Volume/Cilt: 9
Issue/Sayı: 1
ISNN: 2618 6055

2025

-USMTD-
Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi

-IJSET-
International Journal of Sustainable Engineering and Technology

Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi
International Journal of Sustainable Engineering and Technology
ISSN: 2618-6055 / Cilt: 9, Sayı: 1, (2025)

BAŞ EDİTÖR / EDITOR-IN-CHIEF

Prof. Dr. Cengiz ÖZEL
Isparta Uygulamalı Bilimler Üniversitesi

ONURSAL EDİTÖR / HONORARY EDITOR
Prof. Dr. Reşat SELBAŞ
Isparta Uygulamalı Bilimler Üniversitesi

ALAN EDİTÖRLERİ / SECTION EDITORS

Prof. Dr. Hilmi Cenk BAYRAKÇI
Isparta Uygulamalı Bilimler Üniversitesi
Dr. Öğr. Üyesi Tolgahan ERMERGEN
Isparta Uygulamalı Bilimler Üniversitesi
Doç. Dr. Koray ÖZSOY
Isparta Uygulamalı Bilimler Üniversitesi
Dr. Öğr. Üyesi Ali Nadi KAPLAN
Isparta Uygulamalı Bilimler Üniversitesi
Doç. Dr. Berkay ERGENE
Pamukkale Üniversitesi

Doç. Dr. Kubilay TAŞDELEN
Isparta Uygulamalı Bilimler Üniversitesi
Doç. Dr. Bekir AKSOY
Isparta Uygulamalı Bilimler Üniversitesi
Dr. Öğr. Üyesi Merdan ÖZKAHRAMAN
Isparta Uygulamalı Bilimler Üniversitesi
Dr. Öğr. Üyesi Yaşar Kemal ERDOĞAN
Isparta Uygulamalı Bilimler Üniversitesi

YAYIN DANIŞMA KURULU / EDITORIAL ADVISORY BOARD

Prof. Dr. İbrahim ÜÇGÜL
Süleyman Demirel Üniversitesi
Doç. Dr. Shoaib KHANMOHAMMADİ
Kermanshah University of Technology

Prof. Dr. Osman GENÇEL
Bartın Üniversitesi
Dr. Deeb E. A. ALASHGAR
Qatar Environment and Energy Research Institute (QEERI)
Hamad Bin Khalifa University

MİZANPAJ EDİTÖRLERİ / LAYOUT EDITORS

Arş. Gör. Ali Ekber SEVER
Isparta Uygulamalı Bilimler Üniversitesi
Arş. Gör. Nergiz AYDİN
Isparta Uygulamalı Bilimler Üniversitesi

Arş. Gör. Yusuf YAŞIR
Isparta Uygulamalı Bilimler Üniversitesi

DİL EDİTÖRÜ (TÜRKÇE) / LANGUAGE EDITOR (TURKISH)

Arş. Gör. Elifnur ŞAKALAK
Isparta Uygulamalı Bilimler Üniversitesi

DİL EDİTÖRÜ (İNGİLİZCE) / LANGUAGE EDITOR (ENGLISH)

Arş. Gör. Mehmet YÜCEL
Isparta Uygulamalı Bilimler Üniversitesi

TEKNİK EDİTÖR/ TECHNICAL EDITOR

Arş. Gör. Yakup Hakan AYDİN
Isparta Uygulamalı Bilimler Üniversitesi

Arş. Gör. Mert ÖPÖZ
Isparta Uygulamalı Bilimler Üniversitesi

Yazıların tüm bilimsel sorumluluğu yazar(lar)a aittir. Editör, yardımcı editör ve yayıncı dergide yayınlanan yazılar için herhangi bir sorumluluk kabul etmez. Bu dergi, aşağıda listelenen veri tabanları tarafından taranmaktadır.

All the scientific responsibilities of the manuscripts belong to the authors (s). The editor, assistant editor and publisher accept no responsibility for the articles published in the journal. The Journal is indexed by the following abstracting and indexing databases.

Index Copernicus, Google Scholar, Sobiad, Sindex, DRJI

Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi

International Journal of Sustainable Engineering and Technology

Cilt: 9, Sayı: 1, Yıl: 2025

Volume: 9, Number: 1, Year: 2025

İÇİNDEKİLER / CONTENTS

ISSN: 2618-6055 / 9, (1), 1 – 13, 2025
DOI:10.62301/usmtd.1704012

Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi
International Journal of Sustainable Engineering and Technology
ISSN: 2618-6055 / Cilt:9, Sayı:1, (2025), 1 – 13
Araştırma Makalesi – Research Article

INVESTIGATION OF PHYSICAL AND CHEMICAL PROPERTIES OF BITUMEN MODIFIED WITH WASTE ENGINE AND INDUSTRIAL OILS

*Gülşah ÖZ KICI[1], Mehmet SALTAN[1]

[1]Süleyman Demirel University, Faculty of Engineering and Natural Sciences, Department of Civil Engineering, Isparta

(Geliş/Received: 22.05.202, Kabul/Accepted: 27.05.2025, Yayınlanma/Published: 30.06.2025)

ABSTRACT

The modification of bitumen with waste oils encourages the transition5 to sustainable production and consumption strategies by increasing resource efficiency and reducing environmental impacts. In this study, bitumen was modified with waste engine oil (WEO) and waste industrial oil (WIO) at the rate of 2%, 2.5%, 3%, 3.5%, and 4% by weight to increase resource efficiency and contribute to sustainable transportation. Physical tests (penetration, softening point, penetration index, rotational viscometer) were applied to the oil modified bitumens, and the chemical composition of the bitumen as a result of the modification was examined using FTIR, SEM, and EDS analyses. Results indicate that modified bitumens exhibit superior performance in cold climate regions. The modification of 4% WIO and WEO into bitumen reduced the mixing temperature by 3.5% and 5%, respectively, and the compaction temperature by 5% and 10%, respectively. The findings of the study indicate that WIO and WEO can be used in bitumen modification to improve its performance. This study contributes to sustainable production practices by not only utilizing waste materials in a sustainable manner but also reducing the environmental impacts of industrial processes.

Keywords: Sustainability, Bitumen, Waste engine oil, Waste industrial oil.

ATIK MOTOR VE ENDÜSTRİYEL YAĞLAR İLE MODİFİYE EDİLMİŞ BİTÜMÜN FİZİKSEL VE KİMYASAL ÖZELLİKLERİNİN İNCELENMESİ

ÖZ

Bitümün atık yağlar ile modifikasyonu, kaynak verimliliğini artırarak ve çevresel etkileri azaltarak sürdürülebilir üretim ve tüketim modellerine geçişi teşvik etmektedir. Bu çalışmada, kaynak verimliliğini arttırmak ve sürdürülebilir ulaşıma katkı sağlamak amacıyla bitüm ağırlıkça %2, %2.5, %3, %3.5 ve %4 oranlarında atık motor yağı (AMY) ve atık endüstriyel yağ (AEY) ile modifiye edilmiştir. Modifiye edilen bitümlere fiziksel deneyler (penetrasyon, yumuşama noktası, penetrasyon indeksi, dönel viskozimetre) uygulanmış ve modifiye bitümlerin kimyasal içeriği FTIR, SEM ve EDS analizleri ile incelenmiştir. Çalışma sonucunda, modifiye edilen bitümlerin hava sıcaklığının düşük olduğu bölgelerde kullanımının daha uygun olacağı tespit edilmiştir. AEY ve AMY ile %4 oranında modifiye edilen bitümlerin karıştırma sıcaklığı sırasıyla %3.5 ve %5 oranında azalırken, sıkıştırma sıcaklığı ise %5 ve %10 oranında azalmıştır. Çalışmadan elde edilen sonuçlara göre AEY ve AMY'lerin bitüm performansını arttırmak amacıyla bitüm modifikasyonunda kullanılabileceği gösterilmiştir. Bu çalışma, atık malzemelerin sürdürülebilir bir şekilde kullanılmasının yanı sıra endüstriyel süreçlerin çevresel etkilerini azaltarak sürdürülebilir üretim uygulamalarına katkıda bulunmaktadır.

Anahtar Kelimeler: Sürdürülebilirlik, Bitüm, Atık motor yağı, Atık endüstriyel yağ.

***Corresponding Author:** gulsahozkici@gmail.com

ISSN: 2618-6055 / 9, (1), 1 – 13, 2025
DOI:10.62301/usmtd.1704012

1. Introduction

Bituminous binders used in hot mix asphalt (HMA), which exhibit high thermal sensitivity, suffer various thermal changes during the road construction process. In addition, due to the harsh environmental effects and variable traffic loads they are subjected to throughout their service life, changes occur in the properties of bituminous binders both during the construction and use of HMA [1]. This phenomenon, referred to as the aging of the bituminous binder, causes the binder to gradually lose its volatile components and fail to perform as expected [2]. The aging of bituminous binders is also one of the reasons for the deterioration that may occur in the pavement layer [3].

Bituminous binder modification is carried out to enhance the resistance of pavement layers to temperature fluctuations, reduce cracking and deformation to ensure longer service life, increase skid resistance to improve safety, provide protection against environmental impacts, and promote the use of recyclable materials, thereby contributing to environmental sustainability [4].

The use of waste oil in bitumen modification contributes to environmental sustainability [5] while improving the rheological properties and temperature resistance of bitumen, offering a cost-effective solution, enhancing crack and deformation resistance, and facilitating the application process by ensuring mixture homogeneity. Various studies have been conducted considering these effects. The results have shown that bitumen modified with waste engine oil exhibits lower thermal sensitivity, viscosity, and rutting resistance but higher fatigue resistance and temperature sensitivity [6,7]. In their study, Yalçın and Yılmaz [8] aimed to evaluate the physical, mechanical, and rheological properties of waste oil (WO) modification. It was determined that the viscosity and softening point of bituminous binders decreased, while the penetration value increased when modified using waste engine oil and waste vegetable oil. Additionally, as the additive content increased, the complex modulus values decreased for both WO modifications, and the increase in phase angle indicated that the binders exhibited a more viscous behavior. The use of filtered waste engine oil resulted in a 35% reduction in binder hardness compared to base bitumen after short-term aging [9]. In their study, Sani et al. [10] investigated the use of waste engine oil in warm mix asphalt. The performance of modified bitumen (MB), obtained using different percentages of waste engine oil, was evaluated in terms of penetration, softening point, Marshall stability, flow, and stiffness. The results indicated that waste engine oil modification could improve the performance properties of Warm Mix Asphalt, such as stability, flow, and stiffness, while also having a positive effect on penetration and softening point.

Furthermore, in various studies, the combined effects of using waste engine oil alongside different modifiers have been investigated. In their study, Liu, Li, and Wang [11] utilized waste engine oil and polyphosphoric acid for bitumen modification. It was observed that the combined modification improved binder properties, such as rutting resistance and temperature stability, which were weakened by WO modification alone. A mere 2% polyphosphoric acid modification was found to enhance the rutting factor obtained from DSR analysis by 31.2% compared to base bitumen. To enhance the beneficial effects of waste engine oil, phosphogypsum waste was used in bituminous binder modification, demonstrating superior performance in terms of high-temperature deformation resistance, low-temperature crack resistance, and moisture sensitivity [12]. Based on performance test results regarding high-temperature stability, low-temperature crack resistance, and fatigue life, the optimal ratio for the modified asphalt binder was determined to be 2% waste engine oil and 5% phosphogypsum. The combined use of waste engine oil and SBS resulted in a MB with improved fatigue crack resistance and rutting resistance [13]. The combined use of waste vehicle lubricants and waste polymers was found to enhance high-temperature stability and creep resistance [14].

In road construction, the use of different types of waste oils is recommended to reduce the stiffening effect of Reclaimed Asphalt Pavement. Studies conducted for this purpose have shown that waste engine oils can be used as rejuvenating agents in mixtures containing reclaimed asphalt, with notable improvements in moisture sensitivity and stability [15-18]. Additionally, it has been observed that waste engine oils enhance low-temperature properties [19,20]. The similar chemical structure of waste engine oil and asphalt materials allows the recovery of components lost during asphalt aging through waste engine oil modification, thereby improving the properties of aged asphalt in both binder and mixture forms [21-24]. The rejuvenation of Reclaimed Asphalt Binder using waste industrial oil, specifically waste steel rolling oil, has been investigated, showing increased resistance to moisture sensitivity and the ability to reverse aging effects [25].

*Corresponding Author: gulsahozkici@gmail.com

ISSN: 2618-6055 / 9, (1), 1 – 13, 2025
DOI:10.62301/usmtd.1704012

This study aims to comprehensively evaluate the potential of using waste motor and industrial oils in the modification of bitumen through a combination of physical and chemical analyses. The investigation will focus on assessing the influence of waste oil additives on key engineering properties of bitumen, with particular emphasis on penetration, softening point, and viscosity, using quantitative data. In addition, the chemical composition of the waste oils and their interactions within the bituminous matrix will be examined to determine their impact on structural integrity and overall material performance. The experimental results will be systematically compared with findings from the existing literature, providing a critical assessment of the viability of WO modification as a sustainable approach to material enhancement and environmental impact reduction.

2. Material and Method

2.1. Bitumen

In the study, AC 50-70 grade bitumen supplied from the Isparta Municipality Asphalt Plant was used. Standard performance tests were conducted on the bitumen, and the values obtained from these tests are presented in Table 1.

Table 1. Properties of asphalt binder

Properties	Standard	Unit	Results
Penetration (25 °C)	TS EN 1426 [26]	0.1 mm	50.48
Softening Point	TS EN 1427 [27]	°C	47.7
Ductility (5cm/min)	TS EN 13589 [28]	cm	>100
Specific Gravity	TS 15326+A1 [29]	g/cm^3	1.005

2.2. Waste oil

Used industrial or motor oils are oils that have been refined from crude oil and, after a certain period of use in industrial or non-industrial applications, particularly for lubrication purposes, lose their original properties. Used vehicle oils (consisting of gasoline engine, diesel engine, transmission and differential, drivetrain, two-stroke engine, hydraulic brake, antifreeze, grease, and other specialized vehicle oils); industrial oils (including hydraulic systems, turbine and compressor, slideway, open-closed gear, circulation, metal cutting and processing, metal drawing, textile, heat treatment, heat transfer, insulation and protective, rust and corrosion, insulation, transformer, mold, steam cylinder, pneumatic system protective, food and pharmaceutical industry, general-purpose, paper machine, bearing, and other specialized industrial oils and industrial greases); special preparations (thickeners, protectants, cleaners, and similar); and contaminated oil products are included in this definition.

Used industrial or motor oils contain various chemical compounds and additives that degrade over time, leading to a significant environmental burden if improperly disposed of. Due to their complex composition and potential toxicity, the recycling and reutilization of waste oils have gained considerable attention in recent years. In particular, incorporating these waste oils as modifiers or rejuvenators in bituminous materials presents a sustainable approach to both managing waste and enhancing the performance of asphalt binders.

According to 2023 data in Turkey, 318 thousand tons of vehicle oil were supplied to the market, with the potential amount of waste motor and transmission oil estimated to be 190 thousand tons. However, in 2023, only 1,731 tons of waste motor and transmission oil were collected by PETDER (Petroleum Industry and Emobility Association) [30].

Within the scope of the study, the industrial and motor waste oil groups to be used were obtained from Acıöz Petrol Hurdacılık Nakliye Demir Ürünleri San. ve Tic. Ltd. Şti., located in the Selçuklu district of Konya province. The waste oil groups used in the study, namely industrial waste oil and motor waste oil, are shown in Figure 1.

*Corresponding Author: gulsahozkici@gmail.com

ISSN: 2618-6055 / 9, (1), 1 – 13, 2025
DOI:10.62301/usmtd.1704012

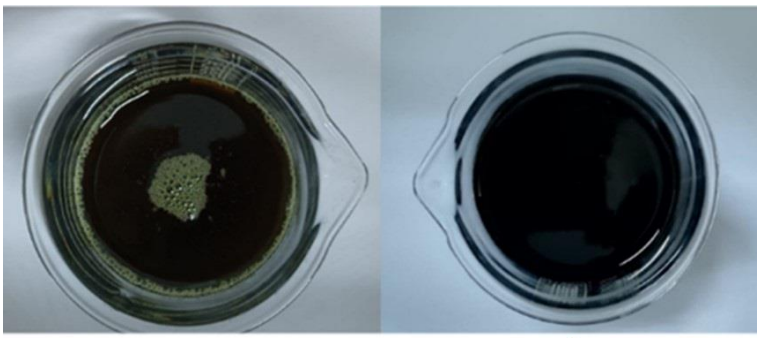


Figure 1. Waste industrial oil (a), Waste engine oil (b)

In the study, the waste industrial and engine oils used were subjected to elemental analysis using the Rotating Disc Electrode Opical Emission Spectrometry (RDE-OES) method (ASTM D6595) (Table 2).

Table 2. The elemental analysis results of the waste industrial and engine oils

	Unit	Calcium (Ca)	Phosphorus (P)	Zinc (Zn)	Ferro (Fe)
Waste Industrial Oil	mg/kg	31,1	181,86	200,6	9,46
Waste Engine Oil		1769,5	847,47	509,79	209,64

2.3. Preparation of Modified Bitumen

For bitumen modification, the bitumen was heated to approximately 145±5 °C, and waste oil from each group was added at weight percentages of 2%, 2.5%, 3%, 3.5%, and 4%. The heated bitumen and waste oil were then transferred to a high-speed mixer mold preheated to 140±5 °C. The high-speed mixer, operating at 1000 rpm, was run for half an hour to ensure a homogeneous mixture of the bitumen and waste oil groups.

Within the scope of the study, the abbreviations for the waste oil groups used in bitumen modification at various weight percentages are provided in Table 3.

Table 3. The abbreviations for the waste oil groups

	Percentages					
	0	%2	%2.5	%3	%3.5	%4
Waste Industrial Oil	B	I2	I2.5	I3	I3.5	I4
Waste Engine Oil		E2	E2.5	E3	E3.5	E4

The effects of changes in the additive and mixture parameters on the properties of bitumen were determined through Penetration (TS EN 1426), Softening Point (TS EN 1427), and Rotational Viscometer (ASTM D 4402) tests [31], and Penetration Index Analysis was conducted. Additionally, Fourier Transform Infrared Spectroscopy (FTIR), Scanning Electron Microscopy (SEM), and Energy Dispersive Spectroscopy (EDS) analyses were performed on the MB, and the chemical structure of the bitumen was compared with that of base bitumen.

3. Results and Discussion

To investigate the effect of using waste engine and industrial oils in bitumen modification on binder consistency, penetration and softening point tests were conducted first. The results of the penetration test are presented in Figure 2. The increase in penetration value with the rise in waste oil percentage indicates that the MB has transitioned to a softer consistency. According to the penetration results, MB containing 2% by weight of waste oil from each oil group remained within the pure bitumen grade (50/70). Additionally, the specimen I2.5 also stayed within the 50/70 penetration grade. The increase in penetration value with the rise in waste oil percentage suggests that the MB has become softer. Based on these findings, it is concluded that the MB obtained through bitumen modification with each waste

*Corresponding Author: gulsahozkici@gmail.com

ISSN: 2618-6055 / 9, (1), 1 – 13, 2025
DOI:10.62301/usmtd.1704012

oil group can be used in cold climate regions. Cold climate bitumens are primarily used in regions exposed to low temperatures, such as highways, roads in mountainous areas, airport runways, and urban roads. These bitumens provide resistance to thermal stresses due to their formulations characterized by high penetration values and low softening points, which enhance their crack resistance. Furthermore, their elastic structure ensures durable performance against frost, icing, and temperature fluctuations.

The physical properties of bitumen intended for use in cold climate conditions play a crucial role in maintaining flexibility at low temperatures and preventing thermal cracking. To this end, bitumens with high penetration values (e.g., 100–150) are typically preferred, as their softness contributes to enhanced elasticity under freezing conditions [32]. Similarly, the softening point of these bitumens is kept relatively low (approximately 35–45 °C), allowing the material to remain workable without becoming brittle [33]. Furthermore, the rotational viscosity at 135 °C is generally limited to below 3000 cP to ensure sufficient workability during production and application phases [34]. Optimization of these parameters enables the material to meet the specific mechanical and environmental demands of cold climate pavements. Recent studies suggest that WO modification may also serve as a promising alternative for enhancing bitumen performance in cold regions. The incorporation of waste lubricating oils has been found to improve low-temperature flexibility and reduce the brittleness of bitumen, offering an environmentally sustainable and economically viable solution for cold climate applications.

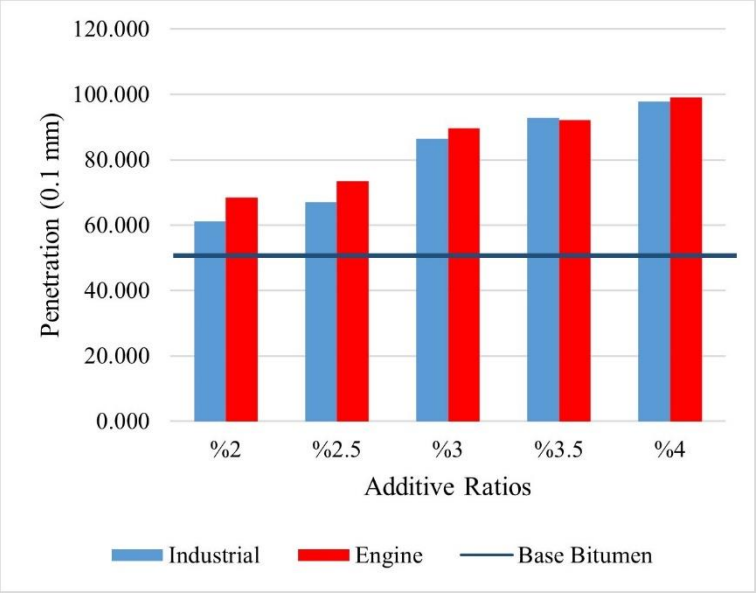


Figure 2. Penetration test results

According to the data obtained from the experimental study, it is observed that bitumen modified with 2.5% by weight of waste oil from each waste oil group has a softening point similar to that of base bitumen (Figure 3). Additionally, in general, the softening points of waste oil-MB are lower compared to base bitumen. As the percentage of waste oil by weight increases, a decrease in the softening point value is observed. The suitability of bituminous binders for specific regions can be inferred based on their softening points. Accordingly, it is concluded that bitumen modifications using waste oil are suitable for use in cold climate regions.

In their study, Gökalp et al. [35] a 2% WO modification resulted in a 12.6% increase in the penetration value compared to base bitumen, while a 4% modification yielded a 33.5% increase. Regarding the softening point, the changes were recorded as a 9.2% decrease for the 2% modification level and an 18.4% decrease for the 4% modification level relative to the base bitumen. In this study, a 2% WO modification resulted in a 45.4% increase in the penetration value compared to base bitumen, while a 4% modification led to a 96.2% increase. As for the softening point, a 2% modification showed a 4.61% increase relative to the base bitumen, whereas a 4% modification resulted in a 1.7% decrease.

*Corresponding Author: gulsahozkici@gmail.com

ISSN: 2618-6055 / 9, (1), 1 – 13, 2025
DOI:10.62301/usmtd.1704012

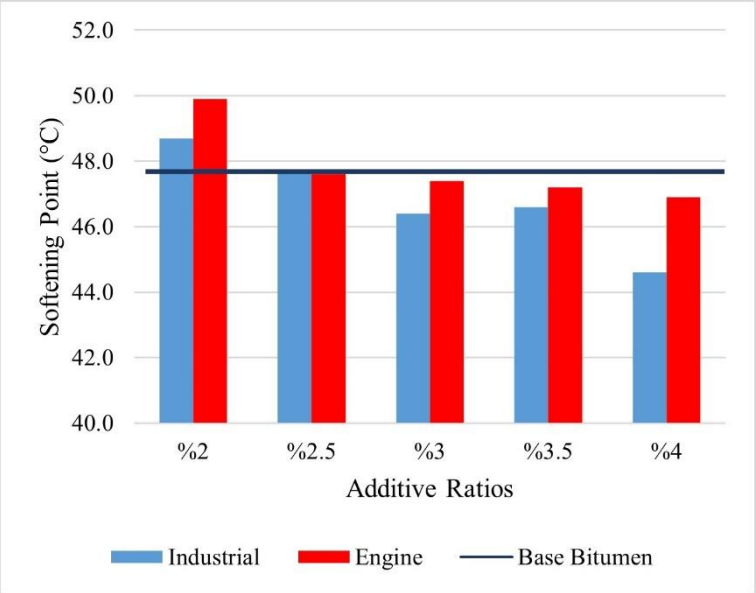


Figure 3. Softening point test results

To determine the temperature sensitivity of bituminous binders, various methods and indices are used. One of these is the penetration index. The calculation of the index utilizes the penetration value and softening point value of the bituminous binder. The penetration index values were calculated using the Equation 1 [36].

$$(20\text{-}PI)/(10+PI) = 50.[\log(800) - \log(Pen) / (TRB-25)] \quad (1)$$

In this equation, Pen 25 ºC represents the penetration value at 25°C, and TRB denotes the softening point. According to the obtained data, the penetration index value for each waste oil group modification falls within the expected range (Figure 4). In particular, it has been observed that the temperature sensitivity of pure bitumen is improved through WO modification.

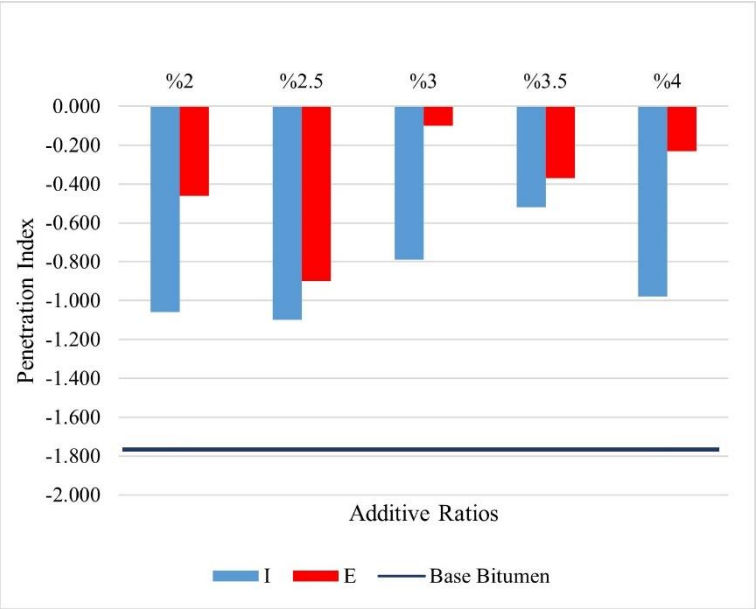


Figure 4. Penetration index results

To determine the mixing and compaction temperatures of the bituminous binders used in the study, a rotational viscometer test was conducted at two different temperatures: 135°C and 165°C. According to the results, the lowest mixing and compaction temperatures were observed at a 4% additive ratio for both waste oil group modifications. The lowest mixing and compaction temperatures were found in the E4 bituminous binder (containing 4% waste engine oil). As shown in Figures 5 and 6, the mixing

*Corresponding Author: gulsahozkici@gmail.com

ISSN: 2618-6055 / 9, (1), 1 – 13, 2025
DOI:10.62301/usmtd.1704012

temperature for pure bitumen is 162.8°C, and the compaction temperature is 154.8°C. As seen in Figure 5, the mixing temperatures for waste industrial oil modification are 161.2°C for I2, 159°C for I2.5, 158.9°C for I3, 158.3°C for I3.5, and 157.6°C for I4. Similarly, the compaction temperatures are 151.9°C for I2, 149.8°C for I2.5, 148.6°C for I3, 147.5°C for I3.5, and 146.8°C for I4.

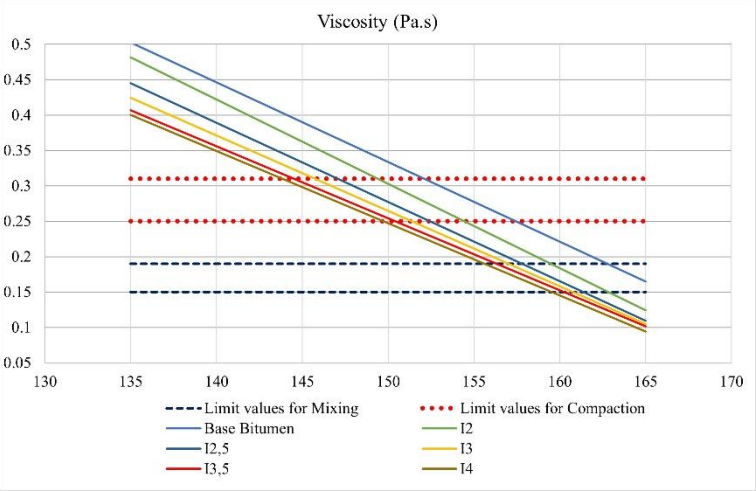


Figure 5. Rotational viscometer values for waste industrial oil modification

As shown in Figure 6, the mixing temperatures for waste engine oil modification are 160.3°C for E2, 158.2°C for E2.5, 157.1°C for E3, 156°C for E3.5, and 155.1°C for E4. The compaction temperatures are 147.8°C for E2, 144.5°C for E2.5, 142.5°C for E3, 141.8°C for E3.5, and 139.4°C for E4. For the highest additive ratio of 4% in waste industrial oil modification (I4), the mixing temperature decreased by approximately 3.5%, and the compaction temperature decreased by approximately 5%. In waste engine oil modification, for the highest additive ratio of 4% (E4), the mixing temperature decreased by 5%, and the compaction temperature decreased by 10%. When evaluating all WO modifications, the lower mixing and compaction temperatures indicate improved workability compared to pure bitumen. The increased workability allows for pavement production and construction activities at lower temperatures, reducing the cost of pavement construction and energy consumption through the modification of bituminous binders with waste oils.

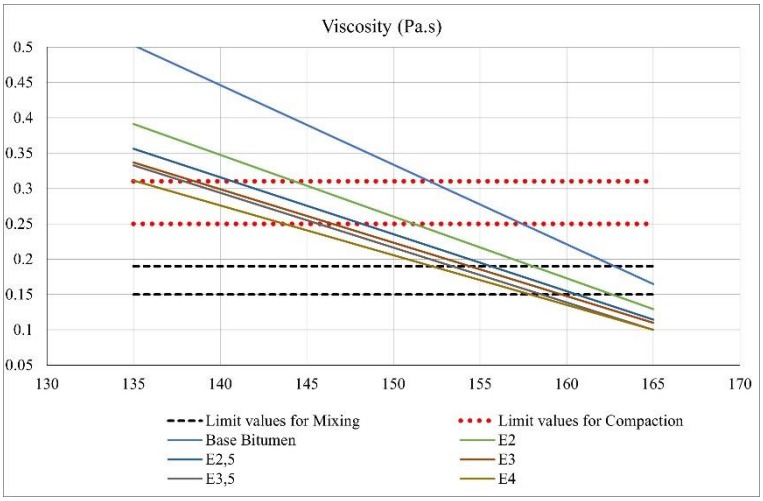


Figure 6. Rotational viscometer values for waste engine oil modification

To determine the chemical structure of the bituminous binders used in the study, FTIR analysis was performed using the JASCO FT/IR-4700 device. The FTIR analysis was conducted in the mid-infrared region (7800-350 cm^{-1}) at room temperature, within a wavenumber range of 400-4000 cm^{-1}, using the FT/IR-4700 device with a resolution higher than 0.4 cm^{-1}. The analysis was completed by measuring the spectrum against percent transmittance and absorbance. FTIR analysis was applied to pure bitumen and the MB samples with the lowest and highest weight percentages for each waste oil group, namely

*Corresponding Author: gulsahozkici@gmail.com

ISSN: 2618-6055 / 9, (1), 1 – 13, 2025
DOI:10.62301/usmtd.1704012

I2, I4, E2, and E4. As shown in Figure 7, the bituminous binders modified with 2% and 4% waste industrial oil exhibit similar and symmetrical transmittance peaks compared to base bitumen.

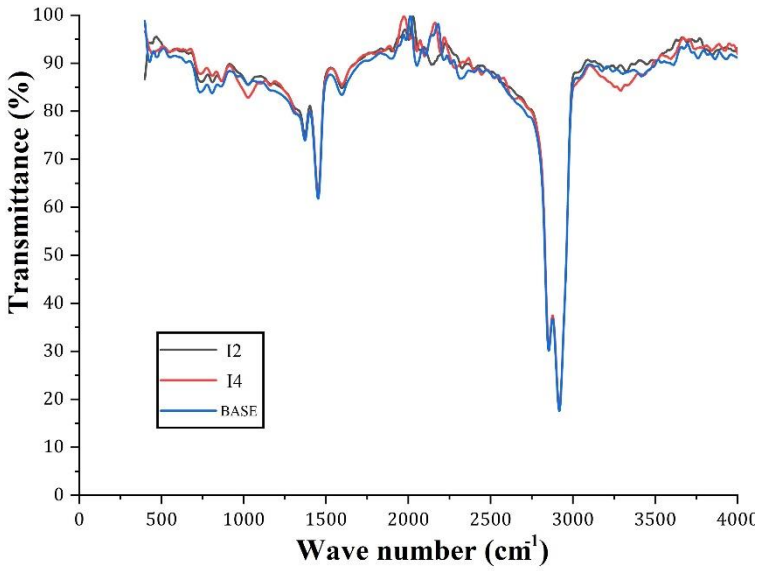


Figure 7. FTIR spectra of waste industrial oil modification

In Figure 8, the bitumen modified with 2% and 4% waste engine oil shows similar and symmetrical characteristics to base bitumen. Based on the evaluation of the FTIR analysis results, it is observed that waste industrial and engine oil modifications are compatible with the bituminous binder. There is no change in the chemical structure of the MB, indicating that the WO modification results from a physical process.

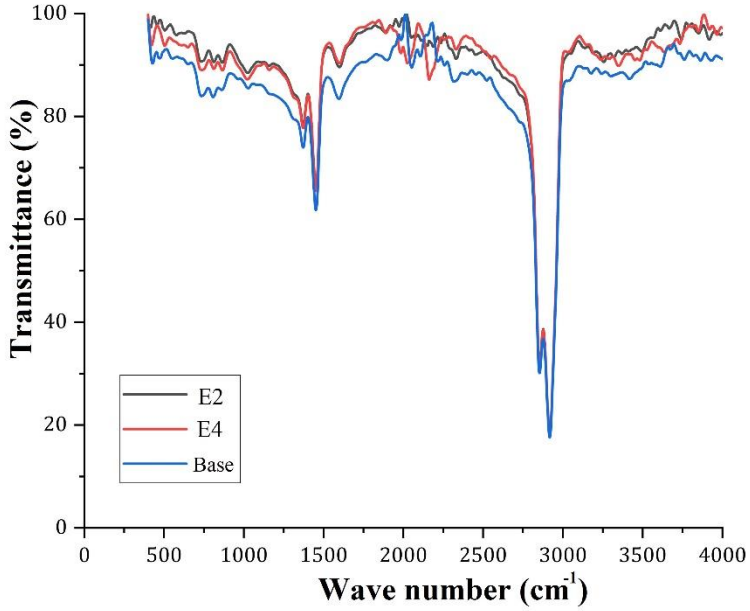


Figure 8. FTIR spectra of waste engine oil modification

To determine the morphological properties of the pure and waste oil-modified bituminous binders used in the study, SEM analysis was performed using the FEI Quanta FEG 250 electron microscope. Additionally, to determine the elemental analysis values of the bituminous binders used in the study, EDS analysis was conducted using the EDAX/EDS detector integrated into the SEM device. The results of the SEM and EDS analyses were presented together on a single figure for each bituminous binder and evaluated collectively. Within the scope of the study, the data obtained from the RDE-OES

*Corresponding Author: gulsahozkici@gmail.com

ISSN: 2618-6055 / 9, (1), 1 – 13, 2025
DOI:10.62301/usmtd.1704012

elemental analysis applied to waste industrial and engine oils align with the results of the EDS analysis. When the obtained data are evaluated, it is observed that the WO modification occurred homogeneously.

According to the SEM image shown in Figure 9, the surface morphology of the base bituminous binder exhibits a uniform and homogeneous structure.

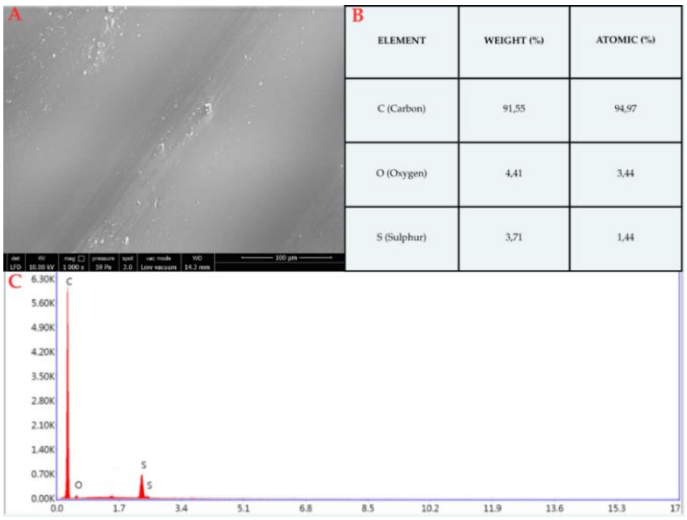


Figure 9. SEM images (A) and EDS analysis results (B and C) of base bitumen

As shown in the SEM image in Figure 10, the surface morphology of the I2 sample exhibits noticeable changes when compared to that of the base bitumen.

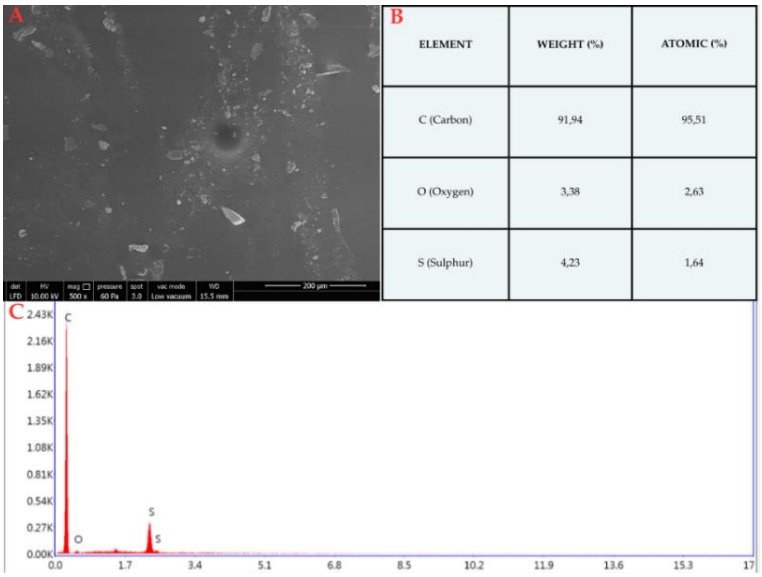


Figure 10. SEM images (A) and EDS analysis results (B and C) of I2

According to the SEM image presented in Figure 11, a distinct pattern formation is observed on the surface morphology of the I4 sample. It has been noted that the pattern becomes more pronounced with the increase in the modification ratio of the waste industrial oil. Furthermore, the waste industrial oil appears to be uniformly dispersed within the bituminous binder without causing any agglomeration. According to the EDS results for sample I4 in Figure 11, different elements that are not present in the structure of base bitumen are observed. The obtained results are in agreement with the findings from the RDE-OES method conducted in this study.

*Corresponding Author: gulsahozkici@gmail.com

ISSN: 2618-6055 / 9, (1), 1 – 13, 2025
DOI:10.62301/usmtd.1704012

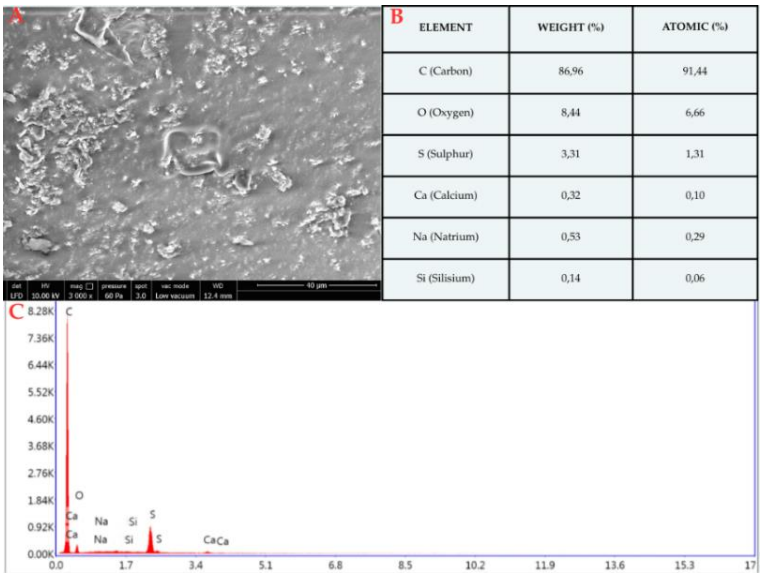


Figure 11. SEM images (A) and EDS analysis results (B and C) of I4

As shown in the SEM image in Figure 12, the surface morphology of the E2 sample exhibits noticeable changes when compared to that of the base bitumen.

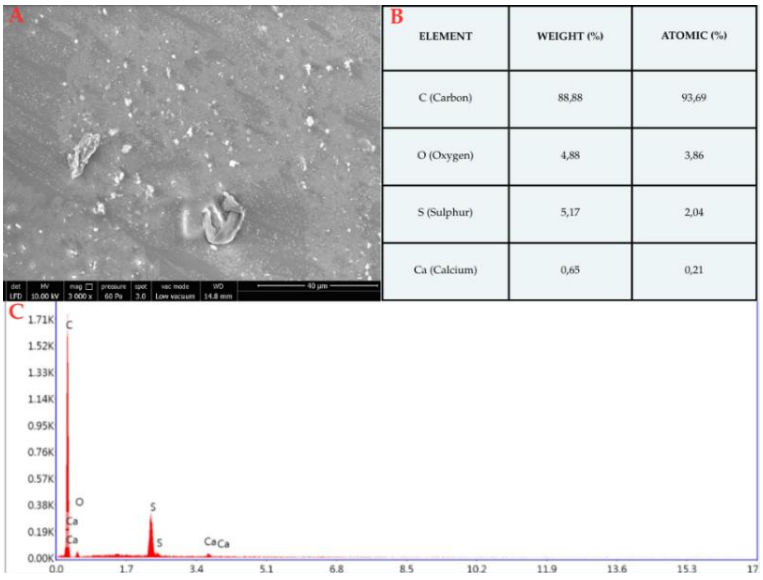


Figure 12. SEM images (A) and EDS analysis results (B and C) of E2

According to the SEM image presented in Figure 13, a distinct pattern formation is observed on the surface morphology of the E4 sample. It has been noted that the pattern becomes more pronounced with the increase in the modification ratio of the waste engine oil. Furthermore, the waste engine oil appears to be uniformly dispersed within the bituminous binder without causing any agglomeration. Furthermore, comparative analysis demonstrated noticeable differences in the surface pattern formations between waste industrial oil and waste engine oil. According to the EDS results for sample E4 in Figure 13, different elements that are not present in the structure of base bitumen are observed. The obtained results are in agreement with the findings from the RDE-OES method conducted in this study.

*Corresponding Author: gulsahozkici@gmail.com

ISSN: 2618-6055 / 9, (1), 1 – 13, 2025
DOI:10.62301/usmtd.1704012

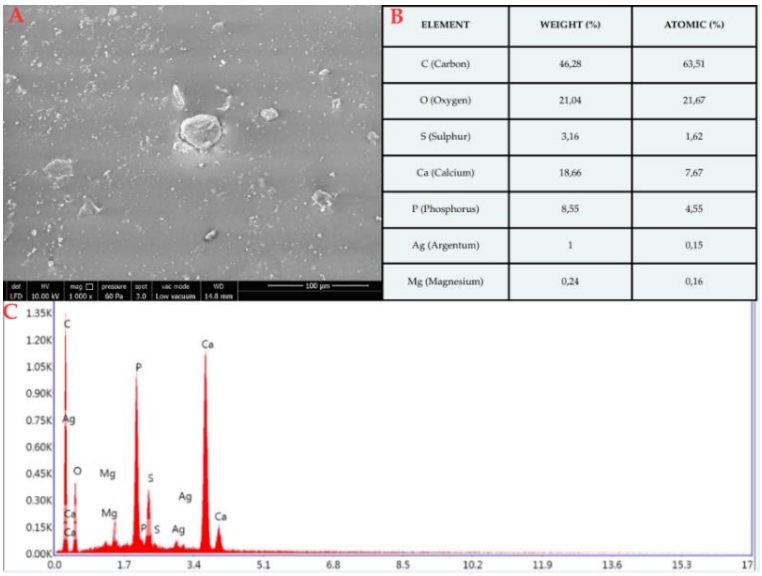


Figure 13. SEM images (A) and EDS analysis results (B and C) of E4

4. Conclusion

The fact that sustainable transportation provides environmental and economic benefits forms the cornerstone of this study. The analyses conducted with this focus aim to contribute to sustainable development goals and guide future transportation strategies.

•It is concluded that the MB obtained through WO modification can be used in cold climate regions.

•In particular, it has been observed that the temperature sensitivity of pure bitumen is improved through WO modification.

•WO modification has increased the workability of the bituminous binder. Thus, working at lower temperatures will reduce pavement construction costs and energy consumption.

•When the elemental analysis results are evaluated, it is observed that waste industrial and engine oil modifications are compatible with the bituminous binder, and the modification process is a result of a physical change.

•The study supports the suitability of using waste industrial and engine oils, which are difficult to store and reuse, in bitumen modification.

In conclusion, the role of waste oils in bitumen modification holds significant potential for environmental sustainability and asphalt performance improvements. The use of waste oils in the asphalt industry not only contributes to solving environmental problems but also enhances the physical and chemical properties of asphalt. Therefore, further research on the effects of waste oils in asphalt modification will increase knowledge in this field and contribute to the development of sustainable practices.

This study has yielded promising results regarding the use of waste oil-MBs, particularly in asphalt mixtures. The findings demonstrate the potential of waste oil additives to enhance the physical properties of bitumen, indicating that such modifications may contribute significantly to sustainable material applications. In light of these results, more comprehensive and long-term studies are planned, particularly focusing on the rejuvenation of aged bitumens using waste oils. Such investigations could provide substantial benefits by both reducing environmental waste and improving the performance of pavement materials.

*Corresponding Author: gulsahozkici@gmail.com

ISSN: 2618-6055 / 9, (1), 1 – 13, 2025
DOI:10.62301/usmtd.1704012

5. Acknowledgements

This study was supported by the Scientific Research Projects Coordination Unit of Süleyman Demirel University under the project number FDK-2022-8810. The first author was also supported by the Council of Higher Education (YÖK) within the scope of the 100/2000 PhD Scholarship Program.

6. Kaynaklar (References)

[1] İ. Gökalp, Y. Özinal, V.E. Uz, Atık bitkisel yemeklik yağların saf bitüm özelliklerine etkisinin araştırılması, Mühendislik Bilimleri ve Tasarım Dergisi 6 (4) (2018) 570–578.

[2] N.H.M. Kamaruddin, M.R. Hainin, N.A. Hassan, M.E. Abdullah, H. Yaacob, Evaluation of pavement mixture incorporating waste oil, Jurnal Teknologi 71 (3) (2014).

[3] İ. Gökalp, V.E. Uz, Sustainable production of aging-resistant bitumen: waste engine oil modification, J. Transp. Eng. Part B: Pavements 147 (4) (2021) 04021055.

[4] T. Khedaywi, M. Melhem, Effect of waste vegetable oil on properties of asphalt cement and asphalt concrete mixtures: an overview, Int. J. Pavement Res. Technol. 17 (1) (2024) 280–290.

[5] M.T. Rahman, A. Mohajerani, F. Giustozzi, Recycling of waste materials for asphalt concrete and bitumen: a review, Materials 13 (7) (2020) 1495.

[6] S. Liu, H. Meng, Y. Xu, S. Zhou, Evaluation of rheological characteristics of asphalt modified with waste engine oil (WEO), Petrol. Sci. Technol. 36 (6) (2018) 475–480.

[7] S. Liu, A. Peng, S. Zhou, J. Wu, W. Xuan, W. Liu, Evaluation of the ageing behaviour of waste engine oil-modified asphalt binders, Constr. Build. Mater. 223 (2019) 394–408.

[8] E. Yalcin, M. Yilmaz, Characteristics of bituminous binders modified with waste oils: physical, chemical, and rheological properties, Int. J. Civ. Eng. 20 (12) (2022) 1377–1395.

[9] T. Shoukat, P.J. Yoo, Rheology of asphalt binder modified with 5W30 viscosity grade waste engine oil, Appl. Sci. 8 (7) (2018) 1194.

[10] W.N.H.M. Sani, R.P. Jaya, B. Bunyamin, Z.H. Al-Saffar, Y. Arbi, Waste motor engine oil–the influence in warm mix asphalt, J. Pendidik. Teknol. Kejuruan 6 (4) (2023) 248–255.

[11] Z. Liu, S. Li, Y. Wang, Characteristics of asphalt modified by waste engine oil/polyphosphoric acid: conventional, high-temperature rheological, and mechanism properties, J. Clean. Prod. 330 (2022) 129844.

[12] Y. Liu, Z. Yang, H. Luo, Effects of anhydrous calcium sulfate whisker and waste engine oil on performance of asphalt binder and its mixture, J. Mater. Civ. Eng. 36 (2) (2024) 04023540.

[13] S. Fernandes, J. Peralta, J.R. Oliveira, R.C. Williams, H.M. Silva, Improving asphalt mixture performance by partially replacing bitumen with waste motor oil and elastomer modifiers, Appl. Sci. 7 (8) (2017) 794.

[14] K.A. Owaid, A.A. Hamdoon, R.R. Matti, M.Y. Saleh, M.A. Abdelzaher, Waste polymer and lubricating oil used as asphalt rheological modifiers, Materials 15 (11) (2022) 3744.

[15] A.A. Mamun, H.I. Al-Abdul Wahhab, M.A. Dalhat, Comparative evaluation of waste cooking oil and waste engine oil rejuvenated asphalt concrete mixtures, Arab. J. Sci. Eng. 45 (2020) 7987–7997.

[16] K.P. Okon, E.O. Inyang, E.O. Mkpa, S.A. Saturday, Optimization of hot mix asphalt using diesel engine waste oil and reclaimed asphalt pavement to improve stability and durability, Asian J. Eng. Appl. Technol. 13 (2) (2024) 7–14.

[17] A. Eltwati, R. Putra Jaya, A. Mohamed, E. Jusli, Z. Al-Saffar, M.R. Hainin, M. Enieb, Effect of warm mix asphalt (WMA) antistripping agent on performance of waste engine oil-rejuvenated asphalt binders and mixtures, Sustainability 15 (4) (2023) 3807.

[18] A. Refaat, A.M. Saleh, R.K. Farag, A.E. Mostafa, Toward sustainable green asphalt pavement mixture using RAP and waste engine oil, Trends Adv. Sci. Technol. 1 (1) (2024) 10.

*Corresponding Author: gulsahozkici@gmail.com

ISSN: 2618-6055 / 9, (1), 1 – 13, 2025
DOI:10.62301/usmtd.1704012

[19] B. Ali, P. Li, D. Khan, M.R.M. Hasan, W.A. Khan, Investigation into the effect of waste engine oil and vegetable oil recycling agents on the performance of laboratory-aged bitumen, Bud. Archit. 23 (1) (2024) 033–054.

[20] A. Villanueva, S. Ho, L. Zanzotto, Asphalt modification with used lubricating oil, Can. J. Civ. Eng. 35 (2) (2008) 148–157.

[21] B. Kuczenski, R. Geyer, T. Zink, A. Henderson, Material flow analysis of lubricating oil use in California, Resour. Conserv. Recycl. 93 (2014) 59–66.

[22] A. Eleyedath, A.K. Swamy, Use of waste engine oil in materials containing asphaltic components, in: Eco-Efficient Pavement Construction Materials, Woodhead Publishing, 2020.

[23] X. Jia, B. Huang, B.F. Bowers, S. Zhao, Infrared spectra and rheological properties of asphalt cement containing waste engine oil residues, Constr. Build. Mater. 50 (2014) 683–691.

[24] A.K. Banerji, D. Chakraborty, A. Mudi, P. Chauhan, Characterization of waste cooking oil and waste engine oil on physical properties of aged bitumen, Mater. Today Proc. 59 (2022) 1694–1699.

[25] S.M.R. Tabatabaei, M. Arabani, G.H. Hamedi, Performance of asphalt mixtures containing high reclaimed asphalt pavement and waste steel rolling oil, Constr. Build. Mater. 441 (2024) 137570.

[26] Türk Standartları Enstitüsü, TS EN 1426. Bitüm ve bitümlü bağlayıcılar – İğne batma derinliği tayini, Ankara, Türkiye, 2015.

[27] Türk Standartları Enstitüsü, TS EN 1427. Bitüm ve bitümlü bağlayıcılar – Yumuşama noktası tayini – Halka ve bilye yöntemi, Ankara, Türkiye, 2015.

[28] Türk Standartları Enstitüsü, TS EN 13589. Bitüm ve bitümlü bağlayıcılar – İşlem görmüş bitümlerin çekme özelliklerinin düktilometre metoduyla tayini, Ankara, Türkiye, 2009.

[29] Türk Standartları Enstitüsü, TS EN 15326+A1. Bitüm ve bitümlü bağlayıcılar-Yoğunluk ve özgül kütle tayini-Kapiler kapaklı piknometre yöntemi. Ankara, Türkiye, 2010.

[30] Petrol Sanayi ve Emobilite Derneği (PETDER), Atık yağların yönetimi projesi 2023 faaliyet raporu. <https://www.petder.org.tr/Uploads/Document/408787db-0002-4186bfbc-cdb96583edbf.pdf>, 2025 (Accessed: Mar. 22, 2025).

[31] ASTM International, ASTM D4402/D4402M-13. Standard test method for viscosity determination of asphalt at elevated temperatures using a rotational viscometer, West Conshohocken, PA, 2013.

[32] Nuroil. (2024). Bitumen 100/150 (Penetration Grade Bitumen) Product Data Sheet. <https://www.nuroil.com/downloads/tds/Bitumen-100-150.pdf> , 2025 (Accessed: Mar. 20, 2025).

[33] Petra Oil. (2024). Bitumen Viscosity Grade AC 2.5. <https://www.petraoilbitumen.com/ac-2-5/> , 2025 (Accessed: Mar. 20, 2025).

[34] Asphalt Institute. (2005). Guidance on AASHTO M320 Specification Limit for Rotational Viscosity. <https://www.asphaltinstitute.org/wp-content/uploads/Guidance-on-AASHTO-M320-Specification-Limit-for-Rotational-Viscosity.pdf> , 2025 (Accessed: Mar. 20, 2025).

[35] İ. Gökalp, A. U. Yıldız, M. B. Eren, V. E. Uz, Atik Motor Yağı Modifiyeli Bitümün Mühendislik Özelliklerinin Araştırılması. Eskişehir Osmangazi Üniversitesi Mühendislik ve Mimarlık Fakültesi Dergisi, 27(3), 184-193, 2019.

[36] N. Kuloğlu, M. Yılmaz, B. V. Kök, Farklı penetrasyon derecelerine sahip asfalt çimentolarının kalıcı deformasyona karşı dayanımlarının ve işlenebilirliklerinin incelenmesi, Uludağ Üniversitesi Mühendislik Fakültesi Dergisi, 13(1), 2008.

***Corresponding Author:** gulsahozkici@gmail.com

ISSN: 2618-6055 / 9, (1), 14 – 24, 2025
DOI:10.62301/usmtd.1691033

Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi
International Journal of Sustainable Engineering and Technology
ISSN: 2618-6055 / Cilt:9, Sayı:1, (2025), 14 – 24
Araştırma Makalesi – Research Article

BLOKZİNCİR DESTEKLİ KİMLİK DOĞRULAMA: MERKEZİ SİSTEMLERE ALTERNATİF BİR YAKLAŞIM

*Didem YANGEÇ[1], Arif KOYUN[1]

[1]Süleyman Demirel Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Bilgisayar Mühendisliği Bölümü, Isparta

(Geliş/Received: 05.05.2025, Kabul/Accepted: 31.05.2025, Yayınlanma/Published: 30.06.2025)

ÖZ

Çağımızın dijital dünyasında güvenliğin temel unsurlarından biri olarak kabul edilen kimlik doğrulama, bir kullanıcının, uygulamanın veya cihazın gerçek kimliğini doğrulamak amacıyla uygulanan bir süreçtir. Bu süreç, yalnızca yetkili bireylerin veya hizmetlerin belirli kaynaklara ve verilere erişmesini sağlamak açısından büyük önem taşımaktadır. Kimlik doğrulama; parolalar (yalnızca kullanıcının bildiği bir şifre), biyometrik veriler (yüz tanıma, parmak izi) veya güvenlik belirteçleri gibi çeşitli yöntemler kullanılarak gerçekleştirilebilir ve böylece sistemlerdeki erişim kontrol mekanizmaları etkin bir şekilde yönetilebilmektedir. Günümüzde kimlik doğrulama süreçleri, güvenlik ve gizlilik açısından büyük önem taşımaktadır. Geleneksel merkezi kimlik doğrulama sistemleri, kullanıcı bilgilerini merkezi bir sunucuda saklayarak erişim yetkilendirme süreçlerini yürütmektedir. Ancak, bu sistemler tek bir güvenlik merkezi içerdiğinden, veri ihlalleri ve siber saldırılara karşı risk altında olabilir. Buna karşılık, blokzincir destekli kimlik doğrulama sistemleri, merkezi bir otoriteye bağlı olmadan kimlik doğrulama sürecini yönetmektedir. Bu sistemler, dağıtık bir ağ üzerinde işlem yaparak, her kullanıcı için şeffaf, değiştirilemez ve izlenebilir bir kayıt defteri oluşturmaktadır. Bu sayede verilerin güvenliğini ve gizliliğini koruyarak işlem yapmaktadır. Böylece hem veri ihlalleri hem de siber saldırılar gibi tehditlere karşı daha dayanıklı bir güvenlik yapısı sunmaktadır. Bu çalışmada, geleneksel kimlik doğrulama sistemleri ile blokzincir destekli çözümler mimari yapı, güvenlik, performans açısından karşılaştırılmıştır. Elde edilen bulgular, her iki yöntemin avantajları ve zorluklarını ortaya koyarak, farklı kullanım senaryoları için en uygun kimlik doğrulama modelini belirlemeye yönelik öneriler sunmaktadır.

Anahtar kelimeler: Geleneksel Kimlik Doğrulama, Blokzincir Teknolojisi, Merkeziyetsizlik, Güvenlik, Performans Karşılaştırması.

BLOCKCHAIN-BASED AUTHENTICATION: AN ALTERNATIVE APPROACH TO CENTRALIZED SYSTEMS

ABSTRACT

Authentication, considered one of the fundamental elements of security in the digital world of our age, is a process applied to verify the true identity of a user, application, or device. This process is of great importance in ensuring that only authorized individuals or services have access to certain resources and data. Authentication can be carried out using various methods such as passwords (a secret known only to the user), biometric data (facial recognition, fingerprints), or security tokens, thus allowing effective management of access control mechanisms in systems. Nowadays, authentication processes are of great importance in terms of security and privacy. Traditional centralized authentication systems manage access authorization processes by storing user information on a central server. However, since these systems contain a single security center, they may be at risk of data breaches and

*Sorumlu Yazar/Corresponding Author: yangecdidem@gmail.com

ISSN: 2618-6055 / 9, (1), 14 – 24, 2025
DOI:10.62301/usmtd.1691033

cyber-attacks. In contrast, blockchain-based authentication systems manage the authentication process without relying on a central authority. These systems create a transparent, immutable, and traceable ledger for each user by operating on a distributed network. In this way, it operates while protecting the security and privacy of the data. Thus, it offers a more resilient security structure against threats such as data breaches and cyber-attacks. In this study, traditional authentication systems and blockchain-based solutions are compared in terms of architecture, security and performance. The findings reveal the advantages and challenges of both methods and provide recommendations for determining the most appropriate authentication model for different usage scenarios.

Keywords: Traditional Authentication, Blockchain Technology, Decentralization, Security, Performance Comparison.

1. Giriş (Introduction)

Kimlik doğrulama, bir kullanıcının iddia ettiği kimliğin gerçekten ona ait olup olmadığını doğrulama sürecidir [1]. Günümüzde bankacılık, sağlık hizmetleri gibi birçok çevrimiçi hizmetlerin yaygınlaşmasıyla birlikte kimlik doğrulamanın önemi giderek artmaktadır. Özellikle hassas verilerin korunması ve yetkisiz erişimlerin önlenmesi için güvenilir kimlik doğrulama yöntemleri kritik bir rol oynamaktadır [2].

Kimlik doğrulama yöntemleri, teknolojinin gelişmesiyle birlikte farklılaşmış ve daha güvenli hale gelmiştir. Bu bağlamda, geleneksel kimlik doğrulama ve blokzincir destekli kimlik doğrulama en çok tartışılan yöntemler arasında yer almaktadır.

Geleneksel kimlik doğrulama yöntemleri, genellikle kullanıcı adı ve şifre kombinasyonuna dayanmaktadır. Bunun yanında, iki faktörlü kimlik doğrulama (2FA) ve çok faktörlü kimlik doğrulama (MFA) gibi ek güvenlik katmanlarıyla da desteklenmektedir [3]. Şifre tabanlı doğrulama, kullanıcıların alışık olduğu basit, entegrasyonu kolay ve ekonomik bir yöntemdir. Fakat şifre çalınması, tahmin edilmesi veya yeniden kullanılması gibi riskleri taşımaktadır [4]. Geleneksel kimlik doğrulama yöntemlerinde kimlik bilgilerinin merkezi sunucularda saklanması, veri ihlalleri durumunda bu bilgilerin toplu halde ele geçirilmesine neden olabilecek ciddi güvenlik açıkları doğurmaktadır [5].

Blokzincir teknolojisi, merkezi olmayan ve değiştirilemez yapısıyla kimlik doğrulamada devrim niteliğindedir. Blokzincir destekli kimlik doğrulama, kimlik bilgilerini dağıtık bir defterde saklayarak güvenliği artırmakta ve veri bütünlüğünü garanti etmektedir. Veriler, merkezi bir sunucuda saklanmadığından tek bir saldırı noktası yoktur. Tüm işlemleri şeffaf ve değiştirilemez şekilde kaydetmektedir. Ancak mevcut sistemlerle entegrasyonunun zor olması ve yüksek maliyet gerektirmesi nedeniyle uygulanması daha karmaşıktır. İşlem doğrulama süresi uzun olması ve ölçeklenebilirlik konusunda iyileştirme gerektirmektedir.

Kimlik doğrulama, bir kullanıcının dijital sistemlere erişimini sağlamak için temel bir güvenlik mekanizmasıdır. Bu alandaki yöntemler, zaman içinde gelişerek çeşitli teknolojik düzeylere evrilmiştir. Geleneksel doğrulama yöntemleri arasında şifreler, akıllı kartlar ve dijital sertifikalar gibi bilgiye dayalı doğrulama teknikleri öne çıkmaktadır. Lal vd. (2016), farklı doğrulama yöntemlerini güvenlik düzeyleri açısından değerlendirmiş ve özellikle şifrelerin yüksek kardinaliteye sahip olması gerektiğini, akıllı kartların ise PIN ile desteklenmesi durumunda daha güvenli hale geldiğini belirtmiştir. Ayrıca, dijital sertifikalar ve çok faktörlü kombinasyonların kullanımıyla güvenlik düzeyinin artırılabileceği vurgulanmıştır [6]. Biyometrik doğrulama ise fiziksel veya davranışsal özelliklere dayalı olarak kullanıcı kimliğini belirlemeyi hedefler. Jain vd. (2004), biyometrik tanımlamanın geleneksel yöntemlere göre daha güvenli bir alternatif sunduğunu ifade etmiştir [7]. Amin vd. (2014) ise mobil cihazlarda kimlik doğrulama konusunu ele alarak geleneksel ve biyometrik yöntemlerin kullanımını değerlendirmiş, davranışsal biyometrinin (yazma ritmi, yürüme tarzı gibi) maliyeti düşürüp kullanılabilirliği artırabileceğini, ancak kişiye özel varyasyonların zorluk oluşturabileceğini belirtmiştir. Bu nedenle çok modlu (multi-modal) sistemlerin geliştirilmesi gerektiği önerilmiştir [8].

Wang vd. (2021), kimlik doğrulama sistemlerine yönelik saldırı türlerini detaylı şekilde inceleyerek, mevcut savunma mekanizmalarının yeterliliğini değerlendirmiştir. Bu çalışmada özellikle parola tabanlı ve biyometrik sistemlerin saldırılara açık olduğu vurgulanmış, savunma mekanizmalarının

*Sorumlu Yazar/Corresponding Author: yangecdidem@gmail.com

ISSN: 2618-6055 / 9, (1), 14 – 24, 2025
DOI:10.62301/usmtd.1691033

geliştirilmesinin gerekliliği ortaya konmuştur. Bu değerlendirme, çalışmamızın ikinci faktör doğrulama bileşenini geliştirirken güvenlik açıklarını dikkate alması açısından önemlidir [9].

Blokzincir teknolojisinin merkeziyetsiz, değiştirilemez ve güvenilir yapısı, kimlik doğrulama alanında yenilikçi çözümler sunmaktadır. Khalid vd. (2020), IoT cihazları için hafif bir blokzincir tabanlı kimlik doğrulama mekanizması önermiş ve mevcut sistemlere göre daha iyi performans sunduğunu göstermiştir [10]. Narayanan vd. (2022) da benzer şekilde IoT ortamında güvenli veri paylaşımı için blokzincir temelli bir model önermiş ve zaman verimliliği, depolama alanı ve işlem hacmi açısından başarılı sonuçlar elde etmiştir [11]. Zhao vd. (2020), dijital eğitim sistemleri bağlamında blokzincir tabanlı iki faktörlü kimlik doğrulama modeli geliştirerek, kimlik verilerinin merkezi olmayan şekilde korunabileceğini göstermiştir [12].

Blokzincirin ikinci faktör doğrulama ile birleştiği diğer bir uygulama alanı da finans ve e-ticaret sistemleridir. Mercan vd. (2021), kredi kartı işlemlerinde SMS doğrulama gibi ikinci faktör yöntemlerinin güvenliğini artırmak için blokzincir temelli bir yapı önermiştir. Bu yapı, hassas doğrulama verilerinin üçüncü taraflarca manipüle edilmesini engellemeyi amaçlamaktadır [13].

Toplumsal katılımın önemli bir alanı olan elektronik oylama sistemlerinde de blokzincir ve çok faktörlü doğrulama kombinasyonları kullanılmaktadır. Abayomi-Zannu vd. (2019), seçmen kimliğinin doğrulanmasında blokzincir destekli çok faktörlü bir sistem önererek, güvenli ve şeffaf mobil oylama çerçevesi sunmuştur [14].

Wanisha vd. (2024), blokzincir temelli çok faktörlü kimlik doğrulama sistemlerinin gizlilik, güvenlik ve kullanılabilirlik açısından avantajlar sağladığını vurgulamıştır. Ancak, bu sistemlerin hala olgunluk aşamasında olduğu ve uygulama karmaşıklığı nedeniyle yaygın kullanım için daha fazla geliştirilmesi gerektiği belirtilmiştir [15].

Literatür incelendiğinde, kimlik doğrulama sistemlerinin güvenliğini artırmak amacıyla çok faktörlü doğrulama yöntemlerinin giderek daha fazla benimsendiği görülmektedir. Aynı zamanda blokzincir teknolojisinin merkeziyetsiz yapı, şeffaflık ve veri bütünlüğü gibi özellikleri sayesinde bu doğrulama süreçlerine entegre edilmeye başlandığı anlaşılmaktadır. Ancak bu çalışmalar genellikle ya sadece merkeziyetsiz yapıyı ele almakta, ya da iki faktörlü doğrulamayı blokzincir dışında ele almaktadır. Bu çerçevede hem geleneksel sistemlerin pratik avantajlarını hem de blokzincir tabanlı yaklaşımların sunduğu güvenlik ve değiştirilemezlik özelliklerini bir araya getiren bütünleşik bir kimlik doğrulama çözümüne duyulan ihtiyaç ortaya çıkmaktadır. Bu çalışmada, merkezi kimlik doğrulama ile blokzincir destekli yapılar birleştirilerek her iki sistemin güçlü yönlerinden faydalanan yenilikçi bir çözüm önerilmektedir. Literatürde bu tür bütünleşik yaklaşımlara az rastlanmakta olup, önerilen sistem bu açıdan özgün bir katkı sunmaktadır.

Bu bağlamda, çalışma kimlik doğrulama sistemlerinin gelişen teknolojiyle birlikte karşılaştığı güvenlik, performans ve kaynak kullanımı gibi kritik konuları ele almakta; geleneksel ve blokzincir tabanlı yaklaşımları bu yönleriyle karşılaştırmalı olarak değerlendirmeyi amaçlamaktadır. Araştırma kapsamında, blokzincir teknolojisinin kimlik doğrulama süreçlerine sağladığı katkılar ile bu sistemlerin performans ve uygulanabilirlik açısından ortaya koyduğu sınırlılıklar detaylı bir biçimde incelenmektedir. Ayrıca, farklı kullanım senaryolarına göre hangi yaklaşımın daha uygun olduğuna dair çıkarımlarda bulunulmaktadır.

2. Materyal ve Metot (Material and Method)

Çalışmanın bu bölümünde, Duende IdentityServer ile geleneksel kimlik doğrulama ve Hyperledger Fabric ile blokzincir destekli kimlik doğrulama için kullanılan bileşenler ve doğrulama süreçlerine ait detaylı bilgiler aşağıdaki alt başlıklarda verilmiştir. Çalışma, iki farklı kimlik doğrulama yönteminin performans karşılaştırması ve analizine dayanmaktadır. Performans ölçümleri için yaygın olarak kullanılan açık kaynaklı bir test aracı olan Apache JMeter kullanılmıştır. JMeter, özellikle web uygulamaları ve hizmetlerinin yük (load) ve stres (stress) testlerini gerçekleştirmek amacıyla kullanılan bir performans test aracıdır. Bu çalışma kapsamında, kimlik doğrulama süreçlerinin farklı kullanıcı yükleri altındaki tepki süreleri ve işlem süreleri JMeter ile senaryolaştırılarak analiz edilmiştir.

*Sorumlu Yazar/Corresponding Author: yangecdidem@gmail.com

ISSN: 2618-6055 / 9, (1), 14 – 24, 2025
DOI:10.62301/usmtd.1691033

2.1. Duende kimlik sunucusu (Duende identityserver)

Çalışmada, geleneksel kimlik doğrulama için Visual Studio platformunda .NET Core 8 kullanılarak geliştirilen Duende IdentityServer ile bir kimlik doğrulama sistemi geliştirilmiştir. Duende IdentityServer, OpenID Connect ve OAuth 2.0 protokollerini destekleyen açık kaynaklı bir kimlik doğrulama ve yetkilendirme sunucusudur. Bu yapı, modern web uygulamaları ve API'ler için merkezi kimlik doğrulama ve yetkilendirme sağlamaktadır [16]. Bu sayede, sistemler arası güvenli kimlik paylaşımı sağlanabilmekte ve kullanıcıların erişim yetkileri merkezi bir yapı üzerinden yönetilebilmektedir. Duende IdentityServer aynı zamanda, yüksek düzeyde esneklik ve özelleştirilebilirlik sunarak farklı güvenlik ihtiyaçlarına uyarlanabilmektedir. Bu esneklik, çalışmada önerilen sistemin hem geleneksel doğrulama mekanizmaları hem de blokzincir tabanlı süreçlerle entegre bir şekilde çalışmasına olanak tanımıştır. Geniş geliştirici topluluğu ve kapsamlı dokümantasyonu sayesinde, geliştirme süreci hızlı ve sorunsuz ilerlemiş, sistemin kurulum ve yönetimi kolaylaştırılmıştır. Ayrıca, erişim denetimi, token yönetimi ve istemci doğrulama gibi güvenlik açısından kritik işlevleri sağlaması nedeniyle, kurumsal düzeyde güvenli bir kimlik doğrulama altyapısı sunma kapasitesi de önemli bir tercih sebebi olmuştur. Duende IdentityServer'ın hem teknik özellikleri hem de .NET Core mimarisiyle olan doğal uyumu sayesinde, çalışmada ihtiyaç duyulan güvenli, esnek ve sürdürülebilir kimlik doğrulama yapısı başarıyla geliştirilmiştir.

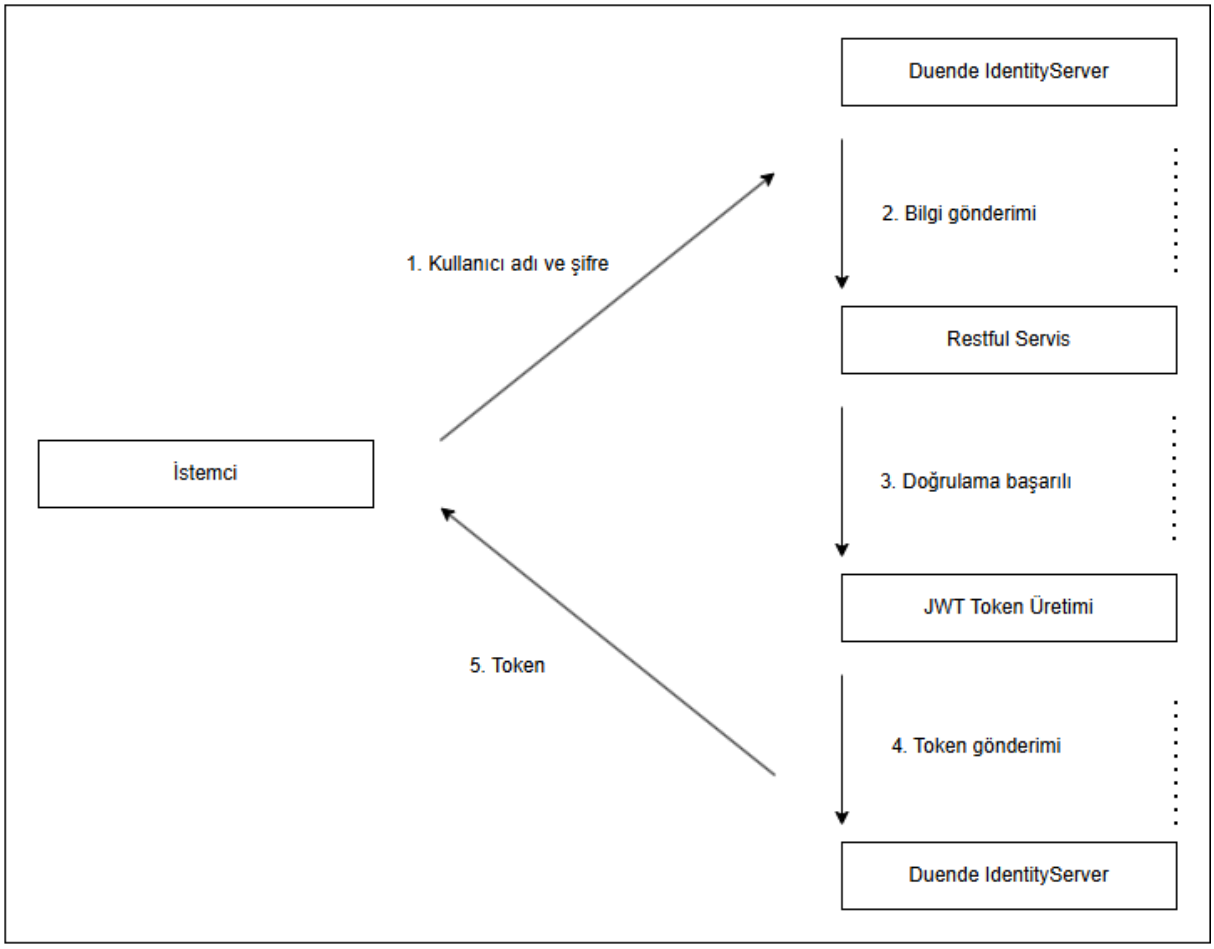
2.1.1. Kimlik doğrulama süreci (Authentication process)

Bu çalışmada, kullanıcı kimlik doğrulaması kullanıcı adı (email) ve şifre bilgileri kullanılarak yapılmıştır. Duende IdentityServer, kullanıcının girdiği bilgileri doğrulamak için bir RESTful servis ile iletişime geçer. Bu doğrulama süreci şu adımlardan oluşmaktadır:
Kullanıcı Girişi ve Bilgi Toplama: Kullanıcı, giriş sayfasında email ve şifre bilgilerini girer. Bu bilgiler, istemci (Client) tarafından Duende IdentityServer'a iletilir.

- RESTful Servis ile Doğrulama: Duende IdentityServer, kullanıcının girdiği bilgileri doğrulamak için bir RESTful servis ile iletişim kurmaktadır. Bu doğrulama işlemi için serviceToken, projectName, email ve password parametrelerini kullanmaktadır. ServiceToken, kimlik doğrulama işleminin güvenliğini sağlamak için kullanılan bir güvenlik anahtarıdır. ProjectName hangi proje için doğrulama yapıldığını belirtmektedir. Email, kullanıcının kimliğini doğrulamak için kullanılmaktadır. Password, kullanıcının kimliğini doğrulamak için gerekli olan şifre bilgisidir.
- Kimlik Doğrulama ve JWT Üretimi: RESTful servis, gönderilen parametreleri kontrol eder ve bilgilerin doğruluğunu teyit eder. Doğrulama başarılı olursa, Duende IdentityServer tarafından JWT (JSON Web Token) üretilmektedir. Bu token, kullanıcının kimliğini doğrulamak ve yetkilendirilmiş kaynaklara erişimini sağlamak için kullanılmaktadır. JWT içerisinde, kullanıcının kimlik bilgileri, yetki düzeyleri ve token süresi gibi bilgiler yer almaktadır.
- Token İletimi ve Yetkilendirme: Üretilen JWT (JSON Web Tokens), istemciye geri gönderilir. İstemci, bu token'ı her istek yaptığında HTTP Header içerisinde ileterek yetkilendirme işlemini gerçekleştirir. Duende IdentityServer, token'ı doğrulayarak kullanıcının kimliğini ve yetkisini teyit eder. Bu sayede, kullanıcı yalnızca yetkili olduğu kaynaklara erişebilir.
  Geleneksel kimlik doğrulama süreci Şekil 1'de gösterilmiştir.

*Sorumlu Yazar/Corresponding Author: yangecdidem@gmail.com

ISSN: 2618-6055 / 9, (1), 14 – 24, 2025
DOI:10.62301/usmtd.1691033



Şekil 1. Geleneksel kimlik doğrulama süreci (Traditional authentication process)

2.2. Blokzincir kimlik doğrulama süreci (Blockchain authentication process)

Blokzincir destekli kimlik doğrulama sistemi, merkezi olmayan ve güvenli bir yapı sunarak, kullanıcıların kimliklerini doğrulama sürecini daha güvenilir hale getirmektedir. Bu yapı, her işlemi dağıtık bir ağda kayıt altına alarak, verilerin değiştirilmesini veya sahtecilik yapılmasını imkânsız hale getirir [17]. Blokzincirinin sunduğu şeffaflık ve güvenlik özellikleri sayesinde, kimlik doğrulama işlemleri daha güvenli ve izlenebilir hale gelir [18].

Bu çalışmada, blokzincir destekli kimlik doğrulama sistemi, Hyperledger Fabric kullanılarak geliştirilmiştir. Bunun için sanal makine kurularak Linux tabanlı bir ortamda çalıştırılmıştır. Çünkü, Hyperledger Fabric, Linux Foundation tarafından geliştirilmiştir ve en güncel destek ile dokümantasyonun bu platformda bulunmasını sağlamaktadır [19].

Hyperledger Fabric, akıllı sözleşme (chaincode) geliştirme sürecinde Go, Java ve JavaScript gibi çeşitli programlama dillerini destekleyerek geliştiricilere esneklik sunmaktadır. Ancak, Hyperledger Fabric'in varsayılan programlama dili Go olduğu için, bu çalışmada kimlik doğrulama sistemi Go dilinde geliştirilmiştir. Bu sayede, sistemin güvenliği ve işleyişi Go'nun verimli ve güçlü yapısından faydalanarak sağlanmıştır.

Bunun yanı sıra, Hyperledger Fabric'in tercih edilme sebeplerinden biri de izinli (permissioned) bir blokzincir olmasıdır. Geleneksel açık blokzincir sistemlerinden farklı olarak, Hyperledger Fabric'de yalnızca yetkilendirilmiş kullanıcılar ağa katılabilir ve işlem gerçekleştirebilir. Bu özellik, sistemin güvenliğini artırarak yalnızca doğrulanmış kullanıcıların erişimine izin verir. Ayrıca, Fabric madencilik (mining) gerektirmediği için işlem onay süreleri çok daha hızlıdır. Geleneksel blokzincirlerde

*Sorumlu Yazar/Corresponding Author: yangecdidem@gmail.com

ISSN: 2618-6055 / 9, (1), 14 – 24, 2025
DOI:10.62301/usmtd.1691033

madencilik işlemi büyük hesaplama gücü gerektirirken, Hyperledger Fabric'in farklı konsensüs mekanizmaları sayesinde işlemler düşük gecikme süresiyle tamamlanır. Bu durum, özellikle kimlik doğrulama gibi anlık ve güvenilir işlem gerektiren sistemler için büyük bir avantaj sağlamaktadır.

Blokzincir destekli kimlik doğrulama süreci, merkezi bir otoriteye ihtiyaç duymadan kullanıcıların güvenli bir şekilde kimliklerini doğrulamalarını sağlamaktadır. Bu süreçte kullanılan en önemli güvenlik özelliklerinden biri, zaman tabanlı tek kullanımlık şifreler (TOTP-Time-Based One-Time Password) kullanımıdır. TOTP, özellikle iki faktörlü kimlik doğrulama (2FA) mekanizmalarında yaygın olarak kullanılan bir güvenlik aracıdır. TOTP, her kullanıcı için belirli bir zaman diliminde geçerli olan ve yalnızca bir kez kullanılabilen dinamik şifreler üretir. Bu dinamik şifreler, kimlik doğrulama sürecine ekstra güvenlik katmanı ekler ve şifrelerin sadece belirli bir süre için geçerli olmasını sağlar, bu da saldırganların şifreyi ele geçirme olasılığını azaltır. Blokzincir destekli kimlik doğrulama süreci şu adımlarla gerçekleştirilir:

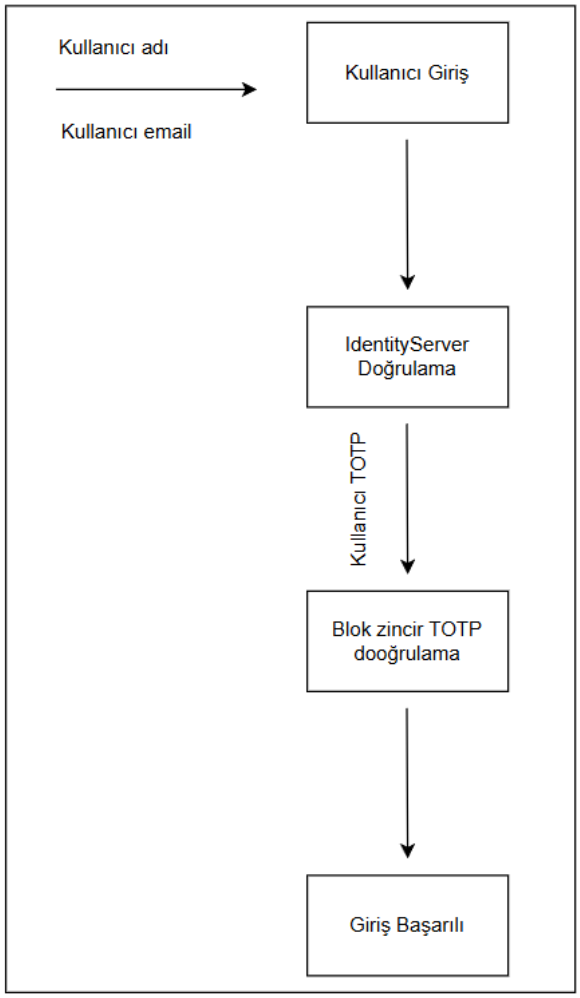
• Kullanıcı kayıt işlemi: Kullanıcı kayıt işlemi sırasında, kullanıcıya ait email adresi ile benzersiz bir TOTP gizli anahtarı (secret key) oluşturulur. Bu bilgiler, Hyperledger Fabric üzerinde çalışan bir chaincode (akıllı sözleşme) aracılığıyla blokzincirine kaydedilir. Kullanıcı bilgileri, anahtar-değer (key-value) formatında depolanarak hem verimlilik hem de güvenlik sağlanır. TOTP gizli anahtarı, her kullanıcı için özel olarak oluşturulur ve böylece kimlik doğrulama süreci her kullanıcı için benzersiz hale gelir.

• Kullanıcı doğrulama: Kullanıcının kimlik bilgileri (email ve şifre) IdentityServer üzerinden doğrulandıktan sonra, kullanıcıdan Google Authenticator gibi bir uygulama aracılığıyla oluşturulmuş olan 6 haneli TOTP kodunu sisteme girmesi istenir. Sistem, blokzincirinde kayıtlı olan kullanıcının TOTP gizli anahtarını kullanarak aynı kodu hesaplar. Hesaplanan TOTP kodu ile kullanıcının girdiği kod eşleşirse, kimlik doğrulama başarılı olur. Eşleşme durumunda, kullanıcı sisteme giriş yapabilir. Eğer kodlar eşleşmezse, giriş işlemi reddedilir.

Blokzincir destekli kimlik doğrulama süreci Şekil 2'de gösterilmiştir.

*Sorumlu Yazar/Corresponding Author: yangecdidem@gmail.com

ISSN: 2618-6055 / 9, (1), 14 – 24, 2025
DOI:10.62301/usmtd.1691033



Şekil 2. Blokzincir destekli kimlik doğrulama süreci (Blockchain-based authentication process)

Bu süreç, blokzinciri teknolojisinin sağladığı dağıtık yapı sayesinde merkezi bir otoriteye ihtiyaç duymadan güvenli ve bağımsız bir kimlik doğrulama mekanizması sunar.

3. Araştırma Bulguları (Research Findings)

Geleneksel kimlik doğrulama sistemi ile blokzincir destekli kimlik doğrulama sisteminin performansları, farklı kullanıcı yükleri altında karşılaştırmalı olarak değerlendirilmiştir. Karşılaştırma, farklı kullanıcı yükleri altında sistemlerin yanıt süresi, CPU ve bellek (RAM) kullanımı gibi metrikler üzerinden gerçekleştirilmiştir. Performans testleri, 1, 10 ve 100 kullanıcı senaryoları için uygulanmıştır. Geliştirilen test uygulamaları, Intel(R) Core(TM) i5-10210U CPU @ 1.60GHz, 16 GB RAM, 64-bit işletim sistemi ve x64 tabanlı 4 çekirdekli bir donanım üzerinde çalıştırılmıştır. Geleneksel kimlik doğrulama sistemi, Duende IdentityServer altyapısı kullanılarak test edilmiştir. Testlerde performans ölçümleri için Apache JMeter kullanılmıştır. Geleneksel kimlik doğrulama sistemi ile blokzincir destekli doğrulama sistemine ait test senaryoları üç aşamada yapılandırılmıştır. Senaryolarda kullanılan temel parametreler şu şekildedir:

- **Kullanıcı sayısı (Threads)**: Aynı anda sisteme erişim sağlayacak kullanıcıların sayısını (clients) ifade etmektedir. Bu çalışma kapsamında sırasıyla 1, 10 ve 100 kullanıcı senaryoları uygulanmıştır.
- **Ramp-Up süresi**: Belirtilen sayıdaki kullanıcının sisteme ne kadar süre içinde dahil olacağını tanımlamaktadır. Tüm senaryolarda bu süre 10 saniye olarak belirlenmiştir. Bu sayede, kullanıcılar aynı anda değil, kısa aralıklarla sisteme giriş yaparak daha gerçekçi bir yük simülasyonu sağlanmıştır.
- **Döngü sayısı (Loop Count)**: Her bir kullanıcının aynı işlemi kaç kez tekrar edeceğini göstermektedir. Bu sayı, kullanıcı başına gönderilecek istek sayısını belirlemektedir.

*Sorumlu Yazar/Corresponding Author: yangecdidem@gmail.com

ISSN: 2618-6055 / 9, (1), 14 – 24, 2025
DOI:10.62301/usmtd.1691033

Uygulanan senaryolar şu şekilde özetlenebilir:

- İlk senaryoda, 1 kullanıcı 10 saniyelik ramp-up süresiyle teste başlamakta ve 100 kez işlem tekrarlayarak toplam 100 istek üretilmiştir.
- İkinci senaryoda, 10 kullanıcı aynı ramp-up süresi içinde sisteme dahil olmakta ve her biri 100 kez işlem yaparak toplamda 1000 istek gerçekleştirmiştir.
- Üçüncü ve en yoğun senaryoda ise 100 kullanıcı, yine 10 saniyelik ramp-up süresiyle sisteme giriş yapmakta ve her biri 50 kez işlem yaparak toplam 5000 istek üretmiştir.

Bu yapılandırma sayesinde, sistemin farklı seviyelerdeki eşzamanlı kullanıcı yüklerine verdiği yanıt süreleri ölçülmüş ve performans analizleri yapılmıştır.

Her test senaryosu yanıt süreleri milisaniye (ms) cinsinden kaydedilmiştir. Elde edilen sonuçlar kullanıcı sayısı, kullanıcı başına istek sayısı, toplam istek sayısı, ortalama yanıt süresi ve standart sapmalar ile birlikte değerlendirilmiştir. Tablo 1 ve Tablo 2'de yer alan performans ölçüm sonuçları, geleneksel kimlik doğrulama sistemi ile blokzincir destekli kimlik doğrulama sisteminin farklı kullanıcı yükleri altındaki tepkilerini ortaya koymaktadır. Her iki sistem için 1, 10 ve 100 kullanıcı senaryolarında ölçülen yanıt süreleri, sistemlerin ölçeklenebilirlik düzeylerini ve kaynak kullanımına verdikleri tepkileri değerlendirmek açısından önemli veriler sunmaktadır.

Duende IdentityServer altyapısı kullanılarak gerçekleştirilen geleneksel kimlik doğrulama testlerinde, tek kullanıcı ve 100 istek senaryosu altında ortalama yanıt süresi 7 ms olarak ölçülmüştür. 10 kullanıcı ve her biri için 100 istek olmak üzere toplam 1000 istek içeren senaryoda ortalama yanıt süresi hafif artış göstererek 9 ms'ye ulaşmıştır. 100 kullanıcı ve kullanıcı başına 50 istek olmak üzere toplam 5000 isteğin gerçekleştirildiği en yoğun senaryoda ise ortalama yanıt süresi 139 ms'ye yükselmiştir. Sistemin yanıt süresinde görülen bu artış, eşzamanlı kullanıcı sayısı ve toplam istek sayısının artışına bağlı olarak beklenen bir performans yavaşlamasıdır. Ancak, bu artışa rağmen 100 kullanıcı ve 5000 toplam istek altında dahi sistemin performansının hala kabul edilebilir seviyelerde olduğu söylenebilir. Standart sapma değerlerinin düşük ve nispeten sabit kalması (1.76 ms'den 134.39 ms'ye) ise, sistem yanıt sürelerinin tutarlı ve öngörülebilir olduğunu göstermektedir.

Hyperledger Fabric tabanlı blokzincir destekli kimlik doğrulama sisteminde ölçülen yanıt süreleri geleneksel sisteme kıyasla daha yüksektir. Tek kullanıcı ve kullanıcı başına 100 istek olmak üzere toplam 100 istek yapılan senaryoda ortalama yanıt süresi 8 ms ile başlamıştır. 10 kullanıcı ve kullanıcı başına 100 istek ile toplam 1000 isteğin gerçekleştirildiği senaryoda ortalama yanıt süresi 11 ms'ye yükselmiştir. En yoğun senaryo olan 100 kullanıcı ve kullanıcı başına 50 istek olmak üzere toplam 5000 istekte ise ortalama yanıt süresi 334 ms'ye çıkmıştır. Standart sapma değerlerinin geleneksel sisteme göre oldukça yüksek olması (7.02 ms'den 267.38 ms'ye) blokzincir sisteminde yanıt sürelerinde daha fazla dalgalanma yaşandığını göstermektedir. Bu durum, blokzincir mimarisinde bulunan işlem doğrulama, konsensüs mekanizmaları ve blok yazma süreçlerinin doğal bir sonucu olarak değerlendirilebilir. Blokzincir sisteminin yüksek eşzamanlılık altında yanıt sürelerinde daha fazla değişkenlik göstermesi, ölçeklenebilirlik ve performans açısından önemli bir sınırlama oluşturmaktadır. Geleneksel kimlik doğrulama sistemi, merkezi mimarisi sayesinde hem düşük gecikme süreleri hem de tutarlı performans göstermektedir. Kullanıcı sayısı ve istek yükü arttıkça yanıt süresi artsa da sistem 100 kullanıcı ve 5000 toplam istek altında ortalama 139 ms yanıt süresi ile kabul edilebilir bir performans sergilemektedir. Standart sapmanın nispeten düşük olması, sistem yanıt sürelerinin öngörülebilir ve stabil olduğunu göstermektedir. Buna karşın, blokzincir destekli kimlik doğrulama sistemi, güvenlik ve veri bütünlüğü avantajları sunmasına rağmen, işlem doğrulama ve konsensüs süreçlerinden dolayı daha yüksek ve dalgalı yanıt sürelerine sahiptir. 100 kullanıcı ve 5000 toplam istek senaryosunda ortalama yanıt süresi 334 ms'ye yükselirken, standart sapma değeri 267 ms gibi yüksek bir seviyede seyretmektedir. Bu yüksek standart sapma ve yanıt süresi dalgalanmaları, blokzincir altyapısının yoğun kullanıcı yüklerinde performans açısından daha değişken ve öngörülemez olduğuna işaret etmektedir. Özetle, geleneksel kimlik doğrulama sistemi, performans ve kararlılık açısından yüksek kullanıcı yüklerinde daha uygun bir çözüm sunarken, blokzincir destekli sistem daha güçlü güvenlik özellikleri sağlamasına rağmen performans açısından belirgin sınırlamalara sahiptir.

*Sorumlu Yazar/Corresponding Author: yangecdidem@gmail.com

ISSN: 2618-6055 / 9, (1), 14 – 24, 2025
DOI:10.62301/usmtd.1691033

Tablo 1. Geleneksel kimlik doğrulama performans ölçüm sonuçları (Traditional authentication performance measurement results)

Kullanıcı Sayısı	Kullanıcı Başına İstek Sayısı	Toplam İstek Sayısı	Ortalama Yanıt Süresi (ms)	Standart Sapma
1	100	100	7	1.76
10	100	1000	9	3.23
100	50	5000	139	134.39

Tablo 2. Blokzincir destekli kimlik doğrulama performans ölçüm sonuçları (Blockchain-based authentication performance measurement results)

Kullanıcı Sayısı	Kullanıcı Başına İstek Sayısı	Toplam İstek Sayısı	Ortalama Yanıt Süresi (ms)	Standart Sapma
1	100	100	8	7.02
10	100	1000	11	9.28
100	50	5000	334	267.38

Tablo 3. Geleneksel kimlik doğrulama ile blokzincir tabanlı kimlik doğrulaması karşılaştırılma kriterleri (Comparison criteria between traditional authentication and blockchain-based authentication)

Metrik	Geleneksel Kimlik Doğrulama (Duende IdentityServer)	Blokzincir Destekli Kimlik Doğrulama (Hyperledger Fabric)
Merkezileşme	Tek bir otorite yönetir.	Dağıtık yapı, merkezi otorite yok
Güvenlik	Orta	Yüksek
Erişim Kontrolü	Yetkilendirme merkezi sunucuya bağlı	Kullanıcılar kendi verilerini yönetebilir
Gizlilik	Merkezi otorite tüm verileri görebilir.	Kullanıcılar verileri üzerinde tam kontrole sahiptir.
Ölçeklenebilirlik	Yüksek	Orta
Erişilebilirlik	Sunucu çökmesi durumunda kesinti olur	Tüm düğümler çalıştığı sürece kesintisiz

Tablo 3'te sunulan karşılaştırma, geleneksel kimlik doğrulama sistemi ile blokzincir tabanlı kimlik doğrulama sisteminin yapısal ve işlevsel farklılıklarını ortaya koymaktadır. Geleneksel kimlik doğrulama sistemleri, merkezi bir otorite tarafından yönetilir ve tüm kimlik doğrulama süreçleri bu otorite üzerinden yürütülür. Bu yapı, hızlı yanıt süreleri ve yüksek ölçeklenebilirlik gibi avantajlar sunsa da güvenlik ve gizlilik açısından bazı sınırlamalar barındırmaktadır. Özellikle merkezi sunucunun çökmesi durumunda sistemin tamamen devre dışı kalması erişilebilirlik açısından zafiyet oluşturabilir. Ayrıca, kullanıcı verilerinin merkezi bir noktada toplanması, gizlilik ve veri ihlali risklerini artırmaktadır.

Öte yandan, blokzincir destekli kimlik doğrulama, verilerin merkezi bir otoriteye bağlı olmadan, dağıtık bir yapı üzerinde saklanması sayesinde yüksek güvenlik ve gizlilik sunmaktadır [20]. Bu sistemlerde güvenlik, temel olarak blokzincirin değiştirilemezlik (immutability), kriptografik imzalar ve işlem geçmişinin şeffaflığına dayanmaktadır. Örneğin, her kimlik doğrulama işlemi kriptografik olarak imzalanır ve blokzincire kaydedildiğinde sonradan değiştirilmesi mümkün değildir. Bu sayede kimlik sahteciliği, veri manipülasyonu ve yetkisiz erişim gibi tehditlere karşı direnç sağlamaktadır. Ayrıca, her kullanıcı verisinin yalnızca gerekli olduğunda ve kullanıcı izniyle paylaşılması, veri gizliliğini arttırmaktadır. Sistem, Sybil saldırılarına karşı düğüm doğrulama mekanizmaları (örneğin Proof-of-Authority veya Proof-of-Stake) ile korunabilir. Buna ek olarak, veri bütünlüğü ve erişim kontrolü gibi güvenlik kriterleri, merkezi sistemlere göre daha güçlü şekilde uygulanabilmektedir. Örneğin, blokzincire entegre edilen akıllı sözleşmeler aracılığıyla sadece yetkili tarafların belirli verilere erişmesine izin verilebilir. Bu tür güvenlik özelliklerinin değerlendirilmesi, sistemin dışarıdan saldırılara karşı direnci, veri sızıntısı ihtimali ve işlem geçmişinin şeffaflığı gibi kriterler göz önünde bulundurularak yapılmıştır. Bu yönleriyle, blokzincir destekli kimlik doğrulama sistemleri, geleneksel sistemlere kıyasla daha yüksek düzeyde güvenlik sağlamaktadır. Ancak, bu yüksek güvenlik düzeyinin

*Sorumlu Yazar/Corresponding Author: yangecdidem@gmail.com

ISSN: 2618-6055 / 9, (1), 14 – 24, 2025
DOI:10.62301/usmtd.1691033

bir maliyeti vardır. Dağıtık yapının doğası gereği, her işlem ağda çok sayıda düğüm tarafından doğrulanmak zorunda olduğu için, işlem yükü artmakta ve yanıt süreleri uzamaktadır. Gerçekleştirilen testlerde, blokzincir destekli sistemlerin geleneksel sistemlere kıyasla daha fazla CPU ve RAM kaynağı tükettiği, dolayısıyla daha yüksek donanımsal gereksinimlere ihtiyaç duyduğu gözlemlenmiştir. Bu nedenle, blokzincir tabanlı kimlik doğrulama sistemleri güvenlik açısından avantajlı olmakla birlikte, performans ve kaynak kullanımı açısından bazı sınırlamalar barındırmaktadır.

4. Tartışma ve Sonuç (Results and Discussion)

Bu çalışmada, geleneksel kimlik doğrulama ile blokzincir destekli kimlik doğrulamanın performans ve güvenlik açısından karşılaştırılması amaçlanmıştır. Yapılan performans testleri, her iki yöntem arasındaki farkları net bir şekilde ortaya koymuştur. Geleneksel kimlik doğrulama, merkezi sunucu yapısı sayesinde daha hızlı yanıt süreleri ve yüksek ölçeklenebilirlik sunmaktadır. Bu, verilerin merkezi bir noktada saklanması ve işlemlerin daha verimli bir şekilde gerçekleştirilmesinden kaynaklanmaktadır. Ancak, merkezi yapının getirdiği güvenlik riskleri, kullanıcı verilerinin gizliliğini tehdit edebilir ve merkezi otoritenin güvenilirliğine bağımlılığı artırabilir. Özellikle sunucu çökmesi veya saldırılar durumunda, sistemin tamamının erişilemez hale gelmesi ciddi bir sorun teşkil etmektedir

Diğer taraftan, blokzincir destekli kimlik doğrulama, dağıtık yapısı sayesinde daha yüksek güvenlik ve gizlilik sağlar. Kullanıcılar, verileri üzerinde tam kontrol sahibidir ve veriler yalnızca ihtiyaç duyulduğunda paylaşılmaktadır. Bu yapı, merkezi otoriteye dayalı sistemlerden bağımsızlık sunar. Ancak, bu avantajların maliyeti, daha uzun yanıt süreleri ve yüksek işlem yükü olarak karşımıza çıkmaktadır. Blokzincir destekli sistemler, ağda her kullanıcı işlemi için doğrulama gerçekleştirdiğinden, ölçeklenebilirlik açısından geleneksel yöntemlere göre sınırlamalar göstermektedir. Ayrıca, daha fazla CPU ve RAM kaynağı gerektiren bu sistemler, performans açısından daha fazla işlem gücü ve zaman tüketmektedir. Sonuç olarak, her iki kimlik doğrulama yönteminin kendine özgü güçlü ve zayıf yönleri vardır. Geleneksel kimlik doğrulama, hızlı ve ölçeklenebilir yapısıyla yüksek performans sunarken, güvenlik ve gizlilik açısından zayıf kalmaktadır. Blokzincir destekli kimlik doğrulama ise daha güvenli ve gizlilik odaklı bir çözüm sunmakta, ancak yüksek işlem gücü ve performans gereksinimleri nedeniyle daha fazla kaynak kullanmaktadır. Bu çalışmanın bulguları, geleneksel kimlik doğrulamanın hız ve ölçeklenebilirlik gerektiren senaryolarda, blokzincir destekli kimlik doğrulamanın ise güvenlik ve gizlilik öncelikli uygulamalarda daha uygun olduğunu göstermektedir.

Gelecekte, blokzincir destekli sistemlerin performansını artırmak ve ölçeklenebilirlik sorunlarını çözmek adına çeşitli iyileştirme çalışmaları yapılabilir. Örneğin, işlem yükünü azaltan ve doğrulama süresini kısaltan Layer-2 çözümleri kimlik doğrulama süreçlerine entegre edilebilir. Ayrıca, akıllı sözleşmelerin daha verimli ve güvenli çalışabilmesi için, statik analiz araçları kullanılarak potansiyel güvenlik açıklarının erken tespiti sağlanabilir. Bunun yanında, hibrit kimlik doğrulama modelleri geliştirerek, kullanıcıların ihtiyaçlarına göre geleneksel ve blokzincir tabanlı sistemler arasında dinamik geçiş yapılmasına olanak tanıyan adaptif mimariler araştırılabilir. Bu tür hibrit sistemlerde, örneğin kritik olmayan işlemlerde geleneksel doğrulama tercih edilirken, hassas veri erişimlerinde blokzincir teknolojisi devreye alınabilir. Son olarak, kullanıcı deneyimini iyileştirmek amacıyla, mobil cihazlara özel optimize edilmiş blokzincir çözümleri geliştirilebilir. Bu doğrultuda, blokzincir destekli kimlik doğrulamanın geleceği yalnızca güvenlik değil, aynı zamanda erişilebilirlik ve kullanıcı dostu çözümlerle de şekillenebilir.

5. Kaynaklar (References)

[1] L. Shinder, M. Cross, Chapter 11. Passwords, vulnerabilities, and exploits, In book: Scene of the Cybercrime (Second Edition) (2008) 467-503 doi.org/10.1016/B978-1-59749-276-8.00011-X.

[2] P. P. Piyush, A. S. Rajawat, S.B. Goyal, P. Bedi, C. Verma, M. S. Raboaca, F. M. Enescu, Authentication and authorization in modern web apps for data security using nodejs and role of dark web, Procedia Computer Science. (2022) 781-790 doi.org/10.1016/j.procs.2022.12.080.

[3] A. Henricks, H. Kettani, On data protection using multi-factor authentication, The International Conference on Information System and System Management, (2019) 1-4. doi.org/10.1145/3394788.339478

*Sorumlu Yazar/Corresponding Author: yangecdidem@gmail.com

ISSN: 2618-6055 / 9, (1), 14 – 24, 2025
DOI:10.62301/usmtd.1691033

[4] A. Ezugwu, E. Ukwandu, M. Ezema, C. Olebara, J. Ndunagu, L. Ofusori, U. Ome, Password-based authentication and the experiences of end users, Scientific Africa, (2023) e01743. doi.org/10.1016/j.sciaf.2023.e01743.

[5] C. Mole, E. Chalstrey, P. Foster, T. Hobson, Digital identity architectures: comparing goals and vulnerabilities, (2023). doi.org/10.48550/arXiv.2302.09988.

[6] N. A. Lal, S. Prasad, M. Farik, A Review of authentication methods, International Journal of Scientific & Technology Research, (2016) 246-249.

[7] A. K. Jain, A. Ross, S. Prabhakar, An introduction to biometric recognition, IEEE Transactions on Circuits and Systems for Video Technology, 14(1) (2004) 4-20. doi.org/10.1109/TCSVT.2003.818349.

[8] R. Amin, T. Gaber, G. Eltaweel, A. E. Hassanien, Biometric and traditional mobile authentication techniques: Overviews and open issues, Chapter in Intelligent Systems Reference Library, (2014) 423-446. doi.org/10.1007/978-3-662-43616-5_16.

[9] X. Wang, Z. Yan, R. Zhang, P. Zhang, Attacks and defenses in user authentication systems: A survey, Journal of Network and Computer Applications, (2021).

[10] U. Khalid, M. Asim, T. Baker, P.C. Hung, M.A. Tariq, L. Rafferty, A decentralized lightweight blockchain-based authentication mechanism for IoT systems, Cluster Comput (2020) 2067-2087. doi.org/10.1007/s10586-020-03058-6.

[11] U. Narayanan, V. Paul, S. Joseph, Decentralized blockchain based authentication for secure data sharing in cloud-IoT, J. Ambient Intell. Humaniz. Comput, (2022) 769-787. doi.org/10.1007/s12652-021-02929-z.

[12] G. Zhao, B. Di, H. He, Design and implementation of the digital education transaction subject two-factor identity authentication system based on blockchain, 22nd International Conference on Advanced Communication Technology, ICACT, IEEE (2020). doi.org/10.23919/ICACT48636.2020.9061393.

[13] S. Mercan, M. Cebe, K. Akkaya, J. Zuluaga, Blockchain-based two-factor authentication for credit card validation, Data Privacy Management, Cryptocurrencies and Blockchain Technology, Springer International Publishing (2021). doi.org/10.1007/978-3-030-93944-1_22.

[14] T.P. Abayomi-Zannu, I.A. Odun-Ayo, T. Barka, A proposed mobile voting framework utilizing blockchain technology and multi-factor authentication, Journal of Physics: Conference Series, IOP Publishing (2019). doi.org/10.1088/1742-6596/1378/3/032104.

[15] I. Wanisha, J. B. James, J. S. Witeno, L. H. M. Bakery, Multi-Factor authentication using blockchain: Enhancing privacy, security and usability (2024). doi.org/10.62951/ijcts.v1i3.24.

[16] Duende Software. https://duendesoftware.com/docs/ (erişim tarihi: 10.02.2025).

[17] S. Saranya, K. Monika, Manikandan, J. Nagaraju, S. Nagendiran, Geetha B. T., Blockchain-based identity management: Enhancing privacy and security in digital identity system, International Conference on Contemporary Computing and Informatics (IC3I) (2024). doi.org/10.1109/IC3I61595.2024.10829044.

[18] G. Zyskind, O. Nathan, A. S. Pentland, Decentralizing privacy: Using blockchain to protect personal data, IEEE Security and Privacy Workshops (2015). doi.org/10.1109/SPW.2015.27.

[19] A Blockchain Platform for the Enterprise (BPE). https://hyperledger-fabric.readthedocs.io/en/release-2.5/index.html (erişim tarihi: 20.02.2025).

[20] A. Torongo, M. Toorani, Blockchain-based decentralized identity management for healthcare systems. Cryptography and Security, (2023) 20s. doi.org/10.48550/arXiv.2307.16239.

*Sorumlu Yazar/Corresponding Author: yangecdidem@gmail.com

ISSN: 2618-6055 / 9, (1), 25 – 40, 2025
DOI:10.62301/usmtd.1696804

Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi
International Journal of Sustainable Engineering and Technology
ISSN: 2618-6055 / Cilt:9, Sayı:1, (2025), 25 – 40
Araştırma Makalesi – Research Article

TECHNICAL DEBT ASSESSMENT IN OPEN SOURCE APPLICATIONS USING STATIC CODE ANALYSIS

*Rafet GÖZBAŞI[1], Kökten Ulaş BİRANT[2]

[1]Dokuz Eylül University, The Graduate School of Natural and Applied Science, Department of Computer Engineering, İzmir
[2]Dokuz Eylül University, Faculty of Engineering, Department of Computer Engineering, İzmir

(Geliş/Received: 10.05.2025, Kabul/Accepted: 02.06.2025, Yayınlanma/Published: 30.06.2025)

ABSTRACT

The management of technical debt is critical to the sustainability of software quality throughout the evolution of software systems. This study investigates how technical debt levels change across multiple versions of four widely used open source software projects, nopCommerce, OrchardCore, RavenDB, and ShareX. Versions 30, 28, 20, and 15 of these projects, respectively, were systematically analyzed using the NDepend static analysis tool to measure technical debt levels. The results reveal that technical debt tends to accumulate, decrease, and stabilize. While nopCommerce exhibits significant increases in technical debt after major version migrations, OrchardCore maintains low levels, demonstrating the effectiveness of modular architecture in debt management. RavenDB exhibits a limited but increasing debt profile, while ShareX tends to reduce its debt level over time. These findings highlight the importance of continuously monitoring and proactively managing technical debt, especially during major structural changes. The study presents empirical findings on the evolution of technical debt in open source projects and demonstrates the value of static analysis tools such as NDepend to software quality management.

Keywords: Static code analysis, Open source applications, Static analysis tools, Technical debt

AÇIK KAYNAK KODLU UYGULAMALARDA STATİK KOD ANALİZİ İLE TEKNİK BORÇ DEĞERLENDİRMESİ

ÖZ

Teknik borcun yönetimi, yazılım sistemlerinin evrimi boyunca yazılım kalitesinin sürdürülebilirliği açısından kritik bir öneme sahiptir. Bu çalışma, yaygın olarak kullanılan dört açık kaynak yazılım projesi olan nopCommerce, OrchardCore, RavenDB ve ShareX'in çoklu sürümleri boyunca teknik borç seviyelerinin nasıl değiştiğini araştırmaktadır. NDepend statik analiz aracı kullanılarak bu projelerin sırasıyla 30, 28, 20 ve 15 sürümü sistematik olarak analiz edilmiş ve teknik borç seviyeleri ölçülmüştür. Elde edilen sonuçlar, teknik borcun birikim, azalma ve stabilite eğilimleri gösterdiğini ortaya koymuştur. nopCommerce, büyük sürüm geçişleri sonrası teknik borçta belirgin artışlar sergilerken, OrchardCore düşük seviyeleri koruyarak modüler mimarinin borç yönetimindeki etkinliğini göstermiştir. RavenDB, sınırlı ancak artış eğilimli bir borç profili sergilerken, ShareX zamanla borç seviyesini azaltma eğiliminde olmuştur. Bu bulgular, özellikle büyük yapısal değişiklikler sırasında teknik borcun sürekli izlenmesi ve proaktif olarak yönetilmesinin önemine dikkat çekmektedir. Çalışma, açık kaynak projelerde teknik borcun evrimine ilişkin ampirik bulgular sunmakta ve NDepend gibi statik analiz araçlarının yazılım kalitesi yönetimine sağladığı değeri ortaya koymaktadır.

Anahtar kelimeler: Statik kod analizi, Açık kaynak kodlu uygulamalar, Statik analiz araçları, Teknik borç

***Sorumlu Yazar/Corresponding Author:** rafetgozbasi@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 25 – 40, 2025
DOI:10.62301/usmtd.1696804

1. Introduction

In the software field, issues such as software maintainability, maintaining code quality, and code analysis are becoming increasingly important. Open source software is a type of software where the source code is publicly available and users can review and modify the software. Today, open source software is constantly updated by a large community of developers, but these updates can compromise the integrity of the software and create undesirable changes in quality metrics. Capiluppi and Ramil [1] investigated version management and maintainability issues in open source projects and analyzed the effects of version changes on the evolution of projects. The research findings emphasize the importance of version management for the maintainability of software development processes. Scacchi [2] evaluated the applicability of process improvement practices in open source projects by examining the unique problems that arise in the quality and process management of open source software and the methods for solving these problems. Stol and Fitzgerald [3] have deeply examined the complexities of open source software development processes and the difficulties that arise in managing these processes and have proposed sustainable software development methods. Beller et al. [4] studied the prevalence, configuration patterns, and time evolution of nine different static analysis tools across more than 168,000 open source projects. Configurations were often based on default settings, and developers rarely defined custom rules. This analysis showed that 65% of the alert rules were related to maintainability and 35% were related to functional errors.

Static code analysis is one of the widely used methods to measure software quality with specific metric values. Static code analysis examines the source code without running the software and allows developers to evaluate various metrics such as method complexity, technical debt, number of methods, etc. Ayewah et al. [5] evaluated the effectiveness of static analysis outputs in preventing bugs using the FindBugs tool and emphasized the importance of integrating these tools into software processes. Basutakara and Jayanthi [6] studied the role of static code analysis in improving software security and systematically summarized the methods, tools, and analysis models used for early vulnerability detection. This study stated that static analysis tools analyze the code according to abstract syntax trees, control flow graphs, and call graphs. Sultanow et al. [7] presented a machine learning-based analysis approach that aims to detect invisible defects in large-scale general software for cases where classical static code analysis falls short.

Yeboah and Popoola [8] analyzed 1600 user reviews of SonarQube, PMD, Checkstyle, and FindBugs using problem modeling to uncover users' concerns and preferences about static analysis tools. Lenarduzzi et al. [9] compared six popular static analysis tools (SonarQube, Better Code Hub, Coverity Scan, FindBugs, PMD, and Checkstyle) on 47 Java projects and examined their detection capability, overlap, and accuracy. Nachtigall et al. [10] studied the usability issues of static analysis tools around the concept of “explainability” and identified six main challenges that prevent developers from communicating effectively with these tools. In this study, seven industrial and seven academic static analysis tools were evaluated for their responses to these challenges.

NDepend is a powerful static analysis tool widely used in the .NET ecosystem. The tool offers advanced metrics for calculating and tracking technical debt. Avgeriou and Taibi [11] evaluated NDepend among the tools that focus on technical debt measurement, stating that it stands out with its comprehensive metric support and code query capabilities. Ernst et al. [12] stated that tools such as NDepend report not only design-level but also code-level rule violations, so developers should be careful to distinguish between design-related debt and routine code errors.

Coulin et al. [13] systematically examined the metrics used to measure architectural quality and revealed the relationship between these metrics and quality attributes such as maintainability, extensibility, and performance. Debbarma et al. [14] evaluated widely used complexity metrics and studied the contribution of static analysis-based metrics to software testing processes. Ludwig et al. [15] analyzed technical debt levels using code metrics such as architectural complexity obtained with the Understand tool, and showed that certain code components accumulate debt over time. Hernandez-Gonzalez et al. [16] discussed the role of fundamental metrics such as coupling, dependency, complexity, testability, and reusability in the context of software engineering in their systematic review. Nuñez-Varela et al. [17] analyzed 226 studies on source code metrics and identified trends in this area.

*Sorumlu Yazar/Corresponding Author: rafetgozbasi@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 25 – 40, 2025
DOI:10.62301/usmtd.1696804

Cunningham [18] first defined the concept of technical debt and drew attention to the problems that short-term decisions can create in the long term. The role and effects of technical debt in the software development process are conceptually expressed in this study. Kruchten, Nord, and Ozkaya [19] thoroughly examined the theoretical background and practical applications of the concept of technical debt and presented a comprehensive study emphasizing the importance of technical debt in software engineering processes. Alves et al. [20] examined technical debt indicators and their reflections on software projects and proposed methods for measuring and monitoring technical debt. Ernst et al. [21] analyzed the effectiveness of static analysis tools in identifying technical debt, supporting the importance of these tools in software engineering processes with empirical data. Fontana et al. [22] stated that the tools provide an index that gives an overall assessment of technical debt in a project and explained how to calculate technical debt indices using 5 different tools.

The concept of technical debt is not just a theoretical framework, but a concrete problem that software teams encounter in their daily practice. Holvitie et al [23] stated that with the widespread use of the agile development approach, teams are exposed to an increase in technical debt due to reasons such as frequently changing customer requirements, time pressure, insufficient test coverage, incomplete documentation and resource constraints. This accumulation leads to a decrease in software quality, increased maintenance costs, and structural complexities that make it difficult to manage technical debt. Recent studies emphasize that technical debt accumulates not only at the code level, but also at the architecture, testing, documentation, and process levels, and that failure to effectively manage these debts threatens the success of software projects [24].

With the acceleration of digitalization, software quality has become one of the priority agendas of both researchers and developers. Kokol [25] analyzed the publications in this field and revealed that studies on software quality have increased rapidly in recent years and that research focuses especially on topics such as the development of software engineering processes, testing techniques, and error prediction based on machine learning. Although static analysis tools are effective in detecting software errors early, they can lose the trust of developers due to high false alarm rates. Kang, Aw, and Lo [26] found that the "Golden Features" feature set proposed to increase the accuracy of these tools appeared to be more successful than it actually was due to data leakage and duplicate data. It was also stated that more reliable evaluation methods should be developed to ensure label accuracy.

The issue of technical debt is related to the fact that choices made for the sake of short-term gains make the maintenance and development process of the software difficult in the long run. Melo et al. [27] stated that there are still significant gaps in the identification and measurement of technical debt, especially in the requirements phase, and this can negatively affect software quality. Addressing the general conceptual confusion in this field, Junior and Travassos [28] systematically examined the existing literature on the definition, types, causes and management of technical debt and argued that a common understanding model should be created. Finally, Yadav et al. [29] reported that they achieved more successful results in metrics such as accuracy and faulty sample detection than classical methods with the HEHO-CLSTM method they developed to predict software quality with artificial intelligence.

The main research problem addressed in this study is to answer the question of how the concept of technical debt changes across different software versions and how these changes affect software quality. The questions of whether changes across different versions of a software show a certain trend as open source software versions evolve and whether commonalities or differences exist in technical debt changes across different types of software will also be addressed. In this context, studies analyzing version-based technical debt across different software types are limited. This study aims to fill this gap in the literature by comparatively analyzing the software development processes of different application types.

The primary objective of this study is to examine how technical debt evolves across different versions of structurally diverse open source software projects and to evaluate its implications on software quality. To this end, the research focuses on four widely used .NET-based applications (nopCommerce, OrchardCore, RavenDB and ShareX) analyzing a total of 93 versions using the NDepend static code analysis tool. By applying a uniform metric framework, this study enables both intra-project and cross-project comparisons. The resulting insights aim to guide software developers and researchers by identifying how architectural choices and versioning strategies impact debt accumulation. Furthermore,

*Sorumlu Yazar/Corresponding Author: rafetgozbasi@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 25 – 40, 2025
DOI:10.62301/usmtd.1696804

the study contributes a replicable and systematic methodology for version-based static analysis that enhances the understanding of software maintainability through empirical evidence.

Although there are many studies analyzing technical debt using static code analysis, few have conducted a longitudinal, version-based investigation across multiple types of open source systems with a consistent metric framework. This study differentiates itself by simultaneously analyzing four structurally distinct applications over a total of 93 versions, using a uniform methodology and metric focus via NDepend. Unlike prior studies that often focus on a single system or metric, this research enables both intra-project and inter-project comparisons of technical debt evolution, offering a comprehensive perspective on how software architecture and versioning strategies influence debt accumulation patterns.

2. Material and Method

In this study, static code analysis method is used to analyze the structural changes of open source software over time. This approach provides numerical data about software quality only from the source code without requiring dynamic execution, and provides a strong evaluation ground, especially for comparisons between versions. In this section, technical debt is introduced, the preferred analysis tool is explained, the open source projects examined are defined, and the analysis process is explained in detail step by step.

2.1. Software metrics and technical debt

Honglei et al. [30] defined software metrics as follows: Software metrics give some quantitative descriptions of the attributes and these attributes are extracted from the software product, software development process and related resources. Technical debt is defined as the cost that software has to pay in the long run due to compromises or shortcuts that developers take to achieve some goals [31].

Technical debt was determined both to track the structural evolution of the code and to enable objective comparison across projects. It was measured with the NDepend tool for each implementation and applied equally to each release using the same methodology.

2.2. Open source applications

In this study, four open source and .NET based applications representing different software categories were examined. The versions of the applications were obtained from their official repositories on GitHub, and the criteria of widespread use, active development, having a clear version history and representing different functional areas were taken into account in the selection process of the applications [32-35]. Thus, the analysis results obtained are intended to be both extensible and comparable across software types. In this context, nopCommerce, an e-commerce platform, OrchardCore, a content management system, RavenDB, a document-based NoSQL database, and ShareX, a screenshot-taking and sharing application, were included in the study.

NopCommerce is a well-known e-commerce platform available as free software that allows interested users to open an online store [36]. It has a layered and modular architecture. Orchard Core is an open-source, modular, multi-tenant application framework and CMS for ASP.NET Core [37]. RavenDB is described as a transactional, open-source document database written in .NET and offers a flexible data model designed to meet your needs [38]. ShareX is known as a monolithic desktop application and is described as a free and open source screenshot and screen recording software for Microsoft Windows.[39]. Since all of these applications are open source applications, their development strategies are basically the same. They are all versioned by individuals or communities.

While performing the analysis, 30 versions were included in the analysis for the nopCommerce application, 28 for OrchardCore, 20 for RavenDB, and 15 for ShareX. When selecting the version, compilable versions that contained functionally significant changes were preferred. In this way, it was possible to examine and interpret the software structure changes of the projects in a healthy way over time. All versions were organized under separate folder structures, preserving the source code integrity; analyzes were performed directly on the source files.

*Sorumlu Yazar/Corresponding Author: rafetgozbasi@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 25 – 40, 2025
DOI:10.62301/usmtd.1696804

In this study, four open source applications were selected to provide functional diversity while maintaining a consistent technical base. All four projects are built on the .NET technology stack, are publicly available under open source licenses, and are accessible via GitHub and provide a structured version history. Moreover, they represent different application domains. This diversity allows for an insightful comparison of how technical debt evolves in software systems with different architectural styles, purposes, and usage contexts.

The number of versions analyzed for each project was determined by both functional relevance and technical feasibility. While our initial goal was to analyze as many versions as possible to capture long-term evolution patterns, certain limitations emerged during the compilation and analysis process. In particular, older versions of some applications caused critical compatibility issues in Visual Studio and could not be compiled successfully. Furthermore, NDepend requires compilable assemblies to perform static code analysis. These conditions were not met due to outdated frameworks, missing dependencies, or deprecated APIs, and therefore these versions could not be included in the analysis. As a result, the number of versions included per application was not uniform, but instead is based on a subset of stable and analyzable versions.

Each selected release was selected based on two key criteria: compilability and architectural or functional importance. To ensure consistency and accuracy in metric calculation, only releases that can be compiled in a modern .NET environment (Visual Studio 2022) without major code changes were included. Priority was also given to releases that introduced significant changes, such as major releases, incremental updates affecting core modules, or structural redesigns. This ensured that the observed technical debt fluctuations were not only measurable, but also meaningfully linked to changes in system design and development decisions.

NopCommerce was specifically chosen for its maturity, commercial adoption, and detailed version history. As one of the most widely used open source eCommerce platforms in the .NET ecosystem, it undergoes regular updates with clear documentation and change logs. Additionally, nopCommerce's migration to ASP.NET Core and extensive use in real-world deployments provide a practical context for interpreting the results, thus increasing the applicability and relevance of the study findings.

2.3. Static code analaysis tools and NDepend

Various static analysis tools have been developed to assess software quality based on code-level metrics. Static analysis tools are widely used to assess maintainability, complexity, and technical debt indicators. Among them, NDepend stands out due to its deep integration with the .NET ecosystem and extensive support for technical debt measurement. Stanković [40] conducted a comparative evaluation of four static analysis tools, SonarCloud, Squore, Sonargraph, and NDepend, for the purpose of identifying technical debt in .NET projects. NDepend stands out because it can identify not only code debt but also architectural and design debt. While most of the tools use the SQALE methodology, NDepend offers more detailed analysis with its own proprietary metrics. In tests conducted on six open source projects, NDepend produced the highest technical debt estimates and detected more extensive issues compared to other tools. In conclusion, although there are differences between the tools, NDepend stands out as a powerful tool for in-depth analysis and architectural assessment.

In the study conducted by Pfeiffer and Lungu [41], while examining how technical debt and sustainability are measured by software tools, 11 popular static analysis tools were examined in detail. One of these tools, NDepend, defines technical debt as the correction period corresponding to the violated rules and calculates it in man-days. The outstanding feature of NDepend is that it offers more than 200 analysis rules in a customizable way with LINQ-based queries. The debt value corresponding to the violation of each rule can be determined by the user. In addition, indicators such as annual interest and technical debt ratio are calculated to score the total debt status of the software. The study reveals that NDepend performs technical debt calculations in a transparent and customizable manner, and in this respect, it is in a rare position in the industry. Despite its high configurability and detailed measurements, the existing literature does not provide direct empirical validation for the accuracy of NDepend's technical debt estimates. The study of Lefever et al. [42] revealed significant variability

*Sorumlu Yazar/Corresponding Author: rafetgozbasi@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 25 – 40, 2025
DOI:10.62301/usmtd.1696804

among tools using similar methodologies, meaning that even advanced tools like NDepend should be used with caution when drawing quantitative conclusions.

Avgeriou et al. [11] identified NDepend as one of the most feature-rich tools in their comparative overview of technical debt quantification platforms, noting its capabilities to calculate principal, interest, and debt ratio using customizable rule definitions. Similarly, Shaukat et al. [43] assessed the extent of architectural and maintainability issues in NDepend by comparing it with specialized analyzers for safety-critical software. Pavlič et al. [44] considered NDepend as a SQALE-based debt estimator alongside SonarQube and Squore, and found it to be compliant with industry standards for technical debt quantification. Furthermore, Saraiva et al. [45] included NDepend in a systematic mapping of TD tools, noting its effectiveness in agile environments and its recognition in previous work on supporting debt detection in evolving systems. These studies collectively confirm the relevance and widespread adoption of NDepend in both academic and industrial settings, while also providing a critical perspective on the challenges of tool-to-tool consistency in measuring algorithmic coverage, reliability, and technical debt. Their findings strengthen the methodological soundness of using NDepend to perform version-based technical debt analysis in .NET-based open source projects in this study.

During the tool selection phase, alternative tools were considered, but most of them were not compatible with .NET projects or did not support advanced technical debt estimation frameworks. Therefore, NDepend was selected based on its methodological depth, platform compatibility, and empirical support in comparative studies.

In this study, NDepend version 2024.2.1 was used and the analyzes were performed on the Visual Studio 2022 IDE in a development environment with Windows 11 operating system. Thanks to the analysis performed, it was possible to make a comparative evaluation between different versions. NDepend was chosen because it provides the level of detailed metrics required by the study and offers the flexibility to perform version-independent analysis. Compared to alternative tools, NDepend's .NET-focused analytical depth shows higher overlap with the purpose of this study.

2.4. Analysis Process and Method Followed

In this study, static code analysis was performed on different versions of selected open source .NET applications and the obtained technical debt values were used to evaluate the variation between versions. The analysis process was carried out systematically by following the same steps for all applications and the comparability of the results was ensured. The details of this process are explained below.

First, after identifying the applications, stable versions of each application were manually downloaded from the official repository on GitHub. Each version was organized under separate folder structures to avoid confusion and preserve the integrity of the source files.

An independent NDepend project file was created for each version and the same analysis rules were applied to all versions. This is especially important for measurement consistency. The set of metrics used during the analysis was kept constant; only the technical debt metric was focused on. The analysis operations were completed efficiently thanks to the user interface provided by NDepend.

After each version analysis, the obtained technical debt values were exported via HTML formatted reports produced by NDepend. The data was compiled in Microsoft Excel, processed in separate sheets for each application and tabulated comparatively across all versions. Then, line charts were used to visualize the changes between versions. During the visualization process, trends of increasing, decreasing and fluctuation of the technical debt value were highlighted.

Since the code sizes of the applications are quite different from each other, an approach based on ratios rather than direct absolute values was adopted in order to make a fair comparison between the applications. The order of the basic operations performed during the analysis process is given in Figure 1 as a flow chart.

*Sorumlu Yazar/Corresponding Author: rafetgozbasi@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 25 – 40, 2025
DOI:10.62301/usmtd.1696804

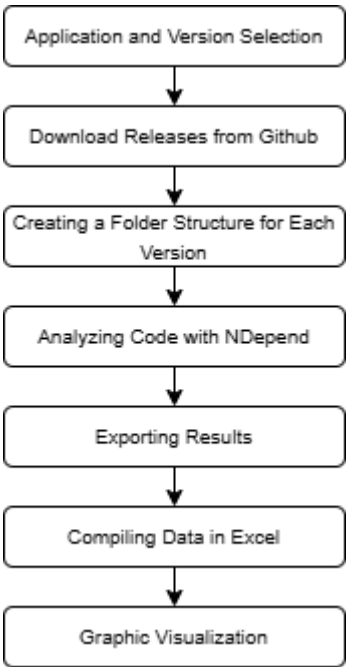


Figure 1. Flow chart of the method followed

3. Result and Discussion

In this study, different versions of four open source software projects, namely nopCommerce, OrchardCore, RavenDB and ShareX, are analyzed in terms of technical debt metric using the NDepend tool. The technical debt trend of each project across versions is evaluated both within itself and in comparison with other applications.

The technical debt value for each version for nopCommerce is given in Table 1. Changes in the analysis between versions are given in Figure 2. As seen in Figure 2, the technical debt in the nopCommerce application, which was initially measured as 5.54%, decreased to below 5% in version 4.30. However, with version 4.40, the technical debt ratio increased to 6.44% and continued to decrease with micro changes in subsequent versions. In version 4.80.0, it changed significantly and approached almost 7%. This trend shows that technical debt increases in major version transitions.

Table 1. Version-technical debt values for nopCommerce

Version	Debt
3.80	5,54%
3.90	5,67%
4.00	5,66%
4.10	5,69%
4.30	4,83%
4.40	6,44%
4.40.1	6,45%
4.40.2	6,45%
4.40.3	6,44%
4.40.4	6,45%
4.50.0	6,35%
4.50.1	6,35%
4.50.2	6,36%
4.50.3	6,35%
4.50.4	6,35%
4.60.0	6,34%
4.60.1	6,34%
4.60.2	6,34%
4.60.3	6,33%
4.60.4	6,33%
4.60.5	6,34%

*Sorumlu Yazar/Corresponding Author: rafetgozbasi@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 25 – 40, 2025
DOI:10.62301/usmtd.1696804

4.60.6	6,34%
4.70.0	6,30%
4.70.1	6,30%
4.70.2	6,30%
4.70.3	6,30%
4.70.4	6,30%
4.70.5	6,33%
4.80.0	6,71%
4.80.1	6,70%

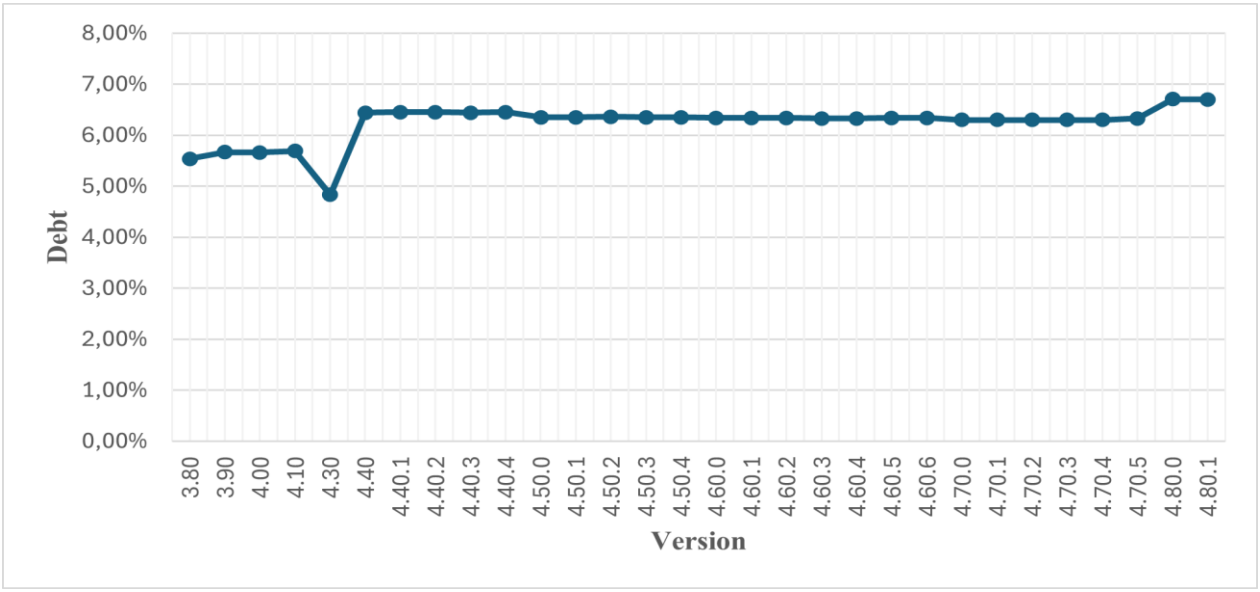


Figure 2. Version-based technical debt values of the nopCommerce application

The technical debt value for each version of OrchardCore is given in Table 2. Changes in the analysis between versions are given in Figure 3. As seen in Figure 3, the technical debt ratio in the OrchardCore application remained stable at 3.5% for a long time. Although it fell below 3.5% starting from version 1.6.0, it approached 3.5% again in version 1.8.0. A significant improvement in technical debt was seen in the transition to version 2.0.0, and the ratio fell to 2.9%. Since this version, the technical debt ratio has remained low with minor fluctuations.

Table 2. Version-technical debt values for OrchardCore

Version	Debt
0.0.1	3,53%
0.0.2	3,53%
0.0.3	3,53%
0.0.4	3,53%
1.0.0	3,53%
1.1.0	3,53%
1.2.0	3,53%
1.2.1	3,53%
1.2.2	3,53%
1.3.0	3,53%
1.4.0	3,58%
1.5.0	3,54%
1.6.0	3,47%
1.7.0	3,38%
1.7.1	3,38%
1.7.2	3,38%
1.8.0	3,45%
1.8.1	3,45%
1.8.2	3,46%

*Sorumlu Yazar/Corresponding Author: rafetgozbasi@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 25 – 40, 2025
DOI:10.62301/usmtd.1696804

1.8.3	3,46%
1.8.4	3,43%
2.0.0	2,92%
2.0.1	2,91%
2.0.2	2,92%
2.1.0	3,08%
2.1.1	3,09%
2.1.2	3,09%
2.1.3	3,09%

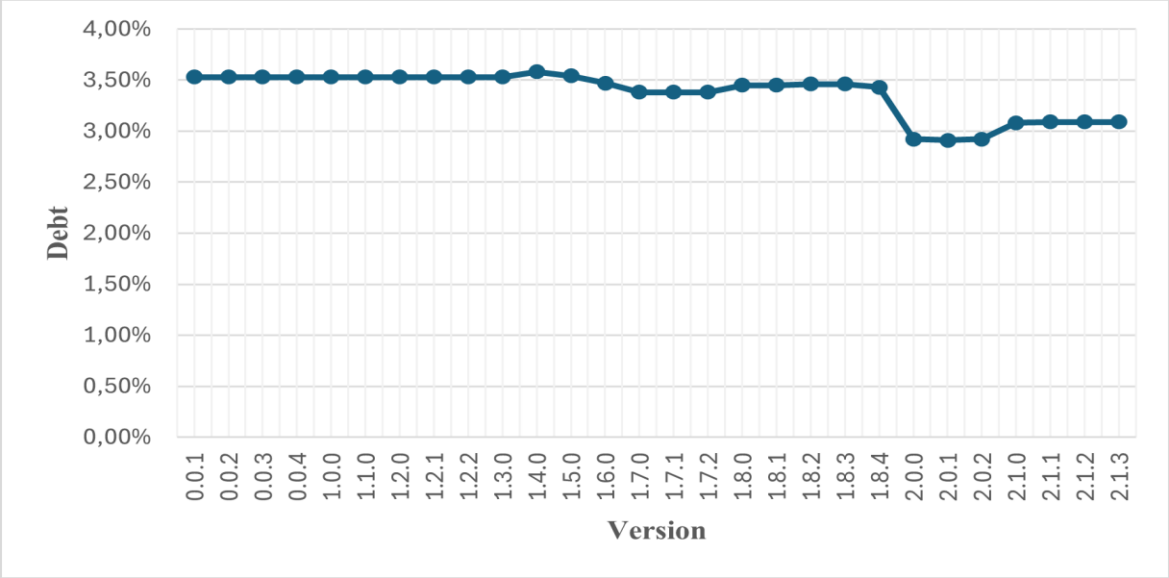


Figure 3. Version-based technical debt values of the OrchardCore application

The technical debt value for each version for RavenDB is given in Table 3. The changes in the analysis between versions are given in Figure 4. As can be seen in Figure 4, the technical debt in the RavenDB implementation was initially measured at 5.6% and increased to 5.84% in version 5.4.104. The technical debt ratio remained constant at this level in all subsequent versions and increased to 5.88% in version 5.4.202 with only minor fluctuations. This indicates a steady but limited improvement in debt management.

Table 3. Version-technical debt values for RavenDB

Version	Debt
5.4.100	5,60%
5.4.101	5,60%
5.4.102	5,61%
5.4.104	5,84%
5.4.105	5,84%
5.4.109	5,84%
5.4.110	5,85%
5.4.111	5,85%
5.4.112	5,85%
5.4.113	5,85%
5.4.114	5,85%
5.4.115	5,85%
5.4.116	5,86%
5.4.117	5,86%
5.4.118	5,85%
5.4.119	5,85%
5.4.202	5,88%
5.4.203	5,88%
5.4.204	5,89%

*Sorumlu Yazar/Corresponding Author: rafetgozbasi@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 25 – 40, 2025
DOI:10.62301/usmtd.1696804

5.4.205	5,89%

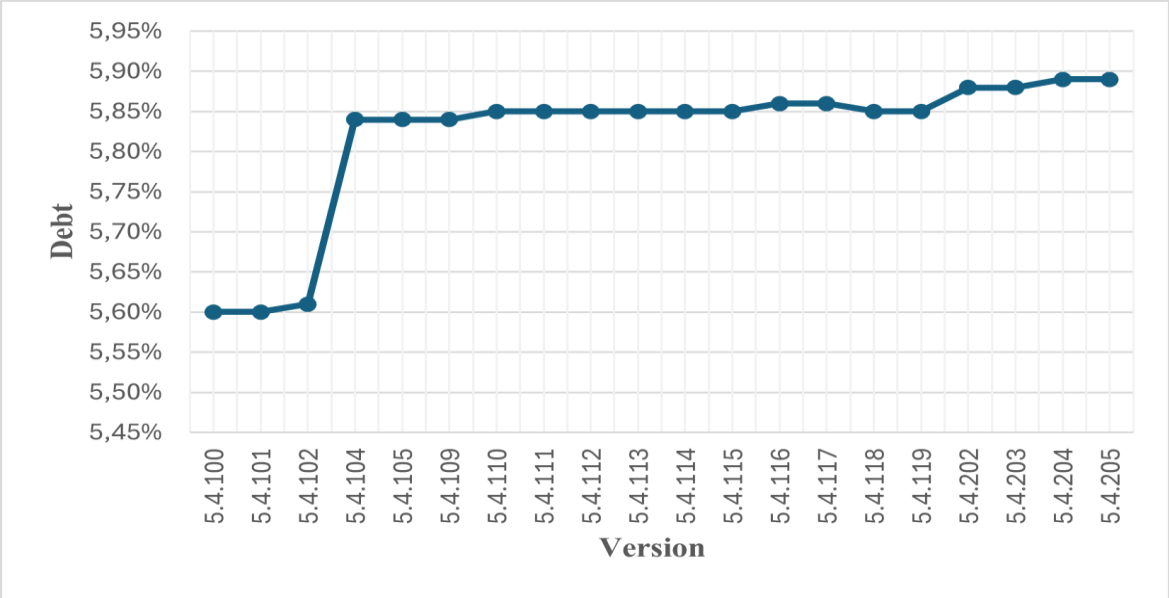


Figure 4. Version-based technical debt values of the RavenDB application

The technical debt value for each version for ShareX is given in Table 4. The changes in the analysis between versions are given in Figure 5. As seen in Figure 5, the technical debt in the ShareX application started at 5.77% in version 13.0.0 and decreased to 5.33% by version 13.2.0. It stabilized at 5.4% in subsequent versions. This trend shows that technical debt has been reduced and managed over time.

Table 4. Version-technical debt values for ShareX

Version	Debt
13.0.0	5,77%
13.0.1	5,77%
13.1.0	5,71%
13.2.0	5,33%
13.2.1	5,33%
13.3.0	5,37%
13.4.0	5,37%
13.5.0	5,38%
13.6.0	5,40%
13.6.1	5,40%
13.7.0	5,40%
14.0.0	5,47%
14.0.1	5,47%
14.1.0	5,46%
15.0.0	5,42%

*Sorumlu Yazar/Corresponding Author: rafetgozbasi@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 25 – 40, 2025
DOI:10.62301/usmtd.1696804

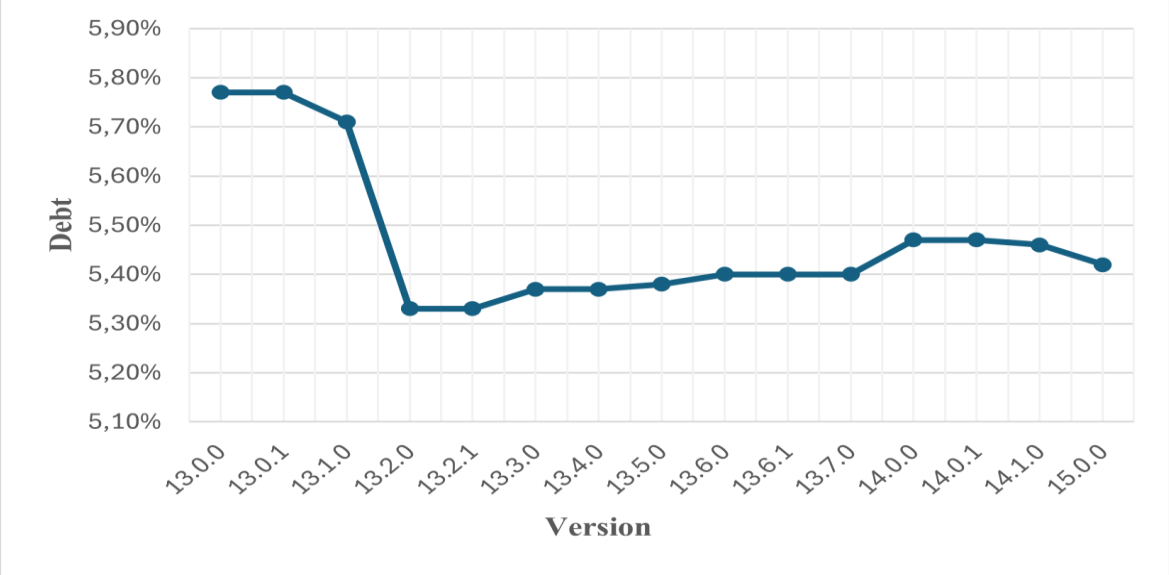


Figure 5. Version-based technical debt values of the ShareX application

The findings of this study reveal distinct technical debt trends across four open source .NET applications (nopCommerce, OrchardCore, RavenDB, and ShareX) analyzed across multiple versions using the NDepend static analysis tool. The analysis showed that technical debt does not follow a uniform pattern across all projects, but instead exhibits application-specific trends shaped by release strategy, architecture, and possibly contributor dynamics.

As seen in Figure 2, nopCommerce's technical debt percentage has shown significant increases in major releases. The debt level started at 5.54% in the first release and decreased to 4.83% in version 4.30, showing some initial improvements. However, there was a sharp increase to around 6.5% with the major release of 4.40, followed by slight decreases in minor releases. The second increase was observed in version 4.80, where the debt reached almost 7%. This pattern suggests that a significant amount of new debt is incurred during major release transitions, most likely due to significant feature additions or rapidly introduced architectural changes. In contrast, intervening minor releases allowed for incremental refactoring or stabilization that modestly reduced the debt. Overall, nopCommerce's trend has been upward over time, meaning that if proactive debt management is not implemented during major upgrades, the project accumulates "debt spikes" that increase the underlying debt level. These results highlight the need for careful planning during major releases. Large transitions often create complexity faster than it pays off, while smaller periods of change allow the team to pay off or stabilize some of its debt.

In Figure 3, OrchardCore has exhibited a remarkably stable and low technical debt profile. Over a long series of releases, the debt ratio has remained stable around 3.5%, which is significantly lower than other projects. There were minor fluctuations with a slight drop below 3.5% in version 1.6.0, then returning to 3.45% in version 1.8.0. In particular, a significant improvement occurred with the release of version 2.0.0. The technical debt ratio dropped to 2.92% and remained low in subsequent releases (with insignificant fluctuations in the range of 2.9–3.1%). This suggests that OrchardCore not only maintained a low debt level for a long time, but also took advantage of a major release to significantly reduce its technical debt. One possible reason is OrchardCore's highly modular architecture and emphasis on code quality, which can limit the spread of debt by isolating components. The modular design appears to be effective in debt management, as seen in the consistently low debt percentage. Moreover, the drop in v2.0.0 suggests that developers are using this milestone for significant refactoring or architectural optimization, fixing delays, and "paying off" debt. In summary, OrchardCore's case demonstrates that when solid architectural practices are in place, a project's technical debt can remain both low and stable, and major releases can serve as an opportunity to proactively reduce debt rather than incur it.

According to Figure 4, RavenDB's technical debt has remained relatively constant across the versions examined, with a slight upward drift. The debt percentage initially rose from 5.6% to 5.84% in version 5.4.104. Later, over the course of many incremental updates, the debt ratio remained essentially constant at this level with only minor fluctuations, gradually increasing to approximately 5.90% in the last

*Sorumlu Yazar/Corresponding Author: rafetgozbasi@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 25 – 40, 2025
DOI:10.62301/usmtd.1696804

analyzed version. This trend suggests that RavenDB's maintainers have managed to avoid any dramatic accumulation of technical debt over time. The largely flat line suggests that as new features or changes are introduced, efforts are made to offset the new debt (e.g., through code reviews or minor refactorings) so that the overall debt remains constant. In other words, the team was likely paying the interest on the technical debt as it accrued, keeping the underlying debt amount under control. The slight increase from 5.6% to 5.89% indicates a modest accumulation. This could be due to the inherent complexity of a database system or a focus on performance optimizations rather than code cleanup. However, the lack of major increases suggests a disciplined development approach where technical debt is constantly monitored and managed with a slow burn. RavenDB's profile can be viewed as a stable state of technical debt. The project has not significantly reduced its current debt or allowed it to increase, reflecting a containment strategy.

Figure 5 shows that ShareX follows the exact opposite trajectory to nopCommerce. Its technical debt has decreased over time. Starting at 5.77% in version 13.0.0, the debt percentage has decreased to 5.33% in version 13.2.0. It has stabilized at 5.4% in subsequent versions and has remained consistently lower than its initial value, albeit with very small increases between versions. This downward trend suggests that ShareX developers actively improved code quality in early 13.x versions and paid off the debt. After version 13.0.0, the team seems to have made efforts to refactor and clean up the code in minor versions, thus eliminating code smells or inefficiencies and resulting in a measurable decrease in debt. Once this low debt base was reached, the project maintained it with only minor increases, suggesting that the new debt introduced in subsequent versions was roughly offset by ongoing maintenance. ShareX's trend exemplifies successful technical debt management during evolution. Contributing factors to this situation may include a relatively small scope, active community contributions focused on quality, or a release policy that values resolving technical debt issues over adding new features. The basic idea is that technical debt does not have to inevitably increase. Technical debt can be reduced and kept under control throughout a project's lifecycle with conscious effort.

In general, debt trends vary according to the type of software and development strategies. While debt can increase rapidly in frequently changing commercial applications such as NopCommerce, debt can remain at low levels in modular and planned projects such as OrchardCore. The RavenDB example shows that debt can progress in a controlled manner, while the ShareX example shows that debt can be reduced. These findings reveal that technical debt follows a different path in each project over time. The changes observed in each project are affected not only by the code itself, but also by the development approach, release frequency, and architectural preferences. Technical debt does not always have to increase; it can be reduced or kept under control with the right strategies.

Researchers have called for moving beyond metaphor to systematic management of technical debt. Kruchten et al. [19] emphasized the importance of technical debt in software engineering and that it should be treated as a fundamental part of project management rather than an abstract concept. Our multi-version analysis supports their claim. We find that unmanaged debt can undermine software quality, but systematically addressing debt (as done in OrchardCore or ShareX) yields tangible quality benefits. Our results highlight the importance of making technical debt visible and measurable in a practical way.

This study is also consistent with the literature investigating how to define and measure technical debt. Alves et al. [20] conducted a systematic mapping study of technical debt management and highlighted the need for indicators to measure and monitor debt in projects. This study addressed this need by using a static analysis tool (NDepend) to continuously measure debt across releases. This allowed for the identification of when and where debt changed, which is exactly the type of monitoring they advocated. This study also aligns with the empirical study of Ernst et al. [21] who examined practitioners' approaches to technical debt and highlighted that many teams struggle with whether to measure or manage it. Importantly, they found that static analysis tools can effectively uncover technical debt items and provide valuable support to developers. Our experience with NDepend confirms this; the tool was effective in uncovering hidden issues and trends that were difficult to track manually. The fact that the ShareX and OrchardCore teams were able to reduce debt suggests that they were likely paying attention to such tool feedback or similar quality signals, while nopCommerce's increases may indicate periods when such feedback was underestimated or overridden by feature pressure.

*Sorumlu Yazar/Corresponding Author: rafetgozbasi@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 25 – 40, 2025
DOI:10.62301/usmtd.1696804

Another important issue is how the tools represent technical debt. Fontana et al. [22] discussed technical debt indices provided by tools and stated that these indices provide an aggregate view of a project's debt and can be calculated by various static analysis platforms. The technical debt percentage from NDepend used in this study is exactly such an index and is a composite measure that condenses a large number of code issues into a single debt percentage. The fact that we use this index across multiple releases demonstrates its practical utility. As suggested by Fontana et al., an index allows for high-level tracking of debt evolution and cross-comparison across projects or releases. By presenting real-world data that explain these concepts, this study strengthens the bridge between theoretical discussions of technical debt and the practical realities observed in open source projects.

4. Conclusion

This study examined the evolution of technical debt across multiple versions of four open-source .NET-based applications using version-based static analysis with the NDepend tool. The results revealed that each project exhibited distinct patterns of technical debt change, shaped not only by the version structure but also by project characteristics and potential architectural decisions. Rather than assuming that technical debt naturally increases over time, the findings show that ShareX and OrchardCore applications achieve debt reduction or stability throughout their evolution, indicating that effective debt control is feasible even in open-source, community-driven environments.

This section discusses what the increases and decreases in technical debt observed in previous sections of the applications mean. NopCommerce's noticeable increase in technical debt after major version migrations suggests that rapid feature delivery may be prioritized over code quality, likely due to limited time for refactoring or internal review. OrchardCore, on the other hand, showed a consistent and well-managed debt profile with a significant decrease around version 2.0. This suggests that the development team made a deliberate effort to improve the internal code structure through extensive cleanup or architectural improvements. RavenDB followed a more stable trajectory where small increases in debt gradually accumulated but no major spikes were observed. This suggests that the team was continuously implementing small improvements to prevent significant increases in technical debt and a strategy of gradual containment rather than aggressive reduction. ShareX presented the opposite case where debt levels gradually decreased over time, likely reflecting successful maintenance practices and attention to code quality between feature additions. These differences suggest that technical debt is not simply a natural consequence of software age or codebase size. Instead, it evolves as a result of project-specific development strategies, architectural decisions, and how proactively a team manages quality during release changes. Projects that incorporate regular refactoring or maintenance windows into their release cycles are in a better position to control or reduce technical debt.

In practical terms, these situations suggest that development teams should integrate technical debt tracking into their release planning, not as a general quality check, but with tools that can continuously monitor metrics across multiple dimensions. It should be noted that in modular platforms that undergo frequent integrations, such as CMS or e-commerce systems, technical debt is more likely to accumulate and should be taken into account during major release transitions. Furthermore, prioritization strategies, such as refactoring before release, should be clearly defined and integrated into the lifecycle. The findings of this study suggest that failure to do so risks long-term quality erosion that may not be apparent in short-term speed gains.

The changes in technical debt levels across versions in the analyzed applications are not only general trends, but also significant numerical differences. For example, in nopCommerce, the technical debt ratio decreased to 4.83% in version 4.30, but increased to 7% in version 4.80.0, approximately 2%, indicating that major version transitions trigger technical debt accumulation. In the OrchardCore application, technical debt remained at 3.5% for a long time, decreasing to 2.92% in version 2.0.0, showing an improvement of nearly 20%. In RavenDB, the debt ratio increased from 5.60% in the first version to 5.84% in version 5.4.104, and to 5.89% in the last version, showing a steady but limited increase of 0.3%. ShareX, on the other hand, achieved an absolute improvement of 0.44%, decreasing the rate from 5.77% in version 13.0.0 to 5.33% in version 13.2.0. These numerical findings reveal the effects of software architecture, version migration strategies, and maintenance processes on technical debt, and indicate that development teams should not only monitor trends but also analyze absolute changes.

*Sorumlu Yazar/Corresponding Author: rafetgozbasi@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 25 – 40, 2025
DOI:10.62301/usmtd.1696804

As a result, it has been shown that technical debt is inevitable in software development processes but can be managed with the right strategies. It is recommended that teams handle development activities together with technical debt management, consider debt risks during version transitions, and develop regular follow-up mechanisms to ensure sustainable quality.

In future studies, similar analyses can be performed on different software types and development approaches, different analysis tools can be compared and debt management strategies can be considered from a broader perspective, and different applications can be analyzed with various metrics. Such studies will provide significant contributions to increasing the sustainability of software quality and to understanding the effects of technical debt on the software life cycle in more depth.

5. Acknowledgements

This study is derived from the master's thesis of Rafet GÖZBAŞI, a student of Dokuz Eylül University.

6. References

[1] A. Capiluppi, J.F. Ramil, Studying the evolution of open source systems at different levels of granularity: Two case studies, in: Proceedings of the 7th International Workshop on Principles of Software Evolution, IEEE, 2004, pp. 113–118.

[2] W. Scacchi, Free/open source software development, in: Proceedings of the 6th Joint Meeting of the European Software Engineering Conference and the ACM SIGSOFT Symposium on the Foundations of Software Engineering, 2007, pp. 459–468.

[3] K.J. Stol, B. Fitzgerald, Inner source—adopting open source development practices in organizations: a tutorial, IEEE Softw. 32 (4) (2014) 60–67.

[4] M. Beller, R. Bholanath, S. McIntosh, A. Zaidman, Analyzing the state of static analysis: a large-scale evaluation in open source software, in: Proceedings of the 23rd IEEE International Conference on Software Analysis, Evolution, and Reengineering (SANER), 2016, pp. 470–481.

[5] N. Ayewah, W. Pugh, J.D. Morgenthaler, J. Penix, Y. Zhou, Using findbugs on production software, in: Companion to the 22nd ACM SIGPLAN Conference on Object-Oriented Programming Systems and Applications, 2007, pp. 805–806.

[6] B.S. Basutakara, P.N. Jeyanthi, A review of static code analysis methods for detecting security flaws, J. Univ. Shanghai Sci. Technol. 23 (6) (2021) 647–653.

[7] E. Sultanow, A. Ullrich, S. Konopik, G. Vladova, Machine learning based static code analysis for software quality assurance, in: Proceedings of the 13th International Conference on Digital Information Management (ICDIM), 2018, pp. 156–161.

[8] J. Yeboah, S. Popoola, Uncovering user concerns and preferences in static analysis tools: a topic modeling approach, in: Proceedings of the 2nd International Conference on Artificial Intelligence, Blockchain, and Internet of Things (AIBThings), 2024, pp. 1–6.

[9] V. Lenarduzzi, F. Pecorelli, N. Saarimaki, S. Lujan, F. Palomba, A critical comparison on six static analysis tools: detection, agreement, and precision, J. Syst. Softw. 198 (2023) 111575.

[10] M. Nachtigall, L. Nguyen Quang Do, E. Bodden, Explaining static analysis—a perspective, in: Proceedings of the 34th IEEE/ACM International Conference on Automated Software Engineering Workshop (ASEW), 2019, pp. 29–32.

[11] P.C. Avgeriou, D. Taibi, A. Ampatzoglou, F.A. Fontana, T. Besker, A. Chatzigeorgiou, et al., An overview and comparison of technical debt measurement tools, IEEE Softw. 38 (3) (2020) 61–71.

[12] N.A. Ernst, S. Bellomo, I. Ozkaya, R.L. Nord, What to fix? Distinguishing between design and non-design rules in automated tools, in: Proceedings of the IEEE International Conference on Software Architecture (ICSA), IEEE, 2017, pp. 165–168.

[13] T. Coulin, M. Detante, W. Mouchère, F. Petrillo, Software architecture metrics: a literature review, arXiv preprint arXiv:1901.09050 (2019).

*Sorumlu Yazar/Corresponding Author: rafetgozbasi@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 25 – 40, 2025
DOI:10.62301/usmtd.1696804

[14] M.K. Debbarma, S. Debbarma, N. Debbarma, K. Chakma, A. Jamatia, A review and analysis of software complexity metrics in structural testing, Int. J. Comput. Commun. Eng. 2 (2013) 129–133.

[15] J. Ludwig, S. Xu, F. Webber, Compiling static software metrics for reliability and maintainability from GitHub repositories, in: Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics (SMC), 2017, pp. 5–9.

[16] E.Y. Hernandez-Gonzalez, A.J. Sanchez-Garcia, M.K. Cortes-Verdin, J.C. Perez-Arriaga, Quality metrics in software design: a systematic review, in: Proceedings of the 7th International Conference in Software Engineering Research and Innovation (CONISOFT), 2019, pp. 80–86.

[17] A.S. Nuñez-Varela, H.G. Pérez-Gonzalez, F.E. Martínez-Perez, C. Soubervielle-Montalvo, Source code metrics: a systematic mapping study, J. Syst. Softw. 128 (2017) 164–197.

[18] W. Cunningham, The WyCash portfolio management system, ACM Sigplan OOPS Messenger 4 (2) (1992) 29–30.

[19] P. Kruchten, R.L. Nord, I. Ozkaya, Technical debt: from metaphor to theory and practice, IEEE Softw. 29 (6) (2012) 18–21.

[20] N.S. Alves, T.S. Mendes, M.G. De Mendonça, R.O. Spínola, F. Shull, C. Seaman, Identification and management of technical debt: a systematic mapping study, Inf. Softw. Technol. 70 (2016) 100–121.

[21] N.A. Ernst, S. Bellomo, I. Ozkaya, R.L. Nord, I. Gorton, Measure it? manage it? ignore it? software practitioners and technical debt, in: Proceedings of the 10th Joint Meeting on Foundations of Software Engineering, 2015, pp. 50–60.

[22] F.A. Fontana, R. Roveda, M. Zanoni, Technical debt indexes provided by tools: a preliminary discussion, in: Proceedings of the 8th IEEE International Workshop on Managing Technical Debt (MTD), IEEE, 2016, pp. 28–31.

[23] J. Holvitie, S.A. Licorish, R.O. Spínola, S. Hyrynsalmi, S.G. MacDonell, T.S. Mendes, V. Leppänen, 2Technol. 96 (2018) 141–160.

[24] P. Avgeriou, I. Ozkaya, A. Chatzigeorgiou, M. Ciolkowski, N.A. Ernst, R.J. Koontz, F. Shull, Technical debt management: the road ahead for successful software delivery, in: Proc. 2023 IEEE/ACM Int. Conf. Softw. Eng.: Future Softw. Eng. (ICSE-FoSE), IEEE, 2023, pp. 15–30.

[25] P. Kokol, Software quality: how much does it matter?, Electronics 11 (16) (2022) 2485.

[26] H.J. Kang, K.L. Aw, D. Lo, Detecting false alarms from automatic static analysis tools: how far are we?, in: Proc. 44th Int. Conf. Softw. Eng. (ICSE), 2022, pp. 698–709.

[27] A. Melo, R. Fagundes, V. Lenarduzzi, W.B. Santos, Identification and measurement of Requirements Technical Debt in software development: a systematic literature review, J. Syst. Softw. 194 (2022) 111483.

[28] H.J. Junior, G.H. Travassos, Consolidating a common perspective on technical debt and its management through a tertiary study, Inf. Softw. Technol. 149 (2022) 106964.

[29] D.C. Yadav, Y. Singh, A.K. Pandey, A. Kannagi, Computerized software quality evaluation with novel artificial intelligence approach, Proc. Eng. 6 (1) (2024) 363–372.

[30] T. Honglei, S. Wei, Z. Yanan, The research on software metrics and software complexity metrics, in: Proceedings of the International Forum on Computer Science-Technology and Applications, IEEE, 2009, pp. 131–136.

[31] S.B. Pandi, S.A. Binta, S. Kaushal, Artificial intelligence for technical debt management in software development, arXiv preprint arXiv:2306.10194 (2023).

[32] nopCommerce, nopCommerce release tags, https://github.com/nopSolutions/nopCommerce/tags, 2025 (accessed 01.02.25).

[33] Orchard Core, Orchard Core release tags, https://github.com/OrchardCMS/OrchardCore/tags, 2025 (accessed 08.02.25).

*Sorumlu Yazar/Corresponding Author: rafetgozbasi@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 25 – 40, 2025
DOI:10.62301/usmtd.1696804

[34] RavenDB, RavenDB release tags, https://github.com/ravendb/ravendb/tags, 2025 (accessed 15.02.25).

[35] ShareX, ShareX release tags, https://github.com/ShareX/ShareX/tags, 2025 (accessed 22.02.25).

[36] A. Juneja, B. Sondhi, A. Sharma, Nopcommerce customization for improved functionality and user experience, 2023.

[37] Orchard Core, Orchard Core Documentation, https://docs.orchardcore.net, 2025 (accessed 27.05.25).

[38] P. Nepaliya, P. Gupta, Performance analysis of NoSQL databases, Int. J. Comput. Appl. 127 (12) (2015) 36–39.

[39] Wikipedia, ShareX, https://tr.wikipedia.org/wiki/ShareX, 2025 (accessed 27.05.25).

[40] M. Stanković, Komparativna analiza alata za statičku analizu koda u svrhu identifikacije i procene tehničkog duga u .NET projektima, Zb. Rad. Fak. Tech. Nauka Novi Sad 37 (8) (2022) 1337–1340.

[41] R.H. Pfeiffer, M. Lungu, Technical debt and maintainability: how do tools measure it?, arXiv preprint arXiv:2202.13464 (2022).

[42] J. Lefever, Y. Cai, H. Cervantes, R. Kazman, H. Fang, On the lack of consensus among technical debt detection tools, in: Proc. 2021 IEEE/ACM 43rd Int. Conf. Softw. Eng.: Softw. Eng. Pract. (ICSE-SEIP), IEEE, 2021, pp. 121–130.

[43] R. Shaukat, A. Shahoor, A. Urooj, Probing into code analysis tools: a comparison of C# supporting static code analyzers, in: Proc. 15th Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST), IEEE, 2018, pp. 455–464.

[44] L. Pavlič, T. Hliš, M. Heričko, T. Beranič, The gap between the admitted and the measured technical debt: an empirical study, Appl. Sci. 12 (15) (2022) 7482.

[45] D. Saraiva, J.G. Neto, U. Kulesza, G. Freitas, R. Reboucas, R. Coelho, Technical debt tools: a systematic mapping study, in: Proc. Int. Conf. Enterp. Inf. Syst. (ICEIS), vol. 2, 2021, pp. 88–98.

***Sorumlu Yazar/Corresponding Author:** rafetgozbasi@isparta.edu.tr

ISSN: 2618-6055 / 1, (9), 41 – 50 , 2025
DOI: 10.62301/usmtd.1701938

Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi
International Journal of Sustainable Engineering and Technology
ISSN: / Cilt: 9, Sayı: 1, (2025), 41 – 50
Araştırma Makalesi – Research Article

KONTROLLÜ DENGESİZLİK SENARYOLARINDA TOPLULUK ÖĞRENME MODELLERİN SİSTEMATİK KARŞILAŞTIRMASI

*Muhammed Abdulhamid KARABIYIK[1]

[1]Niğde Ömer Halisdemir Üniversitesi, Bor MYO, Bilgisayar Teknolojileri Bölümü, Niğde

(Geliş/Received: 19.05.2025, Kabul/Accepted: 06.06.2025, Yayınlanma/Published: 30.06.2025)

ÖZ

Bu çalışma, sınıf dengesizliğinin topluluk öğrenme algoritmaları üzerindeki etkisini kontrollü bir deneysel tasarım ile incelemeyi amaçlamaktadır. Çalışma kapsamında, Iris ve Wine veri setleri üzerinde dört farklı sınıf dağılımı senaryosu (orijinal, hafif, orta ve şiddetli dengesizlik) uygulanmış ve her senaryoda Random Forest, Gradient Boosting ve Bagging algoritmaları test edilmiştir. Değerlendirmelerde yalnızca doğruluk değil, aynı zamanda Macro-F1, Balanced Accuracy, G-Mean ve Cohen Kappa gibi çoklu performans metrikleri kullanılmıştır. Elde edilen bulgular, Gradient Boosting modelinin yüksek dengesizlik düzeylerinde ciddi performans kayıpları yaşadığını; buna karşılık Random Forest algoritmasının tüm senaryolarda kararlı ve güvenilir sonuçlar sunduğunu ortaya koymuştur. Bu yönüyle çalışma, sınıf dengesizliğine karşı dayanıklı model seçiminin ve çok boyutlu metriklerle yapılan değerlendirmelerin önemini vurgulamaktadır.

Anahtar kelimeler: Sınıf dengesizliği, topluluk öğrenme, Random Forest, performans metrikleri, dengesiz veri setleri

A SYSTEMATIC COMPARISON OF ENSEMBLE MODELS UNDER CONTROLLED CLASS IMBALANCE SCENARIOS

ABSTRACT

This study aims to investigate the impact of class imbalance on ensemble learning algorithms through a controlled experimental design. Four different class distribution scenarios (original, mild, moderate, and severe imbalance) were applied to the Iris and Wine datasets, and three ensemble models Random Forest, Gradient Boosting, and Bagging were evaluated in each scenario. Model performance was assessed using not only accuracy but also multiple metrics such as Macro-F1, Balanced Accuracy, G-Mean, and Cohen's Kappa. The findings reveal that Gradient Boosting suffered significant performance degradation under high imbalance conditions, while Random Forest consistently delivered stable and reliable results across all scenarios. These results highlight the importance of selecting imbalance-resilient models and conducting evaluations using multiple performance indicators in imbalanced classification tasks.

Keywords: Class imbalance, ensemble learning, Random Forest, performance metrics, imbalanced datasets

***Sorumlu Yazar/Corresponding Author:** abdulhamidkarabiyik@ohu.edu.tr

ISSN: 2618-6055 / 1, (9), 41 – 50 , 2025
DOI: 10.62301/usmtd.1701938

1. Giriş (Introduction)

Makine öğrenmesi, günümüzde çok çeşitli alanlarda karar verme süreçlerini otomatikleştirmek ve optimize etmek amacıyla yaygın olarak kullanılmaktadır [1]. Bu uygulamaların önemli bir kısmı, sınıflandırma problemleri etrafında şekillenmektedir [2]. Ancak sınıflandırma modellerinin başarımı, yalnızca kullanılan algoritmanın gücüne değil, aynı zamanda verinin yapısal özelliklerine de doğrudan bağlıdır. Gerçek dünya veri setlerinin önemli bir kısmı, sınıflar arasında ciddi örnek sayısı dengesizlikleri içermektedir [3]. Bu durum, özellikle azınlık sınıfların yeterince temsil edilemediği durumlarda, makine öğrenmesi modellerinin bu sınıfları doğru tahmin edememesine neden olur [4]. Özellikle doğruluk (accuracy) gibi yüzeysel metrikler, dengesiz veri setlerinde yanıltıcı sonuçlar doğurabilir. Bu tür durumlarda, topluluk öğrenme yöntemleri hem genel başarıyı artırmak hem de model kararlılığını iyileştirmek amacıyla sıklıkla tercih edilmektedir [5]. Ancak, bu modellerin sınıf dengesizliğine karşı gösterdiği performansın ne ölçüde tutarlı ve güvenilir olduğu halen araştırılmaya değer bir konudur.

Topluluk öğrenme yöntemlerinin sınıflandırma performansını artırmadaki başarısı, literatürde birçok çalışma tarafından gösterilmiştir [6–8]. Bu modellerin dengesiz veri setlerinde kullanımı da farklı araştırmalarda ele alınmış; çoğu zaman örnekleme stratejileri veya ağırlıklandırma teknikleriyle birlikte değerlendirilmiştir. Ancak bu çalışmaların önemli bir bölümü, sınıf dengesizliğini ikili sınıflandırma problemleriyle sınırlı bir çerçevede ele almakta ve analizlerini çoğunlukla yalnızca doğruluk, AUC veya ROC eğrisi gibi sınırlı metrikler üzerinden gerçekleştirmektedir [9–12]. Bununla birlikte, çok sınıflı dengesiz veri setlerinde farklı dengesizlik seviyeleri altında topluluk modellerinin ne ölçüde kararlı ve başarılı sonuçlar verdiğine dair sistematik ve çok metrikli karşılaştırmalara oldukça sınırlı sayıda çalışmada yer verilmektedir. Ayrıca, küçük ölçekli ve çok sınıflı veri setleriyle yapılan kontrollü deneylerin azlığı, bu alandaki genel geçer çıkarımların geçerliliğini de sorgulanabilir hale getirmektedir. Bu çalışmada, çok sınıflı sınıflandırma problemlerinde sınıf dengesizliğinin etkisini sistematik biçimde incelemek amacıyla, farklı seviyelerde kontrollü dengesizlik senaryoları oluşturulmuştur. İki küçük ölçekli veri seti (Iris ve Wine) üzerinde, üç popüler topluluk öğrenme modeli olan Karar Ağaçları Topluluğu (Random Forest), Artırmalı Öğrenme (Gradient Boosting) ve Torbalama (Bagging) yöntemleri değerlendirilmiştir. Her veri seti için eğitim ve test verileri sabit tutulmuş; yalnızca eğitim verisine uygulanan %60, %80 ve %90 oranlarındaki örnek dengesizlikleriyle model kararlılığı gözlemlenmiştir. Böylece modellerin aynı test seti üzerinde, farklı eğitim senaryolarında gösterdiği performans farklılıkları adil bir biçimde karşılaştırılabilmiştir. Değerlendirme sürecinde yalnızca doğruluk değil, aynı zamanda Macro-F1 skoru, Dengeleştirilmiş Doğruluk (Balanced Accuracy), G-Mean, Cohen Kappa, Hassasiyet (Precision) ve Duyarlılık (Recall) gibi çoklu metrikler dikkate alınarak, her modelin güçlü ve zayıf yönleri çok boyutlu bir yaklaşımla analiz edilmiştir.

Sunulan yaklaşımın temel katkısı, topluluk öğrenme yöntemlerinin çok sınıflı sınıflandırma problemlerinde sınıf dengesizliği karşısındaki dayanıklılığını, çok metrikli ve sistematik bir yaklaşımla değerlendirmesidir. Özellikle sabit test seti kullanılarak yalnızca eğitim verisi üzerinde uygulanan farklı dengesizlik seviyeleri sayesinde, modellerin öğrenme davranışları ve genelleme kapasiteleri doğrudan gözlemlenebilmiştir. Ayrıca, çok sayıda değerlendirme metriğinin birlikte kullanılması, yalnızca doğruluk odaklı değil, bütüncül bir performans analizi yapılmasını sağlamıştır. Elde edilen bulgular, hem akademik çalışmalar için referans niteliğinde karşılaştırmalı sonuçlar sunmakta hem de dengesiz veri setleriyle çalışan uygulayıcılar için model seçimi konusunda pratik çıkarımlar üretmektedir. Makalenin devamında, ikinci bölümde araştırmanın yöntemi, kullanılan veri setleri ve uygulama adımları detaylandırılmış; üçüncü bölümde deneysel bulgular sunulmuş dördüncü bölümde sonuçlar tartışılmış son bölümde ise genel değerlendirmelere ve gelecekteki çalışma önerilerine yer verilmiştir.

2. Materyal ve Metot (Material and Method)

Bu bölümde, çalışmada kullanılan veri setleri, uygulanan sınıf dengesizliği senaryoları, tercih edilen topluluk öğrenme modelleri, performans ölçütleri ve deneysel düzenek ayrıntılı olarak açıklanmıştır.

2.1. Veri seti (Dataset)

Sınıf dengesizliğinin etkilerini inceleyebilmek amacıyla, çok sınıflı yapıya sahip iki küçük ölçekli veri seti kullanılmıştır. Bu veri setleri Iris ve Wine veriş setleridir [13,14]. Her iki veri setinin temel

*Sorumlu Yazar/Corresponding Author: abdulhamidkarabiyik@ohu.edu.tr

ISSN: 2618-6055 / 1, (9), 41 – 50 , 2025
DOI: 10.62301/usmtd.1701938

istatistiksel özellikleri Tablo 1'de özetlenmiştir. Dengeli başlangıç yapıları sayesinde her iki veri seti de farklı dengesizlik düzeylerinin kontrollü olarak uygulanmasına olanak sağlamaktadır.

Tablo 1. Kullanılan veri setlerinin temel özellikleri

Veri Seti	Örnek Sayısı	Özellik Sayısı	Sınıf Sayısı	Sınıf DağlımıF1-Score
Iris	150	4	3	50/50/50
Wine	178	13	3	59/71/48

Tablo 1 incelendiğinde, her iki veri setinin de çok sınıflı yapıya sahip olduğu ve sınıflar arasında başlangıçta Iris veri seti tam ve Wine veri seti hafif dengeli bir dağılım gösterdiği görülmektedir. Iris veri seti tamamen dengeli bir yapıya sahipken, Wine veri setindeki küçük dengesizlik bile bazı sınıflarda temsil oranlarını anlamlı biçimde etkileyebilmektedir. Özellik sayısı bakımından Iris veri seti daha sade ve düşük boyutlu bir yapı sergilerken, Wine veri seti daha fazla özelliğe sahip olması nedeniyle modelleme açısından daha karmaşık bir örüntü sunmaktadır. Bu farklılıklar, topluluk öğrenme modellerinin hem sınıf dengesizliği hem de veri karmaşıklığı karşısındaki dayanıklılığını test etme açısından değerli bir karşılaştırma ortamı sağlamaktadır.

Her veri seti eğitim ve test olmak üzere ikiye ayrılmıştır. Dengesizlik senaryoları yalnızca eğitim verisi üzerinde uygulanmış; test seti tüm senaryolarda sabit tutularak modellerin farklı öğrenme koşulları altındaki performanslarının nesnel biçimde kıyaslanması sağlanmıştır.

2.2. Sınıf Dengesizliği Senaryoları (Class Imbalance Scenarios)

Gerçek dünya veri setlerinde sıklıkla karşılaşılan sınıf dengesizliği problemi, modellerin özellikle azınlık sınıfları doğru tahmin etme kapasitesini önemli ölçüde sınırlayabilmektedir. Bu durumu kontrollü bir biçimde analiz edebilmek amacıyla, eğitim verileri üzerinde dört farklı sınıf dağılımı senaryosu oluşturulmuştur: orijinal (tam dengeli) dağılım ve üç yapay dengesizlik düzeyi (hafif, orta ve şiddetli). Bu senaryoların temsil ettiği sınıf oranları Tablo 2'de özetlenmiştir.

Tablo 2. Uygulanan sınıf dengesizlik senaryoları

Senaryo	Sınıf 1 (%)	Sınıf 2 (%)	Sınıf 3 (%)	Açıklama
Orjinal	33.3	33.3	33.3	Tam dengeli eğitim verisi
Hafif	60.0	20.0	20.0	Hafif dengesizlik
Orta	80.0	10.0	10.0	Orta düzey dengesizlik
Şiddetli	90.0	5.0	5.0	Şiddetli dengesizlik

Tablo 2'de verilen oranlar, her bir veri setinin eğitim verisi üzerindeki örnek sayısına göre yeniden örnekleme yoluyla uygulanmıştır. Senaryolar, azınlık sınıfların giderek daha az temsil edildiği kurgusal durumları temsil etmektedir. Özellikle şiddetli dengesizlik senaryosunda, azınlık sınıflar toplamın yalnızca %5'ini oluşturmaktadır. Böylece topluluk öğrenme modellerinin bu zorlu koşullar altındaki kararlılığı ve duyarlılığı sistematik olarak test edilebilmiştir.

Dengesizlik yalnızca eğitim verisine uygulanmış; test verileri ise tüm senaryolarda sabit tutulmuştur. Bu yaklaşım sayesinde, modellerin farklı öğrenme senaryoları karşısında eşit koşullarda değerlendirilmesi sağlanmış ve performans farklarının yalnızca öğrenme sürecinden kaynaklanıp kaynaklanmadığı nesnel biçimde analiz edilebilmiştir.

2.3. Topluluk Öğrenme Modelleri (Ensemble Learning Models)

Bu üç modelin tercih edilmesinin temel nedeni, her birinin farklı bir topluluk stratejisini temsil etmesidir. Random Forest, eğitim sırasında kullanılan veri alt kümeleri ve rastgele seçilen özellikler aracılığıyla çeşitlilik sağlar; bu yönüyle aşırı öğrenmeye (overfitting) karşı dirençli bir yapı sunar. Gradient Boosting, modelleri art arda eğiterek önceki hataları düzeltmeye odaklanan artımlı bir yaklaşım benimser. Bu yapı, daha yüksek doğruluk potansiyeli sunsa da, özellikle azınlık sınıfların yanlış

*Sorumlu Yazar/Corresponding Author: abdulhamidkarabiyik@ohu.edu.tr

ISSN: 2618-6055 / 1, (9), 41 – 50 , 2025
DOI: 10.62301/usmtd.1701938

öğrenilmesine karşı daha hassas bir davranış sergileyebilir. Bagging ise bootstrap örnekleme yöntemiyle oluşturulan veri alt kümeleri üzerinde öğreniciler eğiterek model varyansını azaltır ve öğrenme sürecini daha dengeli hale getirir. Bu çeşitlilik, her bir modelin güçlü ve zayıf yönlerinin farklı dengesizlik düzeylerinde nasıl ortaya çıktığını gözlemleme açısından değerli bir karşılaştırma zemini sunmaktadır.

2.3.1. Random Forest

Random Forest, birden fazla karar ağacının bağımsız olarak eğitilmesi ve sonuçların çoğunluk oyu ile birleştirilmesi esasına dayanan bir topluluk yöntemidir. Ağaçlar, bootstrap yöntemiyle oluşturulan farklı veri alt kümeleri üzerinde eğitilirken, her düğümde rastgele seçilen özellikler üzerinden en iyi bölünme noktası belirlenir. Bu yaklaşım, öğrenici modeller arasında çeşitlilik sağlayarak genelleme başarısını artırır [15,16]. Random Forest, sınıf dengesizliği karşısında göreli olarak kararlı sonuçlar verebilmesi ve sınıf etiketlerinden bağımsız olarak bölünme kararları alabilmesi nedeniyle bu çalışmada değerlendirmeye alınmıştır.

2.3.2. Gradient Boosting

Gradient Boosting, zayıf öğrenicilerin (çoğunlukla karar ağaçlarının) art arda eğitildiği ve her yeni modelin önceki modellerin hatalarını düzeltmeye çalıştığı artımlı bir öğrenme yaklaşımıdır. Modelin her adımda eksik kaldığı örüntüleri telafi etmesi, karmaşık karar sınırlarını öğrenebilmesini sağlar. Ancak bu hassasiyet, sınıf dengesizliğine karşı kırılganlığa da neden olabilir [17,18]. Bu modelin çalışmaya dahil edilme nedeni, doğruluk potansiyelinin yüksek olması ve dengesiz örnek dağılımlarında nasıl davrandığının sistematik biçimde gözlemlenmesinin istenmesidir.

2.3.3. Bagging Classifier

Bagging (Bootstrap Aggregating), farklı veri alt örnekleriyle eğitilen zayıf sınıflandırıcıların bir araya getirilmesiyle oluşturulan bir topluluk modelidir. Her bir öğrenici model, bootstrap yöntemiyle rastgele seçilmiş veri alt kümeleriyle eğitilir; daha sonra tahminler birleştirilerek nihai karar üretilir. Bu yaklaşım model varyansını azaltır ve öğrenme sürecini daha dengeli hale getirir [19,20]. Sınıf dengesizliği durumlarında sınıfları doğrudan ağırlıklandırmadan çalışabilmesi, Bagging yönteminin bu çalışmada kullanılmasını motive eden başlıca nedendir.

2.4. Değerlendirme Metrikleri (Evaluation Metrics)

Sınıf dengesizliği içeren çok sınıflı veri setlerinde, yalnızca doğruluk (accuracy) gibi temel metrikler kullanılarak yapılan değerlendirmeler yanıltıcı olabilir. Bu nedenle, model performansının daha adil ve bütüncül bir şekilde ölçülebilmesi için çok sayıda değerlendirme metriği birlikte kullanılmıştır. Kullanılan metrikler hem genel başarıyı hem de azınlık sınıflara yönelik duyarlılığı yansıtan özellikler taşımaktadır. Tablo 3, çalışmada kullanılan başlıca metrikleri ve her birinin kısa açıklamasını özetlemektedir.

Tablo 3. Çalışmada kullanılan değerlendirme metrikleri ve kullanım gerekçeleri

Metrik	Kullanım Gerekçesi
Accuracy (Doğruluk)	Tüm sınıflar dikkate alındığında doğru tahmin edilen örneklerin toplam içindeki oranıdır. Dengesiz veri setlerinde yüksek görünebilir ama azınlık sınıfları ihmal edebilir [21].
Macro-F1 Skoru	Her sınıf için ayrı ayrı hesaplanan F1 skorlarının ortalamasıdır. Sınıf dağılımından bağımsız olduğu için dengesiz veri setlerinde daha adil bir performans ölçütüdür [22].
Balanced Accuracy	Her sınıf için doğru sınıflandırma oranlarının ortalamasıdır. Özellikle sınıf sayıları dengesiz olduğunda daha güvenilir bir genel başarı ölçüsüdür [23].
Precision (Hassasiyet)	Pozitif tahminlerin ne kadarının gerçekten doğru olduğunu gösterir. Yanlış pozitifleri azaltmak açısından önemlidir [24].
Recall (Duyarlılık)	Gerçek pozitiflerin tahmin başarına odaklanan bir metirktir. Azınlık sınıfların tespiti için kritiktir [24].
G-Mean	Her sınıf için duyarlılık skorlarının geometrik ortalamasıdır. Sınıflar arası dengeyi yansıtır, özellikle dengesiz veri setleri için anlamlıdır [25].

*Sorumlu Yazar/Corresponding Author: abdulhamidkarabiyik@ohu.edu.tr

ISSN: 2618-6055 / 1, (9), 41 – 50 , 2025
DOI: 10.62301/usmtd.1701938

Metrik	Kullanım Gerekçesi
Accuracy (Doğruluk)	Tüm sınıflar dikkate alındığında doğru tahmin edilen örneklerin toplam içindeki oranıdır. Dengesiz veri setlerinde yüksek görünebilir ama azınlık sınıfları ihmal edebilir [21].
Cohen's Kappa	Model başarımını rastgele sınıflandırmayla karşılaştırarak düzeltir. Yüksek değerler, modelin tahminlerinin istatistiksel olarak anlamlı olduğunu gösterir [26].
Log Loss	Sınıflara ait tahmin olasılıklarının doğruluğunu ölçer. Modelin ne kadar güvenle tahmin yaptığına dair bilgi verir [27].

Bu metriklerin birlikte kullanılması, yalnızca genel başarıyı değil, aynı zamanda azınlık sınıfların doğru tanınma düzeyini ve modelin kararlılığını da detaylı biçimde analiz etmeye olanak tanımaktadır. Böylece topluluk öğrenme algoritmalarının sınıf dengesizliği altındaki davranışları daha gerçekçi ve çok boyutlu şekilde değerlendirilmiştir.

2.5. Deneysel Altyapı (Experimental Framework)

Tüm deneysel süreçte veri setleri, sınıf dağılımları korunacak şekilde %80 eğitim – %20 test seti olarak ikiye ayrılmıştır. Bu ayrım sırasında katmanlı örnekleme (stratified sampling) yöntemi kullanılmış ve deneylerin tekrarlanabilirliğini sağlamak amacıyla sabit bir rastgelelik tohumu (random seed) tercih edilmiştir [28]. Değerlendirme sürecinde yalnızca eğitim verileri üzerinde dengesizlik senaryoları uygulanmış; test seti tüm senaryolarda sabit tutularak modellerin farklı öğrenme koşulları altında nesnel biçimde karşılaştırılması sağlanmıştır.

Dengesizlik senaryoları, sınıf oranları önceden belirlenmiş dört düzeye göre yapılandırılmıştır: orijinal, hafif dengesizlik, orta düzey dengesizlik ve şiddetli dengesizlik. Her senaryoda, sınıf dağılımları hedeflenen oranlara ulaşacak şekilde yeniden örneklenmiş; baskın sınıflardan rastgele örnek çıkarılarak rastgele alt örnekleme (random undersampling) uygulanmıştır. Bu yöntem, model eğitiminde sınıf dengesizliğini kontrollü biçimde simüle etmeye olanak tanımıştır.

Bu çalışmada, kullanılan üç topluluk öğrenme algoritması (Random Forest, Gradient Boosting, Bagging) varsayılan hiperparametre değerleri ile değerlendirilmiştir [29–31]. Hiperparametre optimizasyonu bilinçli olarak yapılmamıştır. Bunun temel nedeni, modellerin sınıf dengesizliğine karşı gösterdikleri tepkiyi yalnızca yapısal farklılıkları ve öğrenme stratejileri üzerinden karşılaştırmaktır. Hiperparametre ayarları, özellikle Boosting tabanlı yöntemlerde modelin duyarlılığını önemli ölçüde etkileyebilmekte ve sonuçları değiştirebilmektedir. Ancak bu tür optimizasyonlar, deneysel kontrolü azaltmakta ve model performanslarındaki farkların yalnızca algoritma kaynaklı mı yoksa parametre ayarından mı kaynaklandığını belirsizleştirebilmektedir. Bu çalışmanın amacı, algoritmaların sınıf dengesizliği karşısındaki doğal dayanıklılık düzeylerini doğrudan gözlemlemektir. Bu nedenle tüm modeller aynı koşullarda, ön ayarlarla test edilerek parametrik etkilerden arındırılmış daha nesnel bir karşılaştırma yapılması hedeflenmiştir.

Deneyler Python programlama dili ile gerçekleştirilmiş; modelleme sürecinde Scikit-learn kütüphanesi, veri setlerinin temininde ise Scikit-learn.datasets modülü kullanılmıştır [32,33]. Veri setlerinde yalnızca etiketlerin sayısal forma dönüştürülmesi ve temel biçimsel dönüşümler yapılmış; özellik standardizasyonu gerekli görülmemiştir.

3. Araştırma Bulguları (Research Findings)

Bu bölümde, elde edilen bulgular sistematik olarak sunulmuş ve hem model hem de dengesizlik düzeyine göre karşılaştırmalı analizler yapılmıştır. Ayrıca, modellerin hangi koşullarda daha kararlı ve etkili sonuçlar verdiği detaylı grafiklerle desteklenerek ortaya konulmuştur. Bulgular; dengesizlik düzeyinin artışına karşı metrik bazlı performans değişimleri, modeller arası farklar ve veri seti kaynaklı etkiler çerçevesinde çok boyutlu biçimde tartışılmıştır.

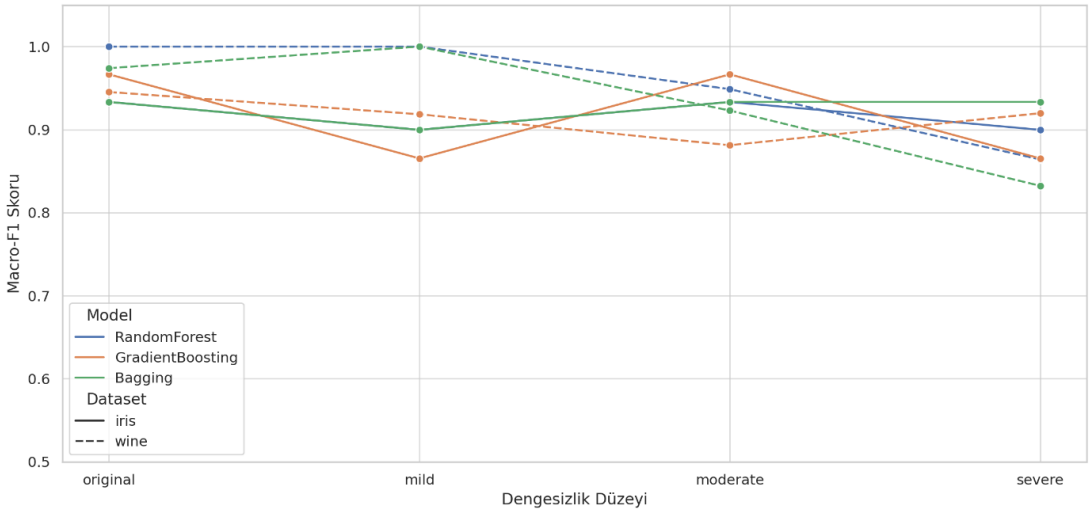
3.1. Modellerin Dengesizlik Düzeylerine Göre Genel Performans (Overall Performance of the Models Based on Imbalance Levels)

Şekil 1'de, dört farklı dengesizlik düzeyi altında her modelin Macro-F1 skorlarındaki değişim çizgisel olarak gösterilmiştir. Grafik incelendiğinde, dengesizliğin artmasıyla birlikte tüm modellerde belirgin

*Sorumlu Yazar/Corresponding Author: abdulhamidkarabiyik@ohu.edu.tr

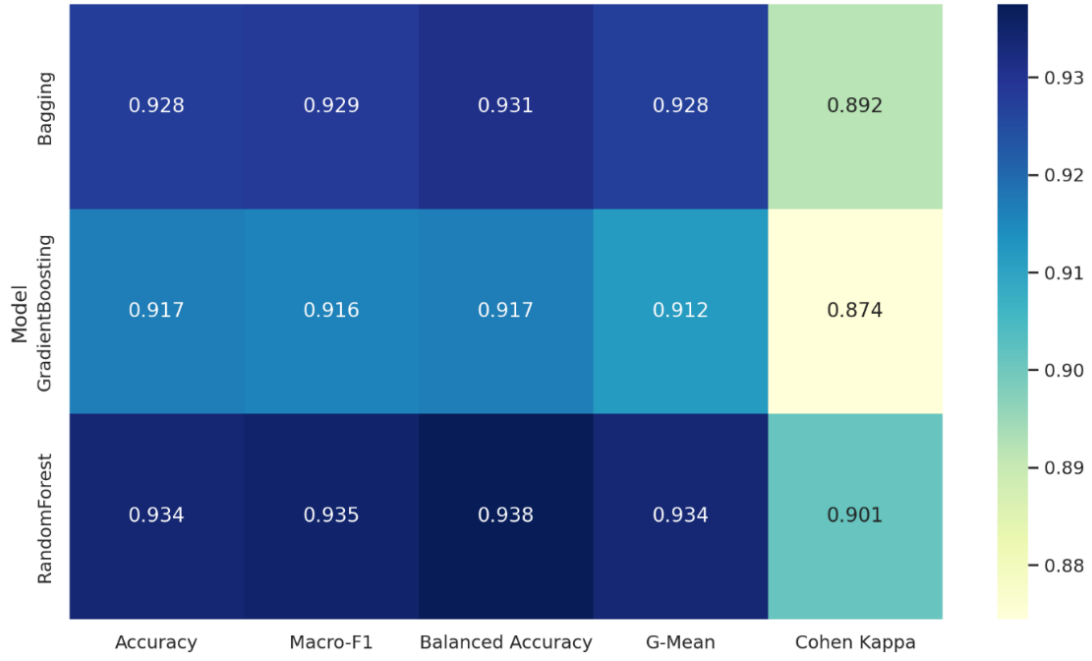
ISSN: 2618-6055 / 1, (9), 41 – 50 , 2025
DOI: 10.62301/usmtd.1701938

bir performans düşüşü gözlemlenmektedir. Bu düşüş, özellikle Gradient Boosting modelinde daha keskin olup, azınlık sınıfların sayıca az olduğu senaryolarda bu modelin duyarlılık kaybına açık olduğunu göstermektedir. Buna karşın, Random Forest ve Bagging algoritmaları daha stabil bir düşüş sergileyerek yüksek dengesizlik oranlarında dahi daha başarılı sonuçlar vermiştir. Iris veri setinde gözlenen düşüş oranları, Wine veri setine kıyasla daha az belirgin olup, bu durum veri yapısındaki özellik sayısının sonuçlar üzerinde etkisini göstermektedir.



Şekil 1. Dengesizlik düzeyine göre Macro-F1 skorlarının değişimi

Modellerin genel başarı düzeylerini özetleyen Şekil 2 ise, beş farklı metriğe göre tüm senaryoların ortalama skorlarını sunmaktadır. Bu ısı haritası, özellikle Random Forest algoritmasının tüm metriklerde en yüksek ortalama değerleri sağladığını açıkça ortaya koymaktadır. Gradient Boosting ise, özellikle Cohen Kappa ve G-Mean gibi dengesizliğe duyarlı metriklerde daha düşük performans göstermiştir. Bagging modeli ise çoğu durumda Random Forest'a yakın sonuçlar üretmiş ancak özellikle Macro-F1 ve Balanced Accuracy gibi metriklerde zaman zaman daha değişken sonuçlar vermiştir.



Şekil 2. Her modelin genel ortalama performans metrikleri

Bu iki görselle birlikte değerlendirildiğinde, sınıf dengesizliği altında topluluk öğrenme modellerinin performans profilinin yalnızca doğruluk metriğine değil, daha duyarlı metriklere göre de farklılaştığı

*Sorumlu Yazar/Corresponding Author: abdulhamidkarabiyik@ohu.edu.tr

ISSN: 2618-6055 / 1, (9), 41 – 50 , 2025
DOI: 10.62301/usmtd.1701938

açıkça görülmektedir. Bu da model seçiminin yalnızca genel doğruluğa değil, değerlendirme bağlamına göre yapılandırılması gerektiğini ortaya koymaktadır.

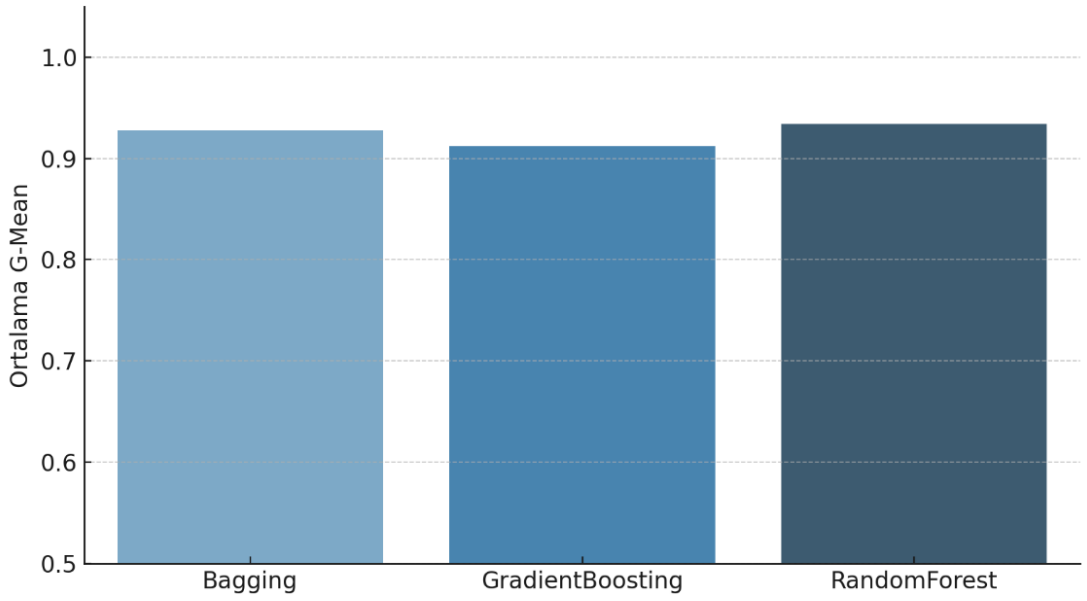
3.2.Model Bazlı Karşılaştırma Analizi (Model-Based Comparative Analysis)

Dört farklı dengesizlik düzeyi altında gerçekleştirilen değerlendirmelerde, topluluk öğrenme modelleri arasında performans açısından anlamlı farklar gözlenmiştir. Özellikle sınıf dağılımının bozulduğu senaryolarda, modellerin dengesizlik karşısındaki dirençleri ve genel kararlılık düzeyleri belirginleşmiştir.

Analiz sonuçları, Random Forest algoritmasının tüm dengesizlik senaryoları boyunca en istikrarlı performansı gösterdiğini ortaya koymaktadır. Bu model, özellikle azınlık sınıfların sayıca zayıf olduğu durumlarda dahi yüksek skorlar üretmiş ve metrikler arası dalgalanma düzeyi düşük kalmıştır. Bagging modeli de genel olarak Random Forest'a yakın sonuçlar vermiştir; ancak bazı senaryolarda (özellikle şiddetli dengesizlik altında) daha fazla performans dalgalanması gözlemlenmiştir.

Buna karşılık, Gradient Boosting algoritması, özellikle orta ve şiddetli dengesizlik senaryolarında belirgin şekilde düşük skorlar üretmiştir. Boosting yapısının, ardışıl öğrenme mekanizması nedeniyle azınlık sınıfları yeterince öğrenememesi bu düşüşün temel nedeni olabilir. Bu durum, özellikle Macro-F1, Balanced Accuracy ve G-Mean gibi dengesizliğe duyarlı metriklerde kendini açık biçimde göstermiştir.

Model karşılaştırmasını daha somut biçimde ortaya koyan Şekil 3, her modelin dört senaryo boyunca elde ettiği ortalama G-Mean skorlarını sunmaktadır. Grafik, Random Forest'ın sınıf dengesini en iyi şekilde koruyan model olduğunu açıkça göstermektedir. Gradient Boosting ise ortalama G-Mean skorlarında diğer modellere göre daha düşük seviyede kalmıştır. Bu da modelin dengesiz veri karşısında daha hassas olduğunu ve özellikle azınlık sınıflarda sınıflandırma başarısının düştüğünü göstermektedir.



Şekil 3. Modellerin Ortalama G-Mean Skorları (Tüm Senaryolar)

Sonuç olarak, sınıf dengesizliği bulunan veri setlerinde yalnızca doğruluk değil, duyarlılığı yüksek metriklere dayalı karşılaştırmalar da yapılmalı; model seçimi bu çok boyutlu değerlendirmeye göre belirlenmelidir. Bu bağlamda Random Forest, kararlılık ve dengesizlik toleransı açısından en güçlü aday olarak öne çıkmaktadır.

3.3. Veri Setine Özgü Bulgular (Findings Specific to the Dataset)

Deneysel analizlerin veri seti düzeyinde ayrıştırılması, sınıf dengesizliğinin etkisinin yalnızca model kaynaklı değil, aynı zamanda veri yapısına bağlı olarak da değiştiğini göstermektedir. Iris veri seti, başlangıçta tam dengeli bir sınıf dağılımına ve yalnızca dört özellik içeren düşük boyutlu bir yapıya

*Sorumlu Yazar/Corresponding Author: abdulhamidkarabiyik@ohu.edu.tr

ISSN: 2618-6055 / 1, (9), 41 – 50 , 2025
DOI: 10.62301/usmtd.1701938

sahiptir. Buna karşılık, Wine veri seti, başlangıçta hafif dengesiz bir yapıya ve 13 özellikten oluşan daha karmaşık bir yapıya sahiptir. Bu yapısal farklılıklar, dengesizlik karşısındaki tepkiyi doğrudan etkilemiştir.

Özellikle Wine veri setinde, dengesizlik düzeyinin artmasıyla birlikte tüm modellerde metrik bazlı performans düşüşleri daha belirgin şekilde gözlemlenmiştir. Bu durum, hem veri setinin zaten başlangıçta tam dengeli olmaması hem de yüksek boyutluluğun modelin karar sınırlarını öğrenmesini zorlaştırmasıyla ilişkilendirilebilir. Gradient Boosting, Wine veri setinde dengesizliğe karşı en zayıf performansı sergileyen model olurken; Random Forest ve Bagging bu veri setinde daha kararlı bir başarı profili çizmiştir.

Iris veri setinde ise dengesizlik düzeylerinin etkisi daha sınırlı kalmış; özellikle Random Forest ve Bagging modelleri senaryolar arası daha az dalgalanma göstermiştir. Bu durum, veri setinin sade yapısı ve sınıflar arası ayrımın daha belirgin olması sayesinde modellerin daha güvenilir karar sınırları oluşturabilmesiyle açıklanabilir. Bununla birlikte, şiddetli dengesizlik senaryosunda Gradient Boosting modelinin performansının ciddi oranda düştüğü gözlemlenmiştir. Bu düşüş, modelin azınlık sınıf temsil sayısındaki azalmaya karşı oldukça hassas olduğunu bir kez daha ortaya koymuştur.

Genel olarak değerlendirildiğinde, veri setine özgü yapısal faktörler (özellik sayısı, sınıf ayrımı belirginliği, başlangıç dengesi) sınıf dengesizliği altında model davranışlarını anlamada kritik rol oynamaktadır. Bu nedenle model değerlendirmeleri yapılırken yalnızca metrik sonuçlara değil, veri setlerinin yapısal özelliklerine de dikkat edilmesi gerektiği ortaya konulmuştur.

4. Tartışma ve Sonuç (Results and Discussion)

Bu çalışmada, sınıf dengesizliğinin topluluk öğrenme algoritmaları üzerindeki etkisi kontrollü bir deneysel düzenekle analiz edilmiştir. Dört farklı dengesizlik düzeyinde, iki çok sınıflı veri seti ve üç yaygın topluluk modeli test edilmiş; doğruluk, duyarlılık ve kararlılığı yansıtan çok sayıda metrik üzerinden değerlendirme yapılmıştır. Elde edilen bulgular, sınıf dağılımı bozuldukça Gradient Boosting modelinin performans kaybına açık olduğunu, buna karşın Random Forest algoritmasının hem yüksek doğruluk hem de metrik bazlı istikrar sağladığını ortaya koymuştur. Bagging ise genel olarak dengeli bir performans göstermiş ancak bazı senaryolarda sınıflar arası dengesizliklere duyarlı olmuştur.

Sonuçlar, model seçiminde yalnızca doğruluk metriğine değil, Macro-F1, G-Mean ve Balanced Accuracy gibi sınıf dengesini daha iyi yansıtan ölçütlere de odaklanılması gerektiğini göstermektedir. Ayrıca, veri setinin yapısı ve sınıf ayrımının belirginliği, dengesizlik karşısında model başarısını doğrudan etkilemektedir. Bu bağlamda, dengesiz veri setleriyle çalışırken hem algoritma hem de metrik seçiminde çok boyutlu ve veriye özgü bir yaklaşım benimsenmesi önerilmektedir.

Gelecek çalışmalarda, daha büyük ve daha karmaşık yapıya sahip çok sınıflı veri setleriyle benzer karşılaştırmaların yapılması önerilmektedir. Ayrıca, örnekleme tabanlı dengeleme stratejilerinin (örneğin SMOTE, Borderline-SMOTE) entegre edildiği senaryolarda modellerin dayanıklılığı daha ayrıntılı şekilde analiz edilebilir. Bunun yanında, hiperparametre optimizasyonunun etkisini izole eden ek deneyler, modeller arası farkların daha adil bir zeminde değerlendirilmesini sağlayabilir. Farklı topluluk algoritmalarının (örneğin XGBoost, LightGBM, AdaBoost) değerlendirme kapsamına dahil edilmesi de çalışmanın kapsamını genişletmeye yönelik önemli bir adım olacaktır.

5. Kaynaklar (References)

[1] Ö. ÇELİK, A Research on Machine Learning Methods and Its Applications, Journal of Educational Technology and Online Learning 1 (2018) 25–40. https://doi.org/10.31681/jetol.457046.

[2] S.B. Kotsiantis, I.D. Zaharakis, P.E. Pintelas, Machine learning: a review of classification and combining techniques, Artif Intell Rev 26 (2006) 159–190. https://doi.org/10.1007/s10462-007-9052-3.

[3] S. Yadav, G.P. Bhole, Handling Imbalanced Dataset Classification in Machine Learning, in: 2020 IEEE Pune Section International Conference (PuneCon), IEEE, 2020: pp. 38–43. https://doi.org/10.1109/PuneCon50868.2020.9362471.

*Sorumlu Yazar/Corresponding Author: abdulhamidkarabiyik@ohu.edu.tr

ISSN: 2618-6055 / 1, (9), 41 – 50 , 2025
DOI: 10.62301/usmtd.1701938

[4] H. Kaur, H.S. Pannu, A.K. Malhi, A Systematic Review on Imbalanced Data Challenges in Machine Learning, ACM Comput Surv 52 (2020) 1–36. https://doi.org/10.1145/3343440.

[5] Z.-H. Zhou, Ensemble Learning, in: Mach Learn, Springer Singapore, Singapore, 2021: pp. 181–210. https://doi.org/10.1007/978-981-15-1967-3_8.

[6] J. Beemer, K. Spoon, L. He, J. Fan, R.A. Levine, Ensemble Learning for Estimating Individualized Treatment Effects in Student Success Studies, Int J Artif Intell Educ 28 (2018) 315–335. https://doi.org/10.1007/s40593-017-0148-x.

[7] A. Mohammed, R. Kora, A comprehensive review on ensemble deep learning: Opportunities and challenges, Journal of King Saud University - Computer and Information Sciences 35 (2023) 757–774. https://doi.org/10.1016/j.jksuci.2023.01.014.

[8] O. Saidani, M. Umer, A. Alshardan, N. Alturki, M. Nappi, I. Ashraf, Student academic success prediction in multimedia-supported virtual learning system using ensemble learning approach, Multimed Tools Appl 83 (2024) 87553–87578. https://doi.org/10.1007/s11042-024-18669-z.

[9] M.R. Khalilpour Darzi, S.T.A. Niaki, M. Khedmati, Binary classification of imbalanced datasets: The case of CoIL challenge 2000, Expert Syst Appl 128 (2019) 169–186. https://doi.org/10.1016/j.eswa.2019.03.024.

[10] W. Lee, K. Seo, Downsampling for Binary Classification with a Highly Imbalanced Dataset Using Active Learning, Big Data Research 28 (2022) 100314. https://doi.org/10.1016/j.bdr.2022.100314.

[11] V. Kumar, G.S. Lalotra, P. Sasikala, D.S. Rajput, R. Kaluri, K. Lakshmanna, M. Shorfuzzaman, A. Alsufyani, M. Uddin, Addressing Binary Classification over Class Imbalanced Clinical Datasets Using Computationally Intelligent Techniques, Healthcare 10 (2022) 1293. https://doi.org/10.3390/healthcare10071293.

[12] S. Cateni, V. Colla, M. Vannucci, A method for resampling imbalanced datasets in binary classification tasks for real-world problems, Neurocomputing 135 (2014) 32–41. https://doi.org/10.1016/j.neucom.2013.05.059.

[13] P. Manikanta, D. Jayaprakash, D. Sharma, R. Bathla, Wine Quality Prediction Using Machine Learning Techniques, in: 2024 2nd International Conference on Advances in Computation, Communication and Information Technology (ICAICCIT), IEEE, 2024: pp. 1015–1018. https://doi.org/10.1109/ICAICCIT64383.2024.10912091.

[14] M. De Marsico, M. Nappi, D. Riccio, H. Wechsler, Mobile Iris Challenge Evaluation (MICHE)-I, biometric iris dataset and protocols, Pattern Recognit Lett 57 (2015) 17–23. https://doi.org/10.1016/j.patrec.2015.02.009.

[15] G. Biau, E. Scornet, A random forest guided tour, TEST 25 (2016) 197–227. https://doi.org/10.1007/s11749-016-0481-7.

[16] A. Parmar, R. Katariya, V. Patel, A Review on Random Forest: An Ensemble Classifier, in: 2019: pp. 758–763. https://doi.org/10.1007/978-3-030-03146-6_86.

[17] Y. Zhang, A. Haghani, A gradient boosting method to improve travel time prediction, Transp Res Part C Emerg Technol 58 (2015) 308–324. https://doi.org/10.1016/j.trc.2015.02.019.

[18] C. Bentéjac, A. Csörgő, G. Martínez-Muñoz, A comparative analysis of gradient boosting algorithms, Artif Intell Rev 54 (2021) 1937–1967. https://doi.org/10.1007/s10462-020-09896-5.

[19] J.G. Dias, J.K. Vermunt, A bootstrap-based aggregate classifier for model-based clustering, Comput Stat 23 (2008) 643–659. https://doi.org/10.1007/s00180-007-0103-7.

[20] T. Hothorn, B. Lausen, Double-bagging: combining classifiers by bootstrap aggregation, Pattern Recognit 36 (2003) 1303–1309. https://doi.org/10.1016/S0031-3203(02)00169-3.

[21] C.W. Fisher, E.J.M. Lauria, C.C. Matheus, An Accuracy Metric, Journal of Data and Information Quality 1 (2009) 1–21. https://doi.org/10.1145/1659225.1659229.

*Sorumlu Yazar/Corresponding Author: abdulhamidkarabiyik@ohu.edu.tr

ISSN: 2618-6055 / 1, (9), 41 – 50 , 2025
DOI: 10.62301/usmtd.1701938

[22] M.C. Hinojosa Lee, J. Braet, J. Springael, Performance Metrics for Multilabel Emotion Classification: Comparing Micro, Macro, and Weighted F1-Scores, Applied Sciences 14 (2024) 9863. https://doi.org/10.3390/app14219863.

[23] V. García, R.A. Mollineda, J.S. Sánchez, Index of Balanced Accuracy: A Performance Measure for Skewed Class Distributions, in: 2009: pp. 441–448. https://doi.org/10.1007/978-3-642-02172-5_57.

[24] H.R. Sofaer, J.A. Hoeting, C.S. Jarnevich, The area under the precision-recall curve as a performance metric for rare binary events, Methods Ecol Evol 10 (2019) 565–577. https://doi.org/10.1111/2041-210X.13140.

[25] H. Guo, H. Liu, C. Wu, W. Zhi, Y. Xiao, W. She, Logistic discrimination based on G-mean and F-measure for imbalanced problem, Journal of Intelligent & Fuzzy Systems 31 (2016) 1155–1166. https://doi.org/10.3233/IFS-162150.

[26] A. Figueroa, S. Ghosh, C. Aragon, Generalized Cohen's Kappa: A Novel Inter-rater Reliability Metric for Non-mutually Exclusive Categories, in: 2023: pp. 19–34. https://doi.org/10.1007/978-3-031-35132-7_2.

[27] A. Aggarwal, Z. Xu, O. Feyisetan, N. Teissier, On Log-Loss Scores and (No) Privacy, in: Proceedings of the Second Workshop on Privacy in NLP, Association for Computational Linguistics, Stroudsburg, PA, USA, 2020: pp. 1–6. https://doi.org/10.18653/v1/2020.privatenlp-1.1.

[28] R. Singh, N.S. Mangat, Stratified Sampling, in: 1996: pp. 102–144. https://doi.org/10.1007/978-94-017-1404-4_5.

[29] M. Moeini, Hyperparameter tuning of supervised bagging ensemble machine learning model using Bayesian optimization for estimating stormwater quality, Sustain Water Resour Manag 10 (2024) 83. https://doi.org/10.1007/s40899-024-01064-9.

[30] R.I. Alkanhel, E.-S.M. El-Kenawy, M.M. Eid, L. Abualigah, M.A. Saeed, Optimizing IoT-driven smart grid stability prediction with dipper throated optimization algorithm for gradient boosting hyperparameters, Energy Reports 12 (2024) 305–320. https://doi.org/10.1016/j.egyr.2024.06.034.

[31] Y. Rimal, N. Sharma, A. Alsadoon, The accuracy of machine learning models relies on hyperparameter tuning: student result classification using random forest, randomized search, grid search, bayesian, genetic, and optuna algorithms, Multimed Tools Appl 83 (2024) 74349–74364. https://doi.org/10.1007/s11042-024-18426-2.

[32] O. Kramer, Scikit-Learn, in: 2016: pp. 45–53. https://doi.org/10.1007/978-3-319-33383-0_5.

[33] H. Bin Mehare, J.P. Anilkumar, N.A. Usmani, The Python Programming Language, in: A Guide to Applied Machine Learning for Biologists, Springer International Publishing, Cham, 2023: pp. 27–60. https://doi.org/10.1007/978-3-031-22206-1_2.

*Sorumlu Yazar/Corresponding Author: abdulhamidkarabiyik@ohu.edu.tr

ISSN: 2618-6055 / 9, (1), 51 – 57, 2025
DOI: 10.62301/usmtd.1700194

Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi
International Journal of Sustainable Engineering and Technology
ISSN: 2618-6055 / Cilt:9, Sayı:1, (2025), 51 – 57
Araştırma Makalesi – Research Article

EAR PATHOLOGIES USING DEEP LEARNING ON OTOSCOPIC IMAGES

Yasin TATLI[1]

[1]Avrasya University, Vocational School of Health Services, Department of Medical Services and Techniques, Audiometry Programme, Trabzon, Türkiye

(Geliş/Received: 16.05.2025, Kabul/Accepted: 14.06.2025, Yayınlanma/Published: 30.06.2025)

ABSTRACT

In this study, the performance of different deep learning architectures is comparatively analyzed for the classification of ear pathologies based on otoscopic images. The dataset included four basic classes: chronic otitis media, ear wax obstruction, myringosclerosis and normal ear structure. The images were normalized at 224×224-pixel resolution and made suitable for the model, and classification was performed using CNN, CNN-LSTM, DenseNet121, ResNet50 and EfficientNet architectures. During the training and validation phases, performance metrics such as accuracy, F1 score, precision, recall and loss values were calculated, and the class discrimination power of the models was evaluated with ROC curves and complexity matrices. According to the results, CNN+LSTM and DenseNet121 architectures showed the best performance with over 94% accuracy and high F1 score in both training and validation sets. Some transfer learning-based architectures such as EfficientNet and ResNet50 showed low generalization performance. This study demonstrates the effectiveness of deep learning-based models for computerized diagnosis of intra-ear diseases and provides an important basis for decision support systems to be developed in this field.

Keywords: Deep Learning, Ear Diseases, Otoscopic Images, Image Classification.

OTOSKOPİK GÖRÜNTÜLER ÜZERİNDE DERİN ÖĞRENME KULLANARAK YAYGIN KULAK PATOLOJİLERİNİN SINIFLANDIRILMASI

ÖZ

Bu çalışmada, otoskopik görüntülere dayalı kulak patolojilerinin sınıflandırılması için farklı derin öğrenme mimarilerinin performansı karşılaştırmalı olarak analiz edilmiştir. Veri kümesi dört temel sınıf içermektedir: kronik otitis media, kulak kiri tıkanıklığı, miringoskleroz ve normal kulak yapısı. Görüntüler 224×224 piksel çözünürlükte normalize edilerek modele uygun hale getirilmiş ve sınıflandırma CNN, CNN-LSTM, DenseNet121, ResNet50 ve EfficientNet mimarileri kullanılarak gerçekleştirilmiştir. Eğitim ve doğrulama aşamalarında doğruluk, F1 skoru, kesinlik, geri çağırma ve kayıp değerleri gibi performans metrikleri hesaplanmış, modellerin sınıf ayırt etme gücü ROC eğrileri ve karmaşıklık matrisleri ile değerlendirilmiştir. Sonuçlara göre, CNN+LSTM ve DenseNet121 mimarileri hem eğitim hem de doğrulama setlerinde %94'ün üzerinde doğruluk ve yüksek F1 skoru ile en iyi performansı göstermiştir. EfficientNet ve ResNet50 gibi bazı transfer öğrenme tabanlı mimariler düşük genelleme performansı göstermiştir. Bu çalışma, kulak içi hastalıkların bilgisayarlı teşhisi için derin öğrenme tabanlı modellerin etkinliğini göstermekte ve bu alanda geliştirilecek karar destek sistemleri için önemli bir temel sağlamaktadır.

Anahtar kelimeler: Derin Öğrenme, Kulak Hastalıkları, Otoskopik Görüntüler, Görüntü Sınıflandırma.

***Sorumlu Yazar/Corresponding Author:** yasin.tatli@avrasya.edu.tr

ISSN: 2618-6055 / 9, (1), 51 – 57, 2025
DOI: 10.62301/usmtd.1700194

1.Introduction

Ear diseases can cause health problems such as hearing loss, balance disorder and infection, which are especially common during childhood [1]. Early and accurate diagnosis of such pathologies is critical for both the success of the treatment process and the quality of life of the patient [2]. Since traditional diagnostic methods are largely based on physician experience, they are open to subjective evaluations and may vary in diagnostic accuracy [3]. Therefore, the use of image processing and artificial intelligence-based diagnostic support systems is becoming increasingly important [4].

The intra-aural diseases targeted for classification in this study were divided into four main categories with visually distinguishable pathological features based on otoscopic images [5]. Chronic otitis media is characterized by prolonged inflammation of the middle ear and often perforation of the eardrum, which can lead to hearing loss, discharge and balance problems [6]. Earwax plug is a condition in which the auditory canal is completely or partially blocked by cerumen (earwax) and usually causes conductive hearing loss [7]. Myringosclerosis is characterized by scar tissue with lime-like white opacities on the tympanic membrane, mostly due to previous infections or ventilation tube placements [8]. Finally, normal ear images represent the condition in which there are no pathological findings, and a healthy eardrum can be observed with clear anatomical structures [9]. Such otoscopic classifications are the basis for the training of computer-aided diagnostic systems [10].

In recent years, the successful results of deep learning algorithms in the analysis of medical images have paved the way for a new approach in the automatic classification of intra-aural pathologies. In this context, in this study, the performance of different deep learning architectures (CNN, CNN-LSTM, DenseNet121, ResNet50, EfficientNet) in the classification of intra-aural diseases over otoscopic images is systematically evaluated and performance measurement is performed with comparative analyses.

The organization of this study consists of five main sections. The first section, Introduction, describes the purpose, scope and literature of the study, emphasizing the importance of ear diseases and the need for automatic classification. In the second section, the literature on the topics examined in this study and the analyses of previous studies are presented in detail. In the third section, the dataset used, image processing steps, model architectures (CNN, CNN-LSTM, DenseNet121, ResNet50, EfficientNet) and training processes are detailed. The fourth section includes the analysis part where the experimental results are presented with graphs and tables and the findings are compared technically. Finally, in the fifth section, an overview of the study is given, the results are summarized and recommendations for future work are presented.

2. Related Works

In recent years, deep learning-based studies on computerized analysis of otoscopic images have led to significant developments in the field of automatic diagnosis of ear diseases. Image-based artificial intelligence systems aim to increase the accuracy and reliability in the detection of middle ear pathologies, and different neural network architectures and data processing techniques have been tested in this context.

Lee et al. [5], in their in-depth study of artificial intelligence-supported image analysis systems in the diagnosis of middle ear diseases, revealed that models trained with transfer learning techniques can provide high classification success even in limited data sets. Similarly, the CNN-based system developed by Khan and colleagues [9] provided over 90% accuracy in distinguishing pathologies such as chronic otitis media in oto-endoscopic images. Such approaches demonstrate the potential of image-based decision support systems in clinical applications.

In Młyńczak and Kashani's study [7], a machine learning supported approach is proposed for the detection of effusions in images obtained using shortwave infrared otoscopy. The study shows that successful results can be obtained in the integration of alternative imaging technologies with diagnostic systems. In addition, Mahdavi [8], in his literature review on inflammatory middle ear opacities, pointed out the importance of identifying the distinctive visual features of lesions such as myringosclerosis and chronic otitis media in otoscopic diagnosis processes.

The CNN-LSTM architecture proposed by Afify et al [4] was developed with the aim of improving diagnostic accuracy by capturing spatial and sequential relationships in otoscopic images. The

*Sorumlu Yazar/Corresponding Author: yasin.tatli@avrasya.edu.tr

ISSN: 2618-6055 / 9, (1), 51 – 57, 2025
DOI: 10.62301/usmtd.1700194

researchers stated that this model offers a strong classification performance especially in sequential data. In addition, Singh and Raghuvanshi [11] obtained high success rates with a soft computing-based classification approach in otoscopic images with fuzzy boundaries and showed that such algorithms can work effectively even in images with uncertainty.

Viscaino et al [10] tested machine learning algorithms for the classification of various pathologies using an otoscopic dataset presented on the Figshare platform, and significant successes were recorded, especially in classical CNN and SVM-based systems. This dataset is used as a standard source in many studies on otoscopic images.

Chan and Stephenson's clinical guideline [1] discusses the role of otoscopic examination in the diagnosis of diseases such as otitis media, which is common in children, and how this can be enhanced by digitally assisted diagnostic systems. Furthermore, Bone [3], in his comprehensive review of otoscopic examination protocols, argues that AI-assisted analyses can improve reliability, especially in the paediatric population.

On the other hand, studies on the adaptation of current deep learning structures such as the EfficientNet architecture for ear pathologies [12] emphasize that each model should be customized according to the data and the problem. In a systematic review by Rony et al. [13], it was revealed that various architectures offer different levels of performance in pathologies such as cerumen impaction and myringosclerosis and that there is no generalized architecture.

In conclusion, current trends in the literature show that deep learning plays an important role in integrating image-based decision support systems into clinical diagnosis processes and that the choice of architecture should be customized according to the data structure. The present study aims to extend this literature and contribute to the classification of intra-aural pathologies with higher accuracy using otoscopic images.

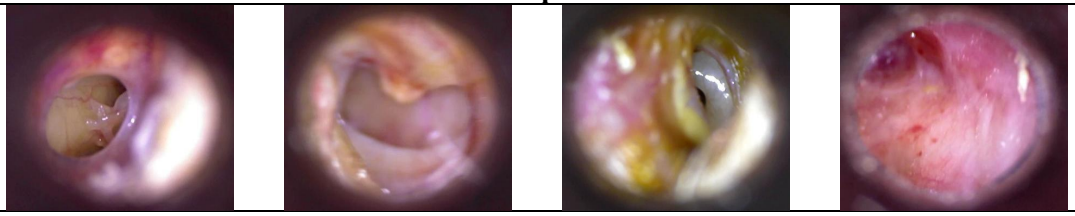
3. Material and Methods

In this section, the dataset used for the classification of intra-ear diseases is introduced. CNN, CNN-LSTM, DenseNet121, ResNet50 and EfficientNet models were applied to train the dataset. The visual data were pre-processed to a fixed size and normalized to be suitable for the model. Each model was trained using the same training and validation set, and the success levels were systematically analyzed with the metrics obtained.

3.1. Dataset

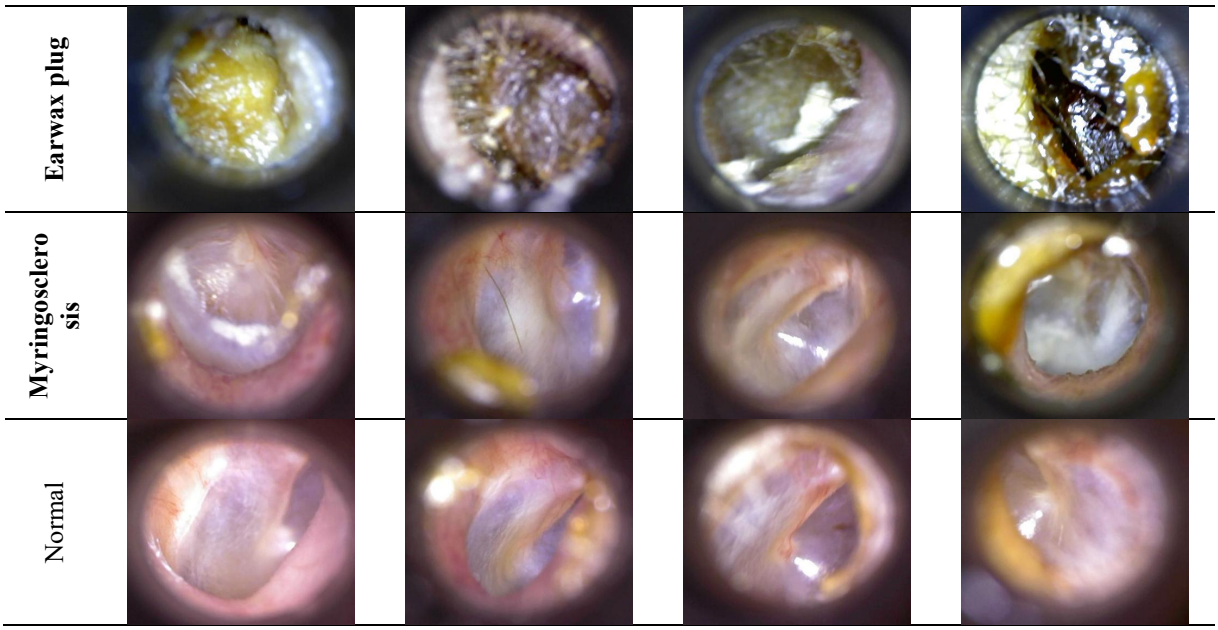
The image data used in this study were obtained from the open access dataset called "Ear Imagery Database", shown in Table 1, shared by Michelle Viscaino and Fernando Auat Cheein on the Figshare platform [14]. The dataset aims to provide a standardized resource for the classification of ear diseases and medical image analysis. The images consist of real patient data acquired with otoscope devices in clinical settings and cover various categories representing different ear disorders. The dataset is organized to include several classes, each representing a specific pathological condition. These classes include Chronic Otitis Media, Chronic Middle Ear Inflammation, Earwax Plug, Myringosclerosis and Normal ear images (800 images). The dataset consisted of 800 otoscopic images, equally distributed among four classes (200 images each for chronic otitis media, earwax plug, myringosclerosis, and normal).

Table 1. Ear Imagery Database samples with classes

Class	Samples
Chronic otitis media	

*Sorumlu Yazar/Corresponding Author: yasin.tatli@avrasya.edu.tr

ISSN: 2618-6055 / 9, (1), 51 – 57, 2025
DOI: 10.62301/usmtd.1700194



Each model was trained using the Adam optimizer with an initial learning rate of 0.0001, a batch size of 32, and a total of 50 epochs. The categorical cross-entropy loss function was used for multiclass classification. Early stopping was applied with a patience value of 10 epochs based on validation loss to avoid overfitting.

4. Findings and Results

Within the scope of the experimental study, the applicability of different deep learning architectures to the in-ear disease classification problem was tested. The results obtained are shown in Table 2, allowing the performance of each architecture in both training and validation phases to be evaluated on multidimensional metrics. CNN+LSTM and DenseNet121 models achieved remarkably high accuracy, F1 score, precision and recall values in both phases, significantly outperforming the other architectures. This indicates that both the learning and generalisation capabilities of these models are quite strong.

Table 2. Machine learning algorithms training and testing metric values

Metrics & Algorithms	CNN+LSTM	CNN	DenseNet121	ResNet50	EfficientNet
Accuracy	1.0000	0.9842	1.0000	0.5741	0.2698
f1_m	1.0000	0.9832	1.0000	0.5703	0.0293
Loss	0.0029	0.0477	$1.53e^{-06}$	0.8785	1.4864
precision_m	1.0000	0.9842	1.0000	0.6735	0.1152
recall_m	1.0000	0.9823	1.0000	0.5082	0.0172
val_accuracy	0.8562	0.8562	0.9438	0.6187	0.2500
val_f1_m	0.8526	0.8562	0.9395	0.5021	$0.0e^{+00}$
val_loss	0.5708	1.1348	0.5325	0.7633	1.4109
val_precision_m	0.8556	0.8562	0.9417	0.5365	$0.0e^{+00}$
val_recall_m	0.8500	0.8562	0.9375	0.4813	$0.0e^{+00}$

The CNN+LSTM architecture has increased its capacity to represent sequential or complex spatial relationships with the effect of LSTM layers that allow the modelling of the time dimension. The model produced a very low loss value of 0.0029 with 100% training accuracy and F1 score. This value shows that the classification process is almost error-free. Similarly, the DenseNet121 architecture also showed 100% success in the same training metrics. However, the accuracy rate (94.38%) and F1 score (93.95%) of DenseNet121 in the validation phase are slightly superior to the validation performance (85.62%) of CNN+LSTM. This result suggests that CNN+LSTM may have a stronger tendency to overfitting. The classical CNN architecture showed a balanced and stable performance in both training and validation phases. In the training phase, a successful result was obtained with an accuracy of 98.42% and an F1 score of 98.32%, while the performance in the validation phase (accuracy: 85.62%, F1: 85.62%) was lower compared to DenseNet121, but quite adequate.

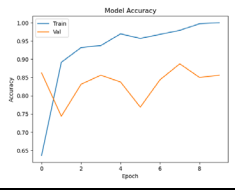
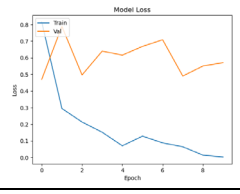
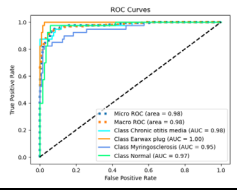
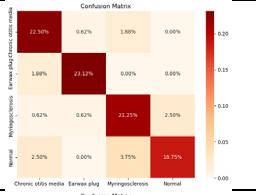
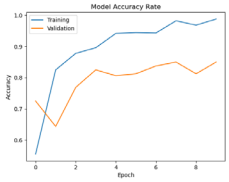
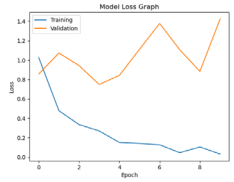
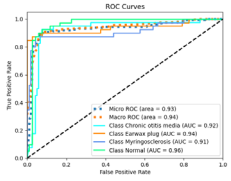
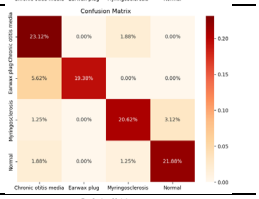
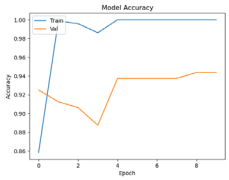
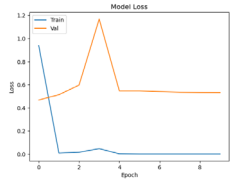
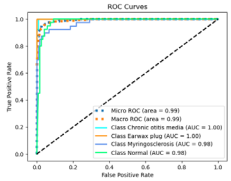
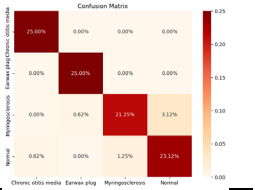
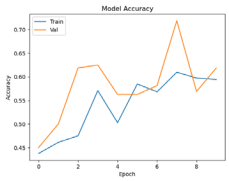
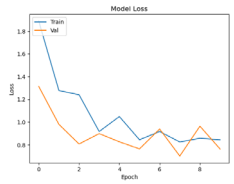
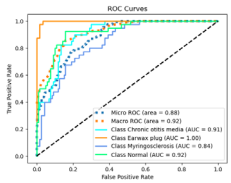
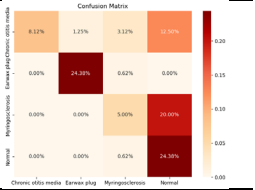
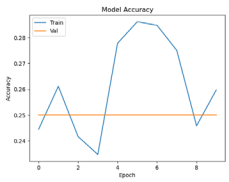
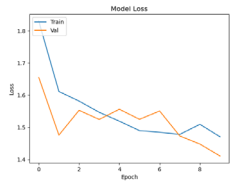
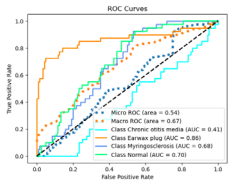
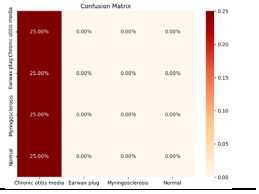
*Sorumlu Yazar/Corresponding Author: yasin.tatli@avrasya.edu.tr

ISSN: 2618-6055 / 9, (1), 51 – 57, 2025
DOI: 10.62301/usmtd.1700194

Furthermore, the validation loss (val_loss: 1.1348) indicates that there is still potential for improvement in the learning curve. The ResNet50 and EfficientNet models did not perform at the expected level within the scope of the study. The training accuracy of ResNet50 was 57.41% and the F1 score was 57.03%, while only 61.87% accuracy and 50.21% F1 score were obtained in the validation phase. This shows that the model does not adapt effectively to either the training data or the validation data. More strikingly, EfficientNet architecture produced very low scores of 26.98% accuracy and only 1.72% sensitivity in training, and almost zero success (F1: 0.00) in the validation phase. These results suggest that EfficientNet is not well suited to this dataset or model configuration and suffers from a serious learning problem.

Machine learning algorithms training and testing graphs shown in Table 3.

Table 3. Machine learning algorithms training and testing graphs

Model	Accuracy Graph	Loss Graph	ROC Curve	Confusion Matrix
CNN+LSTM				
CNN				
DenseNet121				
Resnet50				
EfficientNet				

Accuracy and loss graphs reflect that the CNN+LSTM model shows stable and high performance in both training and validation sets. Accuracy clearly increases throughout the training process, while the loss value remains low and stable. The ROC curve shows that the discrimination power between the classes is quite high (the curves are close to the upper left corner) and the AUC values are around 1.0. The confusion matrix confirms that accurate classification is performed with almost no room for error between classes. The CNN architecture performed an effective learning process with a rapid increase in the accuracy graph and a low loss value in training. In the validation data, although slight fluctuations were observed in the curves, high accuracy and low loss were generally maintained. The ROC curve shows that the discrimination between classes is strong, and the Confusion Matrix shows that the confusion between classes remains at a very low level. The training accuracy graph of the DenseNet121 architecture shows that the model reaches high accuracy very quickly, while the loss graph shows very low values.

The fact that the AUC values are close to 1.0 for almost all classes in the ROC curves proves the discriminative power of the model between classes. Confusion matrix provides almost 100% accuracy

*Sorumlu Yazar/Corresponding Author: yasin.tatli@avrasya.edu.tr

ISSN: 2618-6055 / 9, (1), 51 – 57, 2025
DOI: 10.62301/usmtd.1700194

in almost all classes. In the ResNet50 architecture, accuracy values rise and fall during the training process and fluctuations are observed in the loss graph. This indicates that the model cannot fully adapt to the data and experiences inconsistencies in the learning process. The ROC curves generally have lower AUC values, and the curves are close to the 0.5 line, indicating that the model performs close to randomly predicting some classes. Although the EfficientNet architecture aims for high performance with a low number of parameters in literature, in this study it performed well below expectations. The accuracy graph remains at low levels, and the loss value increases rapidly in the validation phase. The ROC curves clearly show that it cannot distinguish the classes, and the AUC values indicate that the classes are randomly predicted. Since the confusion matrix contains failed predictions in almost all classes, it can be concluded that this model is not compatible with the current data set or problem.

To further investigate the poor performance of the EfficientNet architecture, an ablation study was conducted. We examined different EfficientNet variants (B0, B1) and modified input normalization schemes and batch sizes. None of these adjustments led to significant improvements, suggesting that the model's depth and compound scaling may not align well with the limited dataset size and visual patterns in otoscopic images. Further analysis of misclassified images revealed that EfficientNet tended to confuse wax obstruction and chronic otitis media cases, possibly due to similar background textures.

5. Conclusions

In this study, the performance of various deep learning architectures for the classification of intra-ear diseases on otoscopic images is evaluated in detail. Experimental findings clearly show that CNN+LSTM and DenseNet121 models achieve high accuracy, F1 score, precision and sensitivity values in both training and validation processes. These two models stood out with their capacity to capture and generalise complex spatial patterns in images of intra-aural pathologies. On the other hand, the classical CNN architecture also provided balanced and reliable results, showing that effective classification is also possible with simpler structures. On the other hand, more complex structures based on transfer learning such as ResNet50 and EfficientNet have shown limited success in this context, and EfficientNet could not distinguish between classes in the validation phase. This emphasises that the compatibility of model architectures with specific problem types and datasets is not always directly proportional, and that data set specific configurations are required.

Unlike prior studies that focus predominantly on binary classification (e.g., diseased vs. normal), this work addresses the multiclass classification of four clinically distinct ear pathologies. Moreover, by comparing both custom CNN models and popular transfer learning architectures on a standardized dataset, this study provides a benchmark for future research in otoscopic image analysis.

In conclusion, this study not only demonstrates the significant potential of deep learning in the automatic diagnosis of ear diseases but also contributes to the field by presenting the relative performance of different model approaches in a comparative manner. It is predicted that the classification performance can be further improved in future studies by examining the hyperparameter optimisation of the models in detail.

References

[1] J. Chan, K. Stephenson, Diagnosis and management of middle ear disease in children. *Paediatrics and Child Health 33*(12) (2023) 376-381.

[2] T. Marom, O. Kraus, N. Habashi, S.O. Tamir, Emerging technologies for the diagnosis of otitis media. *Otolaryngology–Head and Neck Surgery 160*(3) (2019) 447-456.

[3] A. Bone, Middle Ear. *Pediatric Physical Examination-E-Book: Pediatric Physical Examination-E-Book* (2023) 192.

[4] H.M. Afify, K.K. Mohammed, A.E. Hassanien, Insight into automatic image diagnosis of ear conditions based on optimized deep learning approach. *Annals of biomedical engineering 52*(4) (2024) 865-876.

[5] D. Song, T. Kim, Y. Lee, J. Kim, Image-based artificial intelligence technology for diagnosing middle ear diseases: a systematic review. *Journal of Clinical Medicine 12*(18) (2023) 5831.

[6] F. Larrosa, L. Pujol, E. Hernández-Montero, Chronic otitis media. *Medicina Clínica (English Edition)*, (2025) 106915.

***Sorumlu Yazar/Corresponding Author:** yasin.tatli@avrasya.edu.tr

ISSN: 2618-6055 / 9, (1), 51 – 57, 2025
DOI: 10.62301/usmtd.1700194

[7] R.G. Kashani, M.C. Młyńczak, D. Zarabanda, P. Solis-Pazmino, D.M. Huland, I.N. Ahmad, T.A. Valdez, Shortwave infrared otoscopy for diagnosis of middle ear effusions: A machine-learning-based approach. *Scientific Reports 11*(1) (2021) 12509.

[8] A. Mahdavi, Diagnostic and imaging findings in inflammatory Opacifications of the middle ear: A review of the literature. *The International Tinnitus Journal 27*(2) (2023) 146-153.

[9] M.A. Khan, S. Kwon, J. Choo, S.M. Hong, S.H. Kang, I.H. Park, S.J. Hong, Automatic detection of tympanic membrane and middle ear infection from oto-endoscopic images via convolutional neural networks. *Neural Networks 126* (2020) 384-394.

[10] M. Viscaino, J.C. Maass, P.H. Delano, M. Torrente, C. Stott, F. Auat Cheein, Computer-aided diagnosis of external and middle ear conditions: A machine learning approach. *Plos one 15*(3) (2020) e0229226.

[11] A. K. Singh, A.S. Raghuvanshi, R. Mehta, A Soft Computing Approach for Efficient Diagnosis of Otitis Media Infection by Mucosal Disease Early Detection and Referrals. In *2024 2nd International Conference on Device Intelligence, Computing and Communication Technologies (DICCT) (*2024) (pp. 687-692) IEEE.

[12] Z. Cao, F. Chen, E.M. Grais, F. Yue, Y. Cai, D.W. Swanepoel, F. Zhao, Machine learning in diagnosing middle ear disorders using tympanic membrane images: a meta-analysis, *The Laryngoscope 133*(4) (2023) 732-741.

[13] M. A. H. Rony, K. Fatema, M. A. K. Raiaan, M. M. Hassan, S. Azam, A. Karim, A. Leach Artificial intelligence-driven advancements in otitis media diagnosis: a systematic review (2024) *Ieee access*.

[14] M. Viscaino and F.A. Cheein , "Ear imagery database," *Figshare*, (2020) [Online] Available: https://figshare.com/articles/dataset/Ear_imagery_database/11886630

***Sorumlu Yazar/Corresponding Author:** yasin.tatli@avrasya.edu.tr

ISSN: 2618-6055 / 9, (1), 58 – 69, 2025
DOI: 10.62301/usmtd.1703696

Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi
International Journal of Sustainable Engineering and Technology
Sayı: 1, Cilt: 9, (2025), 58 – 69
Araştırma Makalesi – Research Article

BULUT BİLİŞİM AĞ TOPOLOJİLERİNDE PERFORMANS VE HATA TOLERANSI ANALİZİ

*Hüseyin Taner GÖKMEN[1], Mevlüt ERSOY[2]

[1]Süleyman Demirel Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, Isparta
[2]Süleyman Demirel Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Isparta

(Geliş/Received: 21.05.2025, Kabul/Accepted: 15.06.2025, Yayınlanma/Published: 30.06.2025)

ÖZ

Günümüzde artan veri trafiği ve ölçeklenebilir sistem gereksinimleri doğrultusunda, bulut bilişim altyapılarında kullanılan ağ topolojilerinin performansı büyük önem taşımaktadır. Bu bağlamda, özellikle veri merkezlerinde tercih edilen Fat Tree, Z-Fat Tree ve RUFT-PL (Reduced Unidirectional Fat Tree with Parallel Links) topolojileri, çok katmanlı yapıları hem ağ içi paket iletim performansı hem de bağlantı arızalarına karşı gösterdikleri dayanıklılık açısından dikkat çekmektedir. Bu çalışmada, söz konusu topolojilerin performansları karşılaştırmalı olarak değerlendirilmiş ve farklı bağlantı arızası senaryoları altında gösterdikleri iletim başarısı, gecikme süreleri ve ağ dayanıklılığı analiz edilmiştir. Elde edilen bulgular, paralel bağlantıların hata toleransını artırdığını, tam bağlantılı yapılarda dengeli performans sağlandığını ve düşük bağlantı yoğunluğunun arıza hassasiyetini yükselttiğini göstermiştir.

Anahtar Kelimeler: Bulut Bilişim, Ağ Topolojileri, Hata Toleransı, Ağ Performansı

PERFORMANCE AND FAULT TOLERANCE ANALYSIS IN CLOUD COMPUTING NETWORK TOPOLOGIES

ABSTRACT

With the increasing volume of data traffic and the growing need for scalable systems, the performance of network topologies used in cloud computing infrastructures has become critically important. In this context, the Fat Tree, Z-Fat Tree, and RUFT-PL (Reduced Unidirectional Fat Tree with Parallel Links) topologies, which are especially preferred in data centers, stand out with their multilayered structures in terms of both packet transmission performance and resilience against link failures. In this study, the performance of these topologies is evaluated comparatively under various link failure scenarios, focusing on transmission success, latency, and network robustness. The findings reveal that parallel links enhance fault tolerance, fully connected structures offer balanced performance, and reduced connection density increases sensitivity to failures.

Keywords: Cloud Computing, Network Topologies, Fault Tolerance, Network Performance

***Sorumlu Yazar/Corresponding Author:** huseyingokmen@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 58 – 69, 2025
DOI: 10.62301/usmtd.1703696

1. Giriş (Introduction)

Bulut bilişim, bilgi teknolojisinin bir üründen hizmete dönüşmesini sağlayarak yazılım, platform ve altyapı kaynaklarını internet üzerinden talep üzerine ölçeklenebilir hizmetler olarak sunar. Uzak bir konumdaki kaynaklara erişim, bunların yapılandırılması ve yönetilmesi, işletmelere fiziksel donanım ve yazılım yükleme ihtiyacını ortadan kaldırır. ABD Ulusal Standartlar ve Teknoloji Enstitüsü (NIST)'e göre, bulut bilişim; sunucular, ağlar ve uygulamalar gibi bilgi işlem kaynaklarına kolayca erişim sağlayan, minimum yönetim çabasıyla hızla sağlanabilen bir modeldir. Bulut bilişim, işletmelerin ihtiyaçlarına esnek çözümler sunarak büyük miktarda veriyi yönetmeyi kolaylaştırır [1-4]. Bulut bilişim altyapısı, kullanıcılara bulut bilişim kaynakları ve hizmetleri sunmak için gerekli olan bilgisayarları, depolama aygıtlarını, ağ tesislerini ve diğer ilgili bileşenleri içerir [5].

Bulut bilişim, son kullanıcılara talep üzerine çeşitli hizmetler sunarak, işletmelerin ve kullanıcıların uygulamaları fiziksel makinelerde barındırmadan kullanmalarını sağlar ve internet üzerinden gerekli kaynaklara kolay erişim imkanı tanır. Hızla gelişen bu teknoloji, giderek daha fazla işletme ve kurumsal uygulamanın barındırılması için tercih edilmektedir [6,7]. Ancak, bulut tabanlı hizmetlerin yaygınlaşması, hem servis sağlayıcılar hem de kullanıcılar için hizmet güvenilirliği ve kullanılabilirlik sorunlarını gündeme getirmektedir. Yüksek performans, ölçeklenebilirlik, güvenilirlik ve verimlilik gibi avantajlar sunduğu halde, performansla ilgili sorunlar, özellikle hata toleransı gibi kritik unsurların etkin bir şekilde ele alınmasını zorunlu kılmaktadır. Bulut ortamında, arızalar ve performans düşüşleri, çeşitli hata türlerinin bir sonucu olarak ortaya çıkabilir [8,9].

Genellikle hata, arıza ve başarısızlık terimleri birbiriyle karıştırılarak kullanılır. Ancak, Hata Toleranslı Hesaplama terimlerinde, bu kavramların her birinin kendine özgü bir anlamı vardır. Fault, bir sistemde herhangi bir anda ortaya çıkabilen temel kusur ya da bozukluktur. Eğer bu arıza ele alınmazsa, sistemin başarısız olmasına yol açabilir. Error, bir arızanın sistemde gözlemlenebilir etkisidir. Failure, arızaların ve hataların ele alınamaması durumunda, sistemin beklenen işlevini yerine getirememesi durumudur. Bulut bilişim, dağıtılmış bilgi işlem yapısı nedeniyle kısmi arızalarla karakterize edilir. Bu arızalar, sistemin tamamen çökmesine veya performans düşüşlerine yol açabilir. Sistem bileşenlerinin, düğümlerin veya ağ elemanlarının herhangi birinde meydana gelen arızalar, sistemin çalışmasını tamamen durdurmaz, ancak kısmi bir aksama oluşturabilir. Bu durum, bulut sistemlerinin güvenilir ve sağlam olmasına rağmen, yüksek performanslı bilgi işlem için uygun arıza toleransı mekanizmalarının zorunlu hale gelmesini gerektirir. Hata toleransı (HT), bir sistemin arızalara rağmen beklenen işlevleri yerine getirme yeteneğini ifade eder ve sistemin, donanım veya yazılım bileşenlerinde oluşan arızalara, güç kesintilerine veya diğer olumsuzluklara karşı dayanıklı olmasını sağlar. Bu şekilde, sistemin arızalardan veya hatalardan etkilenmeden sürekli hizmet sunması sağlanır.

Literatürde, bulut bilişim ortamlarında hata toleransı sağlamak amacıyla birçok yöntem kullanılmıştır. Bu yöntemler arasında; kontrol noktası oluşturma [9], görev kopyalama, iş göçü, kendi kendini iyileştirme [10], güvenlik çantası kontrolleri, S-koruması, görevlerin aynı veya farklı kaynaklarda yeniden denenmesi ve yeniden gönderilmesi, zamanlama kontrolü, kurtarma iş akışı, yazılım yenileme, yeniden yapılandırma, sanal makineler (VM) için yük paylaşımı ve devralma [1] yer almaktadır. Ayrıca kötü arka uç algılama, yumuşak sistemler için kesin olmayan hesaplamaların bir sonraki aşamaya geçirilmesi, kodlama ve kod çözme gibi çeşitli teknikler de uygulanmaktadır. Hata ya da arıza algılama tekniklerinin evrimi, başlangıçta sistem performans kayıtlarının analizine dayanmıştır. Ancak bu yöntemlerin bazı dezavantajları ve veri madenciliğine dayalı algılama sınıflandırmasındaki verimsizlikler nedeniyle, istatistiksel öğrenme teorileri hata algılama amacıyla kullanılmaya başlanmıştır. Bu bağlamda, makine öğrenimi tabanlı tekniklerin kullanımı giderek yaygınlaşmıştır. Destek Vektör Makineleri (SVM), Destek Vektör Veri Açıklaması (SVDD), K-ortalama kümeleme, Bayes sinir ağları gibi yöntemler bu alanda yaygın olarak tercih edilmektedir [11]. Bu yöntemler, hata tespitinde daha yüksek doğruluk oranları ve sistemlere uyum sağlayabilme özellikleri nedeniyle öne çıkmaktadır.

Bu çalışmada, bulut bilişim altyapılarında yaygın olarak kullanılan üç farklı çok katmanlı ağ topolojisi olan Fat Tree, Z-Fat Tree ve RUFT-PL modelleri incelenecek ve karşılaştırmalı olarak analiz edilecektir. Söz konusu topolojiler, veri merkezlerinde yüksek bant genişliği, düşük gecikme, esneklik ve hata toleransı gibi temel gereksinimleri karşılamaya yönelik olarak tasarlanacaktır. Çalışmada, bu yapıların performansları, özellikle paket iletim başarısı, gecikme süreleri ve bağlantı arızalarına karşı dayanıklılık

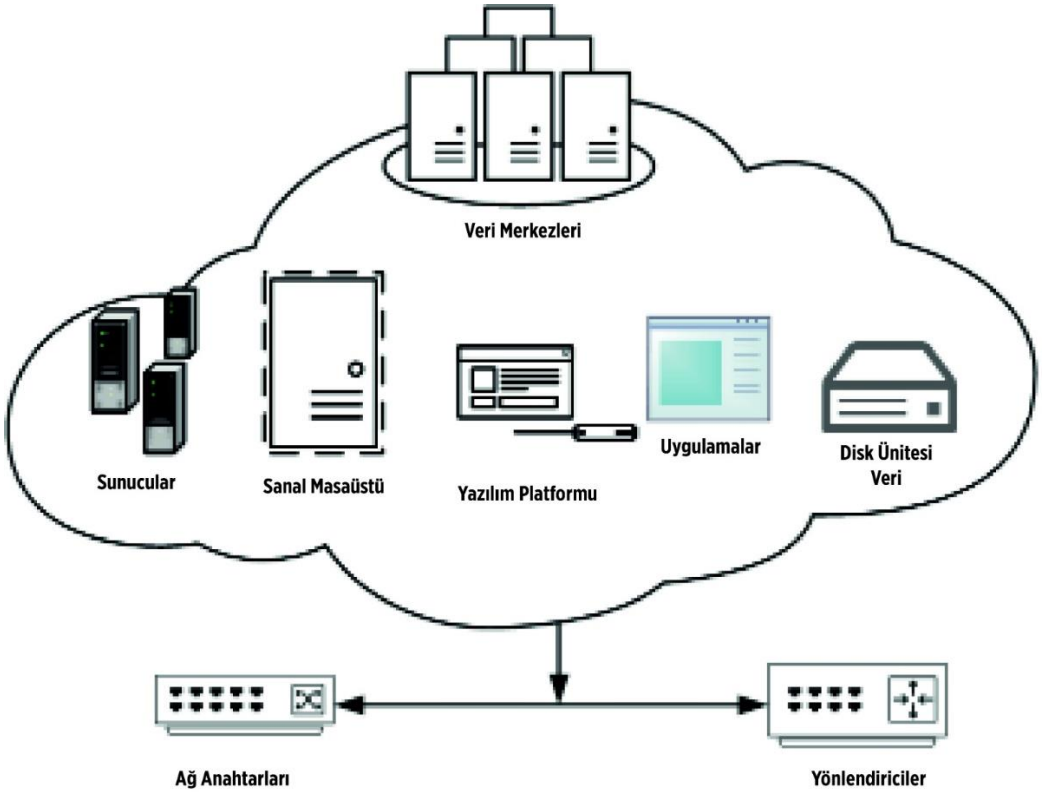
*Sorumlu Yazar/Corresponding Author: huseyingokmen@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 58 – 69, 2025
DOI: 10.62301/usmtd.1703696

kriterleri çerçevesinde değerlendirilecektir. Her bir topoloji için Python ortamında geliştirilen simülasyon senaryoları ile, belirli arıza oranlarında ağın başarımı test edilecek ve sonuçlar grafiksel olarak sunulacaktır. Elde edilen veriler, ağ mimarisi seçiminde topolojik yapıların sistem güvenilirliği ve performansı üzerindeki etkilerini ortaya koyacaktır.

2. Bulut Bilişim Modelleri (Cloud Computing Models)

Bulut bilişim donanım bileşenleri çoğunlukla kurumsal veri merkezlerinde bulunur. Bunlar, güvenlik duvarları, anahtarlar ve yönlendiriciler gibi kararlı depolama ve ağ aygıtları sunan çok çekirdekli sunucuları, katı hal sürücülerini ve sabit disk sürücüsünü kapsamaktadır. Şekil 1.'de bulut bilişim genel mimarisi görülmektedir.

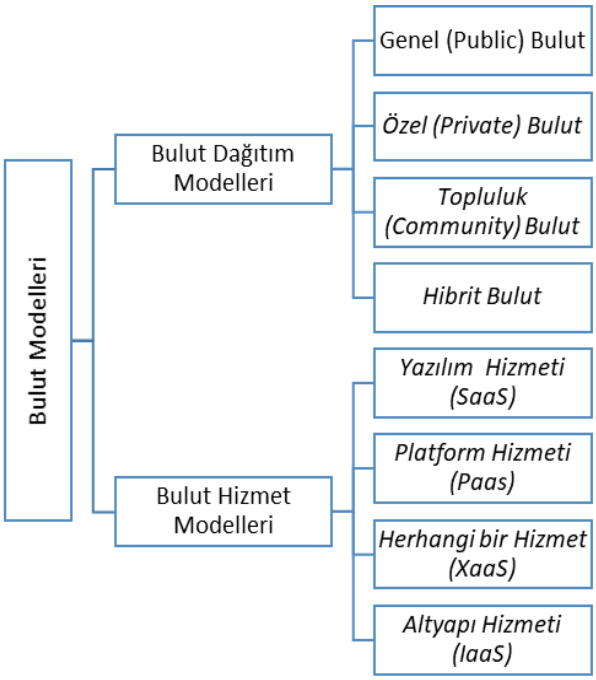


Şekil 1. Bulut Bilişim genel mimarisi

Bu donanım bileşenlerinin dışında, sanallaştırma yazılımı gibi bulut hizmeti modelini destekleyen yazılım bileşenleri de bulut bilişim altyapısı olarak adlandırılır. Sanallaştırma yazılımı, bulut kaynaklarının bir soyutlamasını sağlar ve bu kaynakları genellikle Uygulama Program Arayüzleri (API – Application Programming Interfaces) veya diğer komut satırı ve/veya grafik arayüzleri kullanarak kullanıcılara sunar. Bulut hizmet sağlayıcıları (CSP – Cloud Service Providers) tarafından barındırılan sanallaştırılmış kaynaklar, genellikle İnternet üzerinden kullanıcılara sunulur.

*Sorumlu Yazar/Corresponding Author: huseyingokmen@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 58 – 69, 2025
DOI: 10.62301/usmtd.1703696



Şekil 2. Bulut Bilişim Modelleri

Genel, Özel, Topluluk ve Hibrit Bulut olmak üzere dört temel bulut bilişim dağıtım modeli bulunmaktadır [12]. Genel Bulut, altyapı ve yazılım kaynaklarının internet üzerinden üçüncü parti sağlayıcılar (örneğin AWS, Azure, GCP) aracılığıyla sunulduğu, çok kullanıcılı ve maliyet etkin bir modeldir. Kaynaklar ortak kullanılarak kullanıcılar yalnızca tükettikleri hizmet kadar ödeme yapmaktadırlar [13]. Özel Bulut yapıları, yalnızca bir organizasyona tahsis edilen, genellikle kurum içi veri merkezlerinde ya da özel altyapılarda barındırılan bir modeldir [14]. Daha fazla kontrol, güvenlik sağlamaktadır. Bu durum finansal kurumlar, devlet kurumları gibi yüksek güvenlik gerektiren yapılar için tercih edilmektedir. Ancak, kurulum ve bakım maliyetleri yüksektir. Topluluk Bulut, benzer regülasyonlara, güvenlik ve işlem gereksinimlerine sahip organizasyonlar arasında paylaşılan bir model olup, sağlık, eğitim ve kamu gibi sektörlerde maliyetleri paylaşarak güvenli bir ortak ortam sunmaktadır [15]. Son olarak, Hibrit Bulut modeli, özel ve genel bulut yapılarının birlikte kullanıldığı esnek bir çözümdür. Kritik veriler özel bulutta saklanırken, daha az hassas işlemler için genel bulut kaynakları kullanılır [16]. Bulut bilişim hizmet modelleri, kuruluşların dijital altyapı ve yazılım ihtiyaçlarını esnek, ölçeklenebilir ve maliyet etkin bir şekilde karşılamak amacıyla farklı katmanlarda sunulmaktadır.

Bulut bilişim hizmet modelleri, organizasyonların dijital altyapı ve yazılım ihtiyaçlarını esnek, ölçeklenebilir ve maliyet etkin bir şekilde karşılamak amacıyla farklı katmanlarda sunulmaktadır. Altyapı Hizmeti Olarak Sunulan Hizmet (IaaS) modeli, kullanıcıların fiziksel donanım yatırımı yapmadan sanal sunuculara, veri depolama alanlarına ve ağ kaynaklarına internet üzerinden erişmesini sağlamaktadır [17]. Bu modelde, donanımın yönetimi hizmet sağlayıcısı tarafından gerçekleştirilirken, kullanıcılar sistemler üzerinde geniş kontrol imkanına sahip olur. Platform Hizmeti Olarak Sunulan Hizmet (PaaS) ise altyapı yönetimini tamamen soyutlayarak yazılım geliştiricilere uygulama geliştirme, test etme ve dağıtma süreçlerini yönetebilecekleri bütünleşik bir ortam sunmaktadır [18]. Yazılım Hizmeti Olarak Sunulan Hizmet (SaaS) modeli ise son kullanıcıya, herhangi bir kurulum yapmaya gerek kalmadan doğrudan internet üzerinden erişilebilen yazılım çözümleri sunmaktadır [19]. Yazılımların tüm bakım, güncelleme, yedekleme gibi teknik işlemler hizmet sağlayıcısı tarafından yürütülmektedir. Herhangi bir Hizmet Olarak Sunulması (XaaS) yaklaşımı, güvenlik, analitik, veri yönetimