalitesi yönetimi gibi konularda kritik rol oynar [2]. Session Management Function (SMF), Kullanıcı düzleminde oturum yönetimini üstlenen bu bileşen; IP adresi atama, oturum kurma, değiştirme ve sonlandırma gibi işlevleri yürütür. Ayrıca User Plane Function ile koordineli çalışarak veri düzlemine yön verir [2]. User Plane Function (UPF), Kullanıcı verisinin ağ üzerinden iletiminden sorumludur. Veri yönlendirme, paket inceleme ve hizmet kalitesi (Quality Of Service - QoS) uygulamaları gibi işlevleri gerçekleştirir [2]. Network Repository Function (NRF), Ağ fonksiyonlarının keşfi ve kayıtlarının merkezi olarak tutulmasını sağlayan bu bileşen, diğer ağ fonksiyonlarının birbirlerini bulmasını ve servis bazlı iletişim kurmasını mümkün kılar [2].

Unified Data Management (UDM), Kullanıcı profillerinin yönetimini gerçekleştiren bu bileşen, kimlik doğrulama, yetkilendirme ve kullanıcı verilerine erişim gibi işlevleri kontrol eder. UDR ile doğrudan veri alışverişi içerisindedir [2]. Unified Data Repository (UDR), Kullanıcıya ilişkin kalıcı verilerin (abonelik bilgileri, kimlik verisi vb.) saklandığı merkezi veri deposudur. Özellikle UDM gibi yönetim bileşenlerine veri sağlayıcı rolü üstlenir [2]. Binding Support Function (BSF), Uygulama fonksiyonları ile ağ politikaları arasında ilişkilendirme (binding) kurar. Belirli bir IP akışı için uygun ağ politikalarının uygulanmasını sağlar [2]. Authentication Server Function (AUSF), Kullanıcıların 5G çekirdek ağına erişimi sırasında kimlik doğrulama işlemlerini gerçekleştirir. Kimlik doğrulama süreçlerinde UDM ile birlikte çalışır [2]. Policy Control Function (PCF), Ağ genelinde uygulanacak politikaların belirlenmesi ve yönetilmesinden sorumludur. Hizmet kalitesi (QoS), abone bazlı kısıtlamalar ve erişim denetimi gibi konularda karar verici konumdadır [2]. Network Slice Selection Function (NSSF), Kullanıcıya en uygun ağ dilimini (network slice) seçen bileşendir. Dilim yönetimi, abonelik uyumluluğu ve ağ kaynaklarının verimli kullanımı açısından önem arz eder [2].Web Arabirimi (WEBUI), 5G çekirdek ağ bileşenlerinin ihtiyaç duyduğu abone bilgilerini düzenlemeyi sağlayan web arabimidir [20]. 5G Database (5GDB), UDR gibi 5G ağ bileşenlerinin veri saklama ihtiyacını karşılamak üzere, belge tabanlı, yüksek performanslı bir NoSQL veritabanı olan MongoDB kurulmuş ve yapılandırılmıştır.

Çalışmada kullanılan 5G ağ mimarisi; temel 5G ağ fonksiyonlarını barındıran Çekirdek (Core) ağı, dışarıya açık 5G ağ fonksiyonlarını barındıran Yalıtılmış Ağ Bölgesi (Demilitarized Zone - DMZ) ve 5G kullanıcılarına ait ekipmanların bulunduğu Radyo Erişim Ağı (Radio Access Network – RAN) olmak üzere 3 kısımdan oluşmaktadır.

Baz istasyonu ve kullanıcı cihazlarını (User Equipment - UE) simüle etmek için UERANSIM aracının v3.2.6 versiyonu Ubuntu Linux üzerinde kullanılmıştır [21]. UERANSIM, 5G kullanıcı ekipmanlarını (UE) ve yeni nesil baz istasyonlarını (gNB) yazılım tabanlı olarak simüle edebilen açık kaynaklı bir test aracıdır ve özellikle çekirdek ağların fonksiyonel doğrulama süreçlerinde yaygın olarak kullanılmaktadır [22].

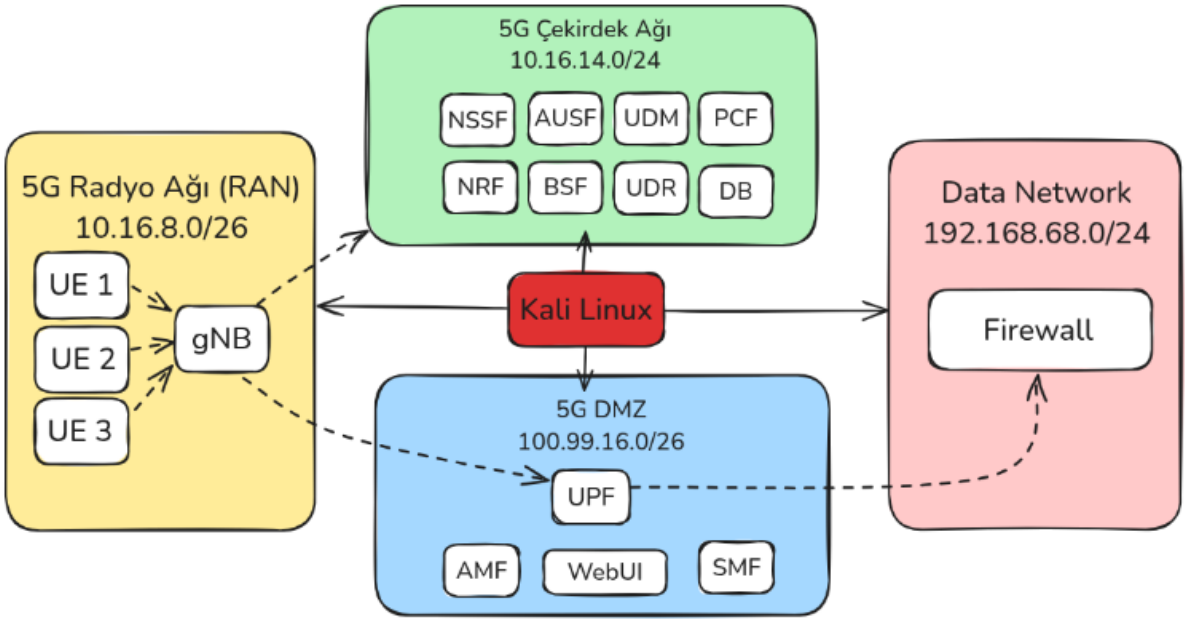
Next Generation Node B (gNB), 5G Yeni Radyo (5G NR) erişim teknolojisinin baz istasyonu olan bu bileşen, kullanıcı ekipmanları (User Equipment) ile 5G çekirdek ağı arasındaki kablosuz iletişimi sağlar. gNB, radyo kaynaklarının yönetimi, veri iletimi, sinyal işleme ve bağlantı kurulumu gibi işlevleri üstlenir. Ayrıca, Central Unit (CU) ve Distributed Unit (DU) gibi alt modüllere ayrılarak merkezi ve dağıtık mimarilerde görev yapabilir. Bu çalışmada Radyo Ağı ve gNB işlevini simüle etmek için UERANSIM uygulaması derlenmiş ve baz istasyonu modunda çalıştırılmıştır.

Kullanıcı Ekipmanları (User Equipment – UE), Linux sistemler üzerinde UERANSIM yazılımı kullanılarak kullanıcıların 5G ağına erişimini sağlayan akıllı telefonlar, tabletler, IoT cihazları veya modemler gibi cihazları simüle edilmiştir.

Tüm bu cihazların iletişimini sağlamak ve 5G ağı ile internet ortamı arasında ayrım oluşturmak amacıyla, sanal ortamda bir adet pfSense güvenlik duvarı (firewall) kurulmuş ve uygun şekilde yapılandırılmıştır. Alternatif olarak, pfSense gibi açık kaynaklı ve FreeBSD tabanlı bir güvenlik duvarı ve yönlendirme yazılımı olan OPNsense de tercih edilebilir.

5G ağ bileşenleri için hazırlanan sanal makineler Vmware NAT (Ağ Adres Çevrimi – Network Address Translation) ağına dahil edilmiştir. Tüm ağ segmentlerine ait trafiklerin, gerçek bilgisayar üzerinde çalışan saldırı tespit uygulaması tarafından izlenebilmesi için tüm sanal sistemler aynı sanal paket anahtarlama cihazına (switch) bağlanmış ve Vmware NAT arabirimi vmnet8, Network Test Access Point (TAP) cihazı gibi çalışacak şekilde yapılandırılmıştır.

Farklı saldırı senaryolarını gerçekleştirebilmek amacıyla Kali Linux dağıtımı kullanılmıştır. Kali Linux, Offensive Security tarafından geliştirilen, sızma testleri ve bilgi güvenliği denetimleri için kullanılan açık kaynaklı bir Linux dağıtımıdır [23]. Bu dağıtım, ağ analizi, tersine mühendislik, zafiyet tarama ve adli bilişim gibi alanlarda kullanılan yüzlerce ön yüklü güvenlik aracını içermektedir [24, 25]. Kali Linux; 5G çekirdek (core) ağı, 5G DMZ ağı, Radyo Erişim Ağı (RAN) ve veri ağı gibi farklı ağ segmentlerine bağlanmıştır. Oluşturulan 5G ağına ait şema Şekil 3'te sunulmaktadır.



Şekil 3: 5G Çekirdek ağ mimarisi

Saldırılar için kullanılan Kali Linux sanal makinesine, bağlı bulunduğu her bir ağ segmenti için uygun IP adreslerinden üç farklı IP adresi verilmiştir. 5G çekirdek ağı ve bileşenlerini hedef alan Komuta ve Kontrol (Command and Control - C2) saldırılarını gerçekleştirmek için Kali Linux üzerinde Metasploit Framework, Sliver ve Netcat gibi araçlar kullanılmıştır. Benzer amaçlarla alternatif olarak Parrot Security OS'de kullanılmaktadır.

Metasploit Framework, sızma testleri ve güvenlik açıklarını değerlendirmek amacıyla Rapid7 tarafından geliştirilen, açık kaynaklı ve modüler bir saldırı simülasyon platformudur [26]. Bishop Fox tarafından geliştirilen Sliver platformu, güvenlik araştırmaları ve saldırı simülasyonları için geliştirilmiş, açık kaynak kodlu bir saldırı sümülasyon aracı olup modern C2 (Command and Control) altyapısı, çapraz platform desteği ve gelişmiş gizlenme teknikleriyle saldırılarda ve sızma testlerinde sıkça kullanılmaktadır [27]. Netcat, TCP veya UDP bağlantıları kurarak veri aktarımı, port tarama ve komut çalıştırma gibi işlemlerde sızma testi uzmanları ve saldırganlar tarafından yaygın olarak kullanılmaktadır [28].

Bu araçlar ile üretilen zararlı uygulamalar 5G sistemlerinde çalıştırılmış ve gerçekleştirilen saldırıların ağ trafiği tshark aracı kullanılarak kayıt altına alınmış ve PCAP (Packet Capture) biçiminde kayıt edilmiştir. PCAPNG (PCAP Next Generation), PCAP formatının geliştirilmiş ve daha fazla özellik eklenmiş bir versiyonu olarak kullanılmaktadır. Tshark, popüler ağ dinleme yazılımı Wireshark'ın komut satırı versiyonudur. Ağ trafiğini gerçek zamanlı olarak analiz etmek, paketleri yakalamak ve filtrelemek amacıyla kullanılır [29]. Yakalanan paketleri detaylı incelemek için Wireshark aracı kullanılabilmektedir [30].

Veri Toplama ve Ön İşleme

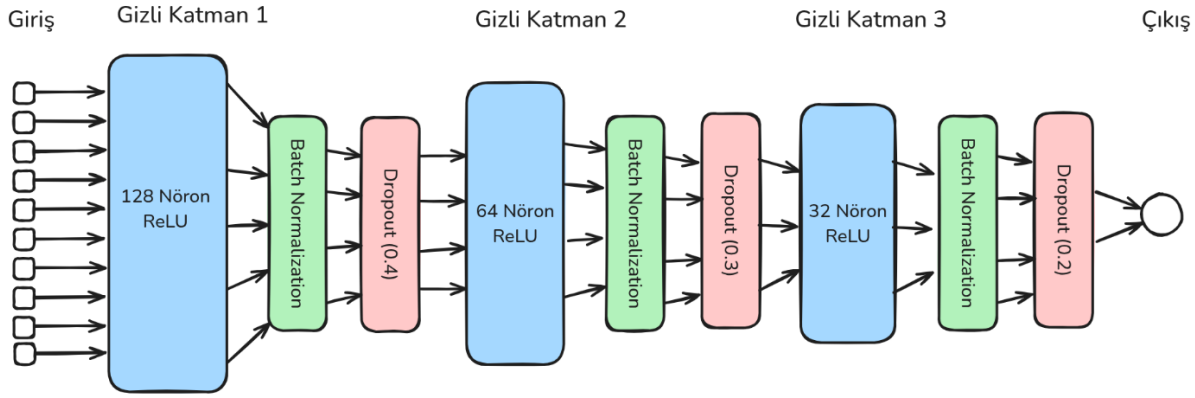
Veri toplama ve eğitim veri setinin oluşturulması için 5G ağ trafiği kaydedilmiş ve yakalanan trafik sınıflandırılmıştır. VMware sanallaştırma ortamında oluşturulan 5G Çekirdek ağının olağan ağ trafiği, Tshark aracı kullanılarak kayıt altına alınmış ve PCAP formatında depolanmıştır.

Saldırıları simule etmek için Kali Linux üzerinde Sliver, Metasploit ve Netcat araçları kullanılmıştır. Bu araçlar ile oluşturulan zararlı yazılımlar 5G sunucularında çalıştırılmış ve ters bağlantı (reverse connect) yöntemi ile Kali Linux'a bağlantı sağlanmıştır. Komuta ve Kontrol özelliğine sahip bu zararlı yazılımların oluşturduğu trafik, tshark aracı ile PCAP formatında kayıt edilmiştir. Kayıt edilen ".pcap" dosyaları bir python betiği ile ayrıştırılmış ve Timestamp, Protocol, Source IP, Source Port, Destination IP, Destination Port, Packet Length, TTL, ICMP Type, ICMP Code ve Flag alanları ayrıştırılarak csv (Comma-Separated Values) dosyalarına kayıt edilmiştir. Elde edilen csv dosyası, saldırılarda kullanılan IP adresleri ve port numaralarına göre filtrelenerek saldırı veya normal şeklinde etiketlenmiştir. Toplam 134.589 pakete ait veriler süzülerek, 95.938 adedi normal ağ trafiği ve 38.651 adedi ise saldırı olarak etiketlenerek eğitim veri seti oluşturulmuştur. 5G çekirdek ağına ait olağan trafiğin ve saldırılara ait trafiğin bulunduğu csv dosyaları Yapay Sinir Ağı (Artificial Neural Network) modelinin eğitimi için kullanılmıştır.

Derin Öğrenme Modeli

Yapay Sinir Ağları, insan beynindeki sinir hücrelerinin işleyişinden esinlenerek geliştirilen, öğrenme ve genelleme yeteneğine sahip hesaplama modelleridir [31]. Bu modeller, özellikle doğrusal olmayan karmaşık ilişkilerin tanımlanmasında ve öğrenilmesinde oldukça etkilidir. Yapay sinir ağları, insan beyninde olduğu gibi, deneyim yoluyla bilgi edinir ve bu bilgiyi daha önce karşılaşılmamış örnekler üzerinde uygulayarak tahmin ve sınıflandırma gibi görevleri yerine getirebilir [32]. Bu çalışmada saldırıları tespit etmek için bir Yapay Sinir Ağı (Artificial Neural Network - ANN) modeli kullanılmıştır ve kullanılan modelin yapısı Şekil 4'te verilmiştir.

Önerilen model, tam bağlantılı (fully connected) ve ileri beslemeli (feedforward) yapay sinir ağıdır. Model, 10 girdi özelliği almakta ve üç gizli katmana sahiptir. Modelin giriş katmanında ise; Protokol (Protocol), Kaynak IP (Source IP), Kaynak Port (Source Port), Hedef IP (Destination IP), Hedef Port (Destination Port), Paket Uzunluğu (Packet Length), TTL (Time to Live), ICMP Tipi (ICMP Type), ICMP Kodu (ICMP Code) ve bayraklar (flags) gibi veriler girdi olarak kullanılmaktadır.



Şekil 4: Yapay Sinir Ağı (ANN) modeli

Protokolleri sayısal olarak ifade etmek için; 6 (TCP), 17 (UDP), 1 (ICMP), 132 (SCTP) gibi protokol numaraları kullanılmıştır. IP Adreslerinin model tarafından işlenebilmesi için noktalı ondalık gösterim (dotted decimal notation) ile yazılmış olan IP adresleri tam sayı değerlere dönüştürülmüştür.

Çalışmada ele alınan Yapay Sinir Ağı (ANN) modeli, üç gizli katmandan oluşmaktadır. İlk katman, giriş verisinin boyutuna göre uyarlanmış 128 düğüm içeren yoğun (Dense) bir katmandır. Bu katman, doğrusal olmayan ilişkileri yakalamak amacıyla Rectified Linear Unit (ReLU) aktivasyon fonksiyonunu kullanmaktadır. Eğitim sürecini hızlandırmak, dahili kovaryans kaymalarını (internal covariate shift) azaltmak ve modelin daha hızlı yakınsamasını sağlamak üzere Batch Normalization (BN) katmanı eklenmiştir. Aşırı uyumu (overfitting) engellemek için Dropout kullanılmış ve %40 oranında rastgele nöron devre dışı bırakılarak model genelleştirilmiştir.

İkinci gizli katman, benzer şekilde tasarlanmış olup 64 düğüm içermekte ve yine ReLU aktivasyonu, Batch Normalization ve %30 oranında Dropout uygulamasıyla desteklenmektedir. Üçüncü katman ise daha düşük kapasiteli olup, 32 düğüm içermekte, ReLU aktivasyonu, Batch Normalization ve %20 Dropout ile tamamlanmıştır. Çıkış katmanı, saldırının varlığı veya yokluğunu ikili (binary) olarak sınıflandırmak üzere tek bir düğümden oluşmaktadır. Bu katman, çıktı değerini [0,1] aralığına sınırlayan sigmoid aktivasyon fonksiyonu kullanmaktadır. Modelin eğitiminde Adam (Adaptive Moment Estimation) optimizasyon algoritması tercih edilmiştir. Bu algoritma, adaptif öğrenme hızı sayesinde hızlı ve etkin bir yakınsama sağlamaktadır. Modelin öğrenme hızı (learning rate) 0.001 olarak belirlenmiştir. Kayıp fonksiyonunda kullanılan İkili Çapraz Entropi (Binary Cross Entropy), modelin normal ve saldırı sınıflarını doğru şekilde ayırt etmesini sağlar. Farklı değişkenlerin ölçeklendirme farklarını gidermek için Standard Scaler kullanılmıştır. Standardizasyon işlemi, tüm değişkenlerin ortalamasını 0, standart sapmasını 1 olacak şekilde dönüştürmeyi sağlar.

Veri kümesi, eğitim için %80 ve test için %20 oranında bölünmüştür. Eğitim verisi, modelin eğitilmesi için; test verisi ise modelin doğruluğunu değerlendirmek için kullanılmıştır. Model, 32 örnekten oluşan mini yığınlar (batch size) ile toplam 20 epok boyunca eğitilmiş ve eğitim süreci 208 saniye sürmüştür. Modelin performansını değerlendirmek için doğruluk (accuracy), kesinlik (precision) ve duyarlılık (recall) gibi metrikler tercih edilmiştir.

Değerlendirme Metrikleri

Modelin performans ölçümü için doğruluk, kesinlik, geri çağırma ve F1 Skoru gibi bazı standart metrikler kullanılmıştır. Kullanılan bu metrikler modellerin performanslarını değerlendirme için yaygın olarak kullanılan ölçütlerdir [33].

Doğruluk (Accuracy) metriği, sınıflandırma işleminin doğruluğunu ölçmeye ve modelin başarımını değerlendirmeye yardımcı olur. Doğruluk, şu şekilde hesaplanır:

$$Doğruluk = \frac{Doğru\ tespit\ edilen\ kayıtlar}{Toplam\ kayıtlar}$$

Kesinlik (Precision), pozitif olarak tahmin edilenlerin ne kadarının gerçekten pozitif olduğunu belirler ve aşağıdaki gibi hesaplanır:

$$Kesinlik = \frac{Gerçek\ pozitif}{Toplam\ tespit\ edilen\ pozitif}$$

Duyarlılık (Recall), pozitif olarak tahmin edilenlerin ne kadarının doğru şekilde sınıflandırıldığını belirlemek için kullanılır. Modelin duyarlılık (recall) değeri, şu şekilde hesaplanır:

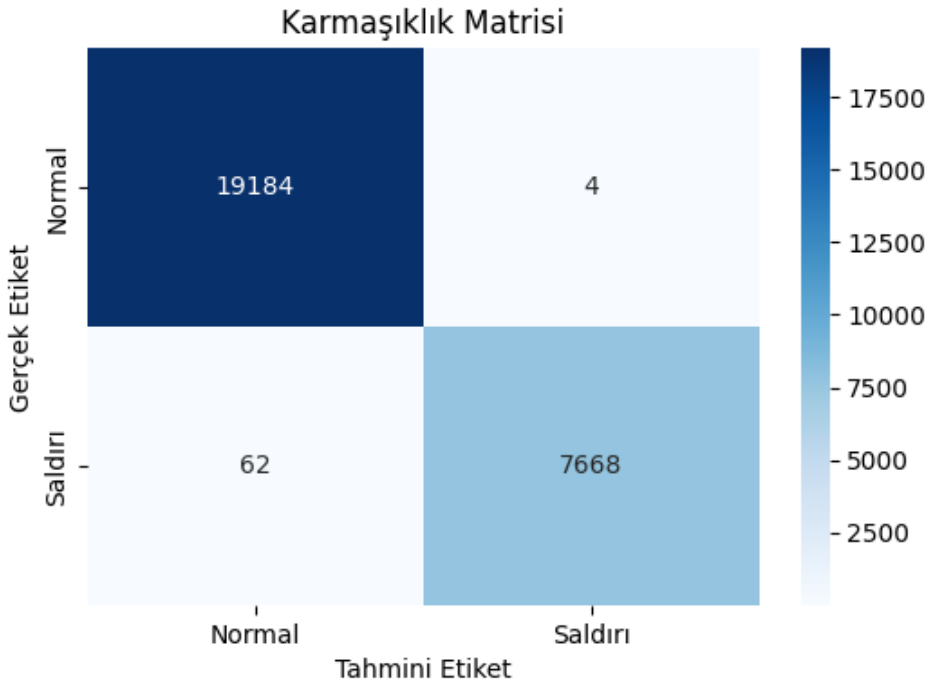
$$Duyarlılık = \frac{Gerçek\ pozitif}{Gerçek\ Pozitif + Yanlış\ Negatif}$$

F1-skoru, kesinlik (precision) ve duyarlılık (recall) arasındaki dengeyi gösterir. F1-skorunun değeri 0 ile 1 arasında olup, 1'e yakın olması modelin yüksek performans gösterdiğini belirtir. F1-skoru, şu şekilde hesaplanır:

$$F1\ Skoru = 2\ x \frac{Kesinlik\ x\ Duyarlılık}{Kesinlik +\ Duyarlılık}$$

Deneysel Sonuçlar

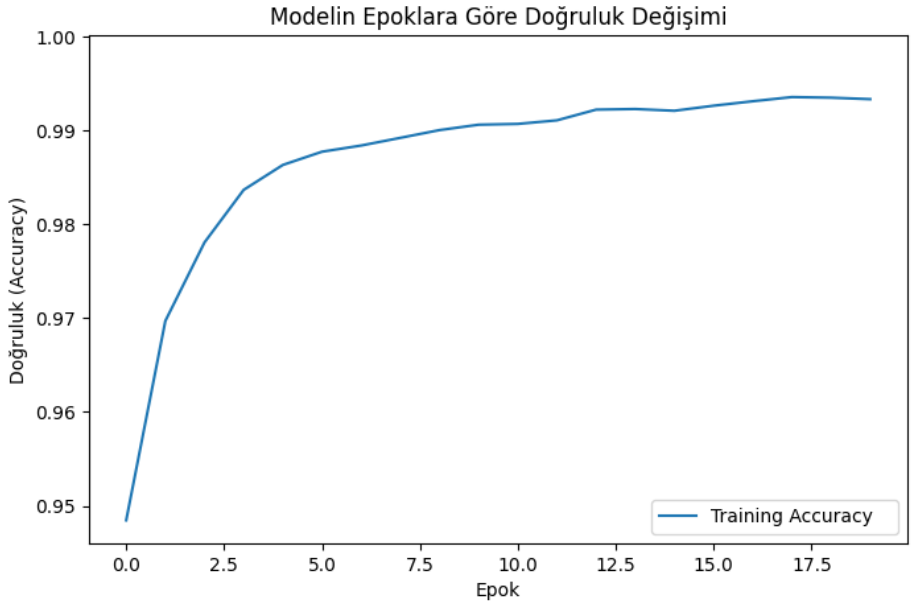
5G çekirdek ağına ait normal trafik ve saldırı trafiğinin kayıtlı bulunduğu CSV dosyaları kullanılarak model eğitimi yapılmış ve model dosyaları kaydedilmiştir. Eğitimin standart metrikleri Tablo 2'de, karmaşıklık matrisi Şekil 5'te ve doğruluk grafikleri Şekil 6'da sunulmuştur.



Şekil 5: Eğitim karmaşıklık matrisi

Tablo 2: Eğitime ait standart metrikler

Metrik	Adet
Doğruluk (Accuracy)	0.9975
Kesinlik (Precision)	0.9995
Duyarlılık (Recall)	0.9920
F1-Skoru	0.9957



Şekil 5: Eğitim doğruluk grafiği

Eğitilmiş modeli kullanarak saldırıları tespit edecek bir uygulama geliştirilmiştir. Eğitilmiş model kullanılarak geliştirilen bu uygulama, hem önceden PCAP formatında kaydedilmiş ağ trafiği üzerinde saldırı tespiti yapabilmekte hem de dinlediği ağ arabirimi üzerinden gerçek zamanlı ve etkileşimli saldırı tespiti gerçekleştirebilmektedir. Geliştirilen uygulama, 5G Çekirdek ağının çalışmakta olduğu Vmware NAT ağını dinleyerek ve yakaladığı paketleri sınıflandırarak saldırıları tespit etmiştir. Uygulamayı test etmek için eğitim veri setindeki saldırılardan farklı olacak şekilde saldırılar tasarlanmış ve uygulanmıştır. Kali Linux'a verilen ve saldırılar için kullanılan IP Adresleri Tablo 3'de verilmiştir.

Tablo 3: Saldırgan IP adresleri

Ağ Segmenti	**IP Adresleri**
Çekirdek Ağı (CORE)	10.16.14.21/24 10.16.14.81/24 10.16.14.98/24
DMZ Ağı	10.16.14.21/24 10.16.14.81/24 10.16.14.98/24
Radyo Erişim Ağı (RAN)	10.16.8.20/26 10.16.8.24/26 10.16.8.25/26
Data Ağı / Internet	10.16.14.21/24 10.16.14.81/24 10.16.14.98/24

Netcat ve Bash ile ters bağlantı (reverse shell) saldırıları yapılmıştır. Sliver ve Metasploit Framework araçları ile çeşitli IP adreslerine çeşitli portlardan ters bağlantı sağlayan zararlılar oluşturulmuş ve bu zararlılar 5G sunucularında çalıştırılarak modelin etkinliği test edilmiştir. Oluşturulan zararlıların türleri, bağlandığı IP adresleri, kullanılan protokoller ve portlar Tablo 4'de verilmiştir. Bu zararlılar, Erişim ve Mobilite Yönetim Fonksiyonu (Access and Mobility Management Function - AMF), Oturum Yönetim Fonksiyonu (Session Management Function - SMF), Kullanıcı Düzlemi Fonksiyonu (User Plane Function - UPF), Birleşik Veri Yönetimi (Unified Data Management - UDM) ve Politika Kontrol Fonksiyonu (Policy Control Function - PCF) sistemlerine yüklenerek çalıştırılmış ve zararlı yazılımlar tarafından Kali Linux'a bağlantı sağlanmıştır.

Tablo 4: Test zararlılarına ait bilgiler tablosu

Saldırgan IP	Yöntem	Protokol	Port
10.16.14.21	Sliver	http	9107
10.16.14.81	Sliver	mtls	22519
10.16.14.98	Netcat (Ters Kabuk)	tcp	5791
10.16.14.108	Netcat (Ters Kabuk)	tcp	61492
10.16.14.127	Meterpreter	tcp	7304
100.99.16.13	Sliver	http	30007
100.99.16.29	Sliver	mtls	2757
100.99.16.33	Netcat (Ters Kabuk)	tcp	18738
100.99.16.42	Netcat (Ters Kabuk)	tcp	62847
100.99.16.60	Meterpreter	tcp	2525
10.16.8.20	Sliver	http	15119
10.16.8.24	Sliver	mtls	2627
10.16.8.25	Netcat (Ters Kabuk)	tcp	3796
10.16.8.31	Netcat (Ters Kabuk)	tcp	2966
10.16.8.35	Meterpreter	tcp	15121
192.168.68.137	Sliver	http	1335
192.168.68.17	Sliver	mtls	12782
192.168.68.77	Netcat (Ters Kabuk)	tcp	7914
192.168.68.161	Netcat (Ters Kabuk)	tcp	9770
192.168.68.237	Meterpreter	tcp	49204

Elde edilen sliver bağlantılarına ait ekran görüntüsü Şekil 7'de ve Meterpreter bağlantılarına ait ekran görüntüsü ise Şekil 8'de görülebilir.

```
ID        Transport  Remote Address        Hostname  Username  Operating System  Health
========  =========  ====================  ========  ========  ================  ========
0090472b  mtls       10.16.14.14:54316     5g-pcf    root      linux/amd64       [ALIVE]
025b925a  mtls       10.16.14.5:40602      5g-amf    root      linux/amd64       [ALIVE]
20b72c05  http(s)    100.99.16.6:35966     5g-smf    root      linux/amd64       [ALIVE]
2c0818af  http(s)    10.16.14.7:57980      5g-upf    root      linux/amd64       [ALIVE]
313a7439  mtls       100.99.16.5:59206     5g-amf    root      linux/amd64       [ALIVE]
39914a28  mtls       10.16.14.6:52212      5g-smf    root      linux/amd64       [ALIVE]
44cc8a76  http(s)    100.99.16.5:48674     5g-amf    root      linux/amd64       [ALIVE]
4f0e4ac9  mtls       10.16.14.15:41528     5g-udm    root      linux/amd64       [ALIVE]
67be606c  http(s)    10.16.14.14:33342     5g-pcf    root      linux/amd64       [ALIVE]
7f1ba70a  http(s)    100.99.16.7:48936     5g-upf    root      linux/amd64       [ALIVE]
80e5a2e4  http(s)    10.16.14.6:58020      5g-smf    root      linux/amd64       [ALIVE]
83b1edfc  http(s)    10.16.14.15:37212     5g-udm    root      linux/amd64       [ALIVE]
8faa2f92  mtls       10.16.14.7:47836      5g-upf    root      linux/amd64       [ALIVE]
9619defc  mtls       10.16.14.14:56396     5g-pcf    root      linux/amd64       [ALIVE]
a17b415c  http(s)    10.16.14.14:45818     5g-pcf    root      linux/amd64       [ALIVE]
b73dc008  http(s)    10.16.14.15:36534     5g-udm    root      linux/amd64       [ALIVE]
bb88ccdf  mtls       10.16.14.15:43448     5g-udm    root      linux/amd64       [ALIVE]
c86f5566  mtls       100.99.16.7:34238     5g-upf    root      linux/amd64       [ALIVE]
d67065d1  mtls       100.99.16.6:43056     5g-smf    root      linux/amd64       [ALIVE]
f406ca96  http(s)    10.16.14.5:43892      5g-amf    root      linux/amd64       [ALIVE]

[server] sliver > 
```

Şekil 7: Sliver bağlantı listesi

Çalışma kapsamında geliştirilen yapay zeka uygulaması deney sırasında aktif olarak çalıştırılmıştır. Uygulama canlı ağ trafiğini analiz ederek yapılan saldırılara ait trafikleri tespit etmiş ve günlük (log) dosyalarına yazarak kayıt altına almıştır. Bu çalışmada geliştirilen uygulamaya ait ekran görüntüsü Şekil 9'da gösterilmiştir.

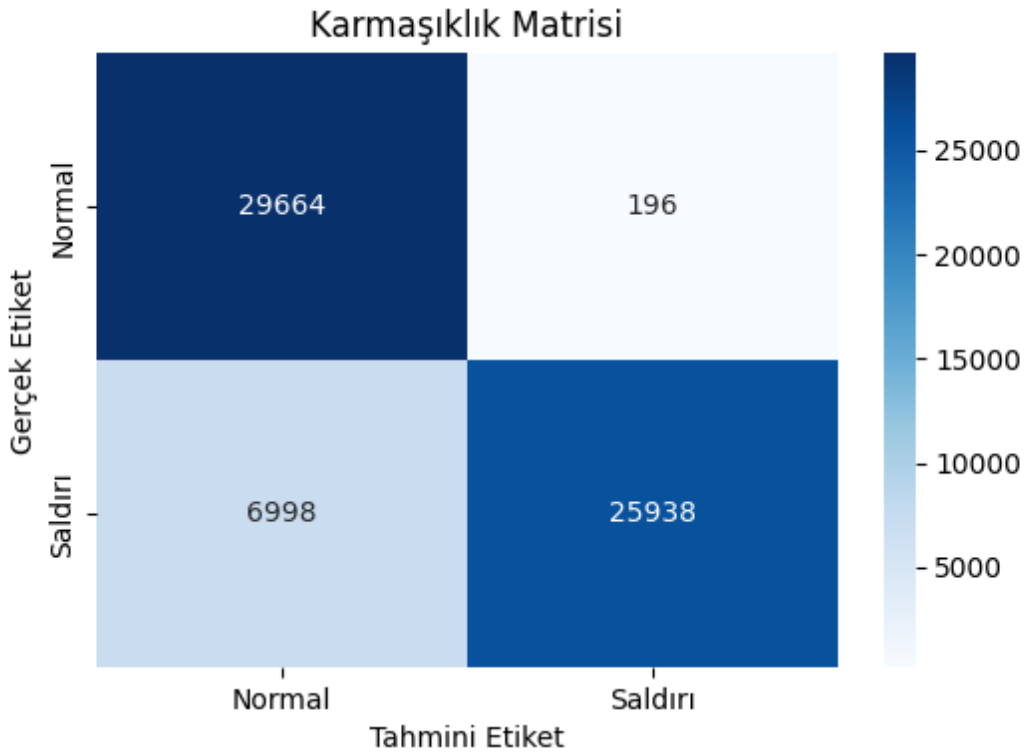
```
msf6 exploit(multi/handler) > sessions

Active sessions
===============

  Id  Name  Type                   Information          Connection
  --  ----  ----                   -----------          ----------
  1         meterpreter x64/linux  root @ 10.16.14.5    10.16.14.127:7304 → 10.16.14.5:47524 (10.16.14.5)
  2         meterpreter x64/linux  root @ 10.16.14.5    100.99.16.60:2525 → 100.99.16.5:33184 (100.99.16.5)
  3         meterpreter x64/linux  root @ 10.16.14.7    10.16.14.127:7304 → 10.16.14.7:44600 (10.16.14.7)
  4         meterpreter x64/linux  root @ 10.16.14.7    100.99.16.60:2525 → 100.99.16.7:47860 (100.99.16.7)
  5         meterpreter x64/linux  root @ 10.16.14.6    10.16.14.127:7304 → 10.16.14.6:36972 (10.16.14.6)
  6         meterpreter x64/linux  root @ 10.16.14.6    100.99.16.60:2525 → 100.99.16.6:44568 (100.99.16.6)
  7         meterpreter x64/linux  root @ 10.16.14.15   10.16.14.127:7304 → 10.16.14.15:33722 (10.16.14.15)
  9         meterpreter x64/linux  root @ 10.16.14.14   10.16.14.127:7304 → 10.16.14.14:53462 (10.16.14.14)

msf6 exploit(multi/handler) >
```

Şekil 8: Metasploit Meterpreter bağlantı listesi

```
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 10ms/step
Packet: Ether / IP / TCP 10.16.14.7:51854 > 10.16.14.81:22519 A -> Normal Traffic
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 10ms/step
Packet: Ether / IP / TCP 10.16.14.81:22519 > 10.16.14.7:51854 A -> Normal Traffic
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 12ms/step
Packet: Ether / IP / TCP 100.99.16.13:30007 > 10.16.14.14:41728 PA / Raw -> ATTACK DETECTED!
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 12ms/step
Packet: Ether / IP / TCP 10.16.14.14:41728 > 100.99.16.13:30007 FA -> ATTACK DETECTED!
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 11ms/step
Packet: Ether / IP / TCP 10.16.14.14:41728 > 100.99.16.13:30007 FA -> ATTACK DETECTED!
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 11ms/step
Packet: Ether / IP / TCP 100.99.16.13:30007 > 10.16.14.14:41728 FA -> ATTACK DETECTED!
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 11ms/step
Packet: Ether / IP / TCP 10.16.14.14:41728 > 100.99.16.13:30007 A -> ATTACK DETECTED!
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 12ms/step
Packet: Ether / IP / TCP 10.16.14.14:41728 > 100.99.16.13:30007 A -> ATTACK DETECTED!
1/1 ━━━━━━━━━━━━━━━━━━━━ 0s 12ms/step
Packet: Ether / IP / TCP 10.16.14.14:57656 > 100.99.16.13:30007 S -> ATTACK DETECTED!
```

Şekil 9: Canlı ağ trafiğinde ANN ile saldırı tespiti

Ek olarak, oluşturulan zararlı yazılım trafiği daha sonra analiz edilebilmesi amacıyla tshark aracı ile yakalanmış ve test sürecinde elde edilen ve ağ trafiği PCAP formatında kayıt edilmiştir. Toplam 62.796 adet ağ paketi kayıt edilmiştir. Bu kayıtların 36.662'si "normal" trafik olarak, 26.134'ü ise "saldırı" olarak etiketlenmiştir. 29.664 adet normal trafik doğru bir şekilde negatif (normal) olarak sınıflandırılmış, 25.938 adet saldırı verisi ise doğru bir şekilde pozitif (saldırı) olarak tespit edilmiştir. Saldırı aktivitesine ait olan 6.998 adet paket, model tarafından hatalı şekilde normal (False Negative) olarak değerlendirilmiştir. Ayrıca, 196 adet normal trafik ise model tarafından hatalı şekilde (False Positive) saldırı olarak sınıflandırılmıştır. Modelin ağ trafiğinden elde ettiği sonuçlara ait karmaşıklık matrisi Şekil 10'da ve standart değerlendirme metrikleri Tablo 5'de verilmiştir.



Şekil 10: Canlı ağ trafiği ile yapılan teste ait karmaşıklık matrisi

Tablo 5: Ağ testine ait standart metrikler

Metrik	Adet
Doğruluk (Accuracy)	0.8854
Kesinlik (Precision)	0.9925
Duyarlılık (Recall)	0.7875
F1-Skoru	0.8782

Sonuçlar

Çalışmada geliştirilen yapay zeka temelli saldırı tespit uygulaması, test ortamında gerçek zamanlı olarak çalıştırılmış ve hem normal hem de saldırı trafiğini başarılı şekilde ayrıştırmıştır. Gerçekleştirilen deneyler sonucunda uygulama, saldırı trafiğini tespit edip günlük dosyalarına kaydederek olayların belgelenmesini sağlamıştır. Ayrıca, test sürecinde üretilen ağ trafiği tshark aracıyla PCAP formatında yakalanarak kayıt altına alınmış ve bu sayede model çıktıları hem çevrim içi hem de çevrim dışı analizlerle doğrulanabilir hale getirilmiştir. Deney süresince toplam 62.796 adet ağ paketi toplanmıştır. Bu paketlerin 36.662'si «normal», 26.134'ü ise «saldırı» olarak etiketlenmiştir.

Model, bu veriler üzerinden yaptığı sınıflandırma sonucunda; 29.664 adet normal paketi doğru şekilde negatif (True Negative) ve 25.938 adet saldırı paketini doğru şekilde pozitif (True Positive) olarak sınıflandırmıştır. Ancak 6.998 saldırı paketi hatalı şekilde "normal" olarak tanımlanmış (False Negative), 196 normal paket ise saldırı olarak sınıflandırılmıştır (False Positive). Bu veriler, modelin ağırlıklı olarak yüksek kesinlik oranına sahip olduğunu; ancak

duyarlılık (recall) açısından bazı zayıflıklar barındırdığını ortaya koymaktadır. Elde edilen %88.54 Doğruluk (Accuracy) oranı, modelin genel sınıflandırma performansının yüksek olduğunu göstermektedir. Ancak bu oranın diğer metriklerle birlikte değerlendirilmesi gereklidir.

Kesinlik (Precision) değeri; modelin saldırı olarak sınıflandırdığı paketlerin ne kadarının gerçekten saldırı olduğunu gösterir. %99.25'lik yüksek Kesinlik (Precision) oranı, modelin yanlış alarm üretme olasılığının oldukça düşük olduğunu, modelin temkinli tahmin yaptığını ve emin olmadıkça saldırı olarak sınıflandırma yapmadığını göstermektedir. Bu durum, özellikle ağ güvenliği açısından kritik olup, sistemin gereksiz uyarı üretmemesi için oldukça önemlidir. Duyarlılık (Recall), Gerçek saldırıların ne kadarının doğru tespit edildiğini gösterir. %78.75 oranındaki duyarlılık, modelin belirli bir oranda saldırı trafiğini gözden kaçırdığını, yani saldırıların yaklaşık %21'inin tespit edilemediğini göstermektedir. Bu durum, modelin savunma açısından eksik kalabileceğini göstermektedir. F1 Skoru, Kesinlik ve duyarlılığın harmonik ortalamasını ifade eder. F1 skoru, modelin genel başarımını dengeli bir şekilde yansıtır. %87.82'lik F1 skoru, yüksek kesinlik ve orta düzeyde duyarlılığın birleşimiyle elde edilmiştir.

Geliştirilen yapay sinir ağı modelinin, özellikle normal ağ trafiğini öğrenme ve tespit etme konusunda oldukça başarılı olduğu, ağ üzerinde şüpheli olmayan trafiği büyük oranda doğru şekilde tanımlayabildiği görülmektedir. Ancak duyarlılık oranının görece düşük kalması, modelin saldırı trafiğini tam anlamıyla yakalayamadığını ve saldırılara ait bazı trafikleri kaçırarak "normal" olarak kabul ettiğini göstermektedir.

Gelecekteki çalışmalarda, modelin duyarlılık oranının artırılmasına yönelik stratejilerin uygulanması önerilmektedir. Özellikle, farklı saldırı türlerinin dengeli bir şekilde temsil edildiği, daha geniş kapsamlı veri kümeleri ile yeniden eğitim yapılması, modelin genelleme yeteneğini geliştirebilir. Ayrıca, yanlış negatif (false negative) oranını azaltmaya yönelik olarak, özellik mühendisliği, LSTM (Long Short-Term Memory), GRU (Gated Recurrent Unit) gibi zaman serisi odaklı farklı mimarilerin denenmesi ve hibrit sistemlerin kullanılması değerlendirilebilir. Bunun yanı sıra, RVFL (Random Vector Functional Link) [34] ve ELM (Extreme Learning Machine) [35] tabanlı çalışmalar, önerilen modelin performansını artırma potansiyeline sahiptir. Ayrıca, açıklanabilir yapay zekâ (XAI) yöntemlerinin entegrasyonu [36], modelin karar süreçlerinin daha anlaşılır ve şeffaf hale gelmesini sağlayarak, kullanıcı güvenini artırabilir ve modelin hata analizini kolaylaştırabilir. Sonuç olarak, geliştirilen sistem, gerçek zamanlı ağ trafiği izleme ve saldırı tespiti açısından umut verici bir performans sergilemekte olup, özellikle düşük yanlış alarm oranı sayesinde pratik güvenlik sistemlerine entegre edilebilecek niteliktedir.

Çıkar Beyanı**:** Yazarlar çıkar çatışması olmadığını beyan etmişlerdir.

Yazar Katkıları: Yazarlar makaleye eşit oranda katkı sağlamış olduklarını beyan etmişlerdir.

REFERANSLAR

[1] J. G. Andrews, S. Buzzi, W. Choi, et al., "What Will 5G Be?" IEEE Journal on Selected Areas in Communications, vol. 32, no. 6, pp. 1065–1082, 2014. [Online]. Available: https://ieeexplore.ieee.org/document/6824752.

[2] TSGS, "TS 123501-V17.11.0-5G; System architecture for the 5G System (5GS)," 3Gpp, vol. 0, 2024. [Online]. Available: https://portal.etsi.org/TB/ETSIDeliverableStatus.aspx.

[3] M. Alshaikh, W. Alsemaih, S. Alamri, and Q. Ramadan, "Using supervised learning to detect command and control attacks in IoT," International Journal of Cloud Applications and Computing, vol. 14, pp. 1–19, Jan. 2024. DOI: 10.4018/IJCAC.334214.

[4] T. S. Othman and S. M. Abdullah, "Intrusion Detection Systems for IoT Attack Detection and Identification Using Intelligent Techniques," International Journal of Computer Networks and Applications, vol. 10, no. 1, pp. 130–143, 2023. ISSN: 23950455. DOI: 10.22247/ijcna/2023/218517.

[5] M. Prasad, S. Tripathi, and K. Dahal, "An efficient feature selection based Bayesian and rough set approach for intrusion detection," Applied Soft Computing, vol. 87, p. 105980, 2020. ISSN: 15684946. DOI: https://doi.org/10.1016/j.asoc.2019.105980. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1568494619307616.

[6] M. A. Ferrag, L. Maglaras, S. Moschoyiannis, and H. Janicke, "Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study," Journal of Information Security and Applications, vol. 50, p. 102419, 2020. ISSN: 2214-2126. DOI: https://doi.org/10.1016/j.jisa.2019.102419. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214212619305046.

[7] R. Khan, P. Kumar, D. N. K. Jayakody, and M. Liyanage, "A survey on security and privacy of 5G technologies: Potential solutions, recent advancements, and future directions," IEEE Communications Surveys & Tutorials, vol. 22, no. 1, pp. 196–248, 2020. DOI: 10.1109/COMST.2019.2933899.

[8] A. Gupta, R. K. Jha, P. Gandotra, and S. Jain, "Bandwidth spoofing and intrusion detection system for multistage 5G wireless communication network," IEEE Transactions on Vehicular Technology, vol. 67, no. 1, pp. 618–632, 2018. DOI: 10.1109/TVT.2017.2745110.

[9] R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat, and S. Venkatraman, "Deep learning approach for intelligent intrusion detection system," IEEE Access, vol. 7, pp. 41525–41550, 2019. DOI: 10.1109/ACCESS.2019.2895334.

[10] S. Fonyi, "Overview of 5G security and vulnerabilities," The Cyber Defense Review, vol. 5, no. 1, pp. 117–134, 2020. [Online]. Available: https://www.jstor.org/stable/10.2307/26902666.

[11] Y. Zhang, P. Li, and X. Wang, "Intrusion detection for IoT based on improved genetic algorithm and deep belief network," IEEE Access, vol. 7, pp. 31711–31722, 2019. DOI: 10.1109/ACCESS.2019.2903723.

[12] H. Liu and B. Lang, "Machine learning and deep learning methods for intrusion detection systems: A survey," Applied Sciences, vol. 9, no. 20, p. 4396, 2019. DOI: 10.3390/app9204396. [Online]. Available: https://doi.org/10.3390/app9204396.

[13] K. Gai, M. Qiu, L. Tao, and Y. Zhu, "Intrusion detection techniques for mobile cloud computing in heterogeneous 5G," Security and Communication Networks, vol. 9, no. 16, pp. 3049–3058, 2015. DOI: 10.1002/sec.1224. [Online]. Available: https://doi.org/10.1002/sec.1224.

[14] S. Samarkoon, Y. Siriwardhana, P. Porambage, et al., "5G-NIDD: A Comprehensive Network Intrusion Detection Dataset Generated over 5G Wireless Network," arXiv preprint arXiv:2212.01298, 2022. [Online]. Available: https://arxiv.org/abs/2212.01298.

[15] A. Gupta, R. K. Jha, and S. Jain, "Attack Modeling and Intrusion Detection System for 5G Wireless Communication Network," International Journal of Communication Systems, vol. 30, no. 10, e3237, 2017. DOI: 10.1002/dac.3237.

[16] C. Hamroun, A. Fladenmuller, M. Pariente, and G. Pujolle, "Intrusion Detection in 5G and Wi-Fi Networks: A Survey of Current Methods, Challenges, and Perspectives," IEEE Access, vol. 13, pp. 40950–40994, 2025. DOI: 10.1109/ACCESS.2025.3546338.

[17] A. Imanbayev, S. Tynymbayev, R. Odarchenko, et al., "Research of Machine Learning Algorithms for the Development of Intrusion Detection Systems in 5G Mobile Networks and Beyond," Sensors, vol. 22, no. 24, p. 9957, 2022. DOI: 10.3390/s22249957.

[18] N. Yadav, S. Pande, A. Khamparia, and D. Gupta, "Intrusion Detection System on IoT with 5G Network Using Deep Learning," Wireless Communications and Mobile Computing, vol. 2022, Article ID 9304689, 13 pages, 2022. DOI: 10.1155/2022/9304689.

[19] Open5GS Project. (n.d.). Open5GS: Open Source 5G/4G Core Network Project. [Online]. Available: https://open5gs.org/.

[20] Open5GS Project, Install the WebUI of Open5GS. [Online]. Available: https://open5gs.org/open5gs/docs/guide/01-quickstart/.

[21] A. Güngör, UERANSIM. [Online]. Available: https://github.com/aligungr/UERANSIM.

[22] K. Du, L. Wang, X. Wen, Y. Liu, H. Niu, and S. Huang, "ML-SLD: A message-level stateless design for cloud-native 5G core network," Digital Communications and Networks, vol. 9, no. 3, pp. 743–756, 2023. ISSN: 2352-8648. DOI: https://doi.org/10.1016/j.dcan.2022.04.026. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2352864822000815.

[23] Offensive Security, Kali Linux: Penetration Testing and Ethical Hacking Linux Distribution, 2025. [Online]. Available: https://www.kali.org/.

[24] V. Visoottiviseth, P. Akarasiriwong, S. Chaiyasart, and S. Chotivatunyu, "PENTOS: Penetration testing tool for Internet of Thing devices," in TENCON 2017 – 2017 IEEE Region 10 Conference, 2017, pp. 2279–2284. DOI: 10.1109/TENCON.2017.8228241.

[25] H. Parmaksız, "Cyber Security Basics: Threats, Protection Methods, Open Source Software, Security Compliance Management with Artificial Intelligence and Trend Analysis," In Current Topics in Management Information Systems , Gaziantep: Özgür Yayinlari, 2024, pp.35-54. [Online]. Available: https://www.ozguryayinlari.com/site/catalog/download/498/2314/5357

[26] Rapid7, Metasploit Framework: Penetration Testing Software, 2025. [Online]. Available: https://www.metasploit.com/.

[27] B. Fox, Sliver: Open Source Adversary Emulation Framework, 2025. [Online]. Available: https://sliver.sh/ (visited on 04/04/2025).

[28] G. Giacobbi, Netcat: The Network Swiss Army Knife. [Online]. Available: https://netcat.sourceforge.net/.

[29] Wireshark Foundation, TShark: Terminal-Based Network Protocol Analyzer, 2025. [Online]. Available: https://www.wireshark.org/docs/man-pages/tshark.html.

[30] Wireshark Foundation, Wireshark: The World's Foremost Network Protocol Analyzer, 2025. [Online]. Available: https://www.wireshark.org/.

[31] G. Zhang, B. Eddy Patuwo, and M. Y. Hu, "Forecasting with artificial neural networks: The state of the art," International Journal of Forecasting, vol. 14, no. 1, pp. 35–62, 1998. ISSN: 0169-2070. DOI: https://doi.org/10.1016/S0169-2070(97)00044-7. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169207097000447.

[32] S. Haykin, Neural Networks and Learning Machines, 3rd ed., Pearson, 2009. ISBN: 978-0-13-147139-9.

[33] D. Powers, "Evaluation: From precision, recall and f-factor to ROC, informedness, markedness correlation," Mach. Learn. Technol., vol. 2, Jan. 2008.

[34] A. K. Malik, R. Gao, M. A. Ganaie, M. Tanveer, and P. N. Suganthan, "Random vector functional link network: Recent developments, applications, and future directions," Appl. Soft Comput., vol. 143, p. 110377, 2023.

[35] H. Parmaksız and C. Karakuzu, "Performance analysis of Extreme Learning Machine classifiers on radio frequency fingerprinting," BSEU J. Eng. Res. Technol., 2022. [Online]. Available: https://bseujert.bilecik.edu.tr/index.php/bseujert/issue/download/5/23

[36] H. Parmaksız and Ö. Öztürk, "Feature selection strategies: A comparative analysis of SHAP-XAI and importance-based methods in customer churn data," Proc. Int. Conf. Appl. Econ. Finance (ICOAEF-XII), Berlin, Germany, Oct. 18-19, 2024, pp. 316-342, ISBN 978-975-06-4614-0. [Online]. Available: https://www.icoaef.com/wp-content/uploads/2024/12/ICOAEF-berlin-2024-tam-KI_TA-BI-ISBN-SON-HALI-10-ARALIK-1.pdf

This publication contains the peer-reviewed full-text papers and abstracts presented at the 5th International Conference on Artificial Intelligence and Data Science, organized through the collaboration of İzmir Kâtip Çelebi University's Artificial Intelligence and Data Science Application and Research Center and Zonguldak Bülent Ecevit University's Karaelmas Cybersecurity Application and Research Center. The publication covers a wide range of topics, including machine learning algorithms, AI-based threat detection and defense mechanisms, data analytics in cybersecurity, ethical considerations in artificial intelligence, and the applications of AI and data science in diverse sectors such as engineering, healthcare, education, law, finance, and urban planning. We are confident that this conference, which we deem significant as the collaboration between two distinct university research centers in "Artificial Intelligence" and "Cybersecurity" marks a first in our country, will contribute to the "Digital Türkiye" vision and the "National Technology Initiative" development goals outlined in the "National Artificial Intelligence Strategy."

İKÇÜ YAYINEVİ
ikcupress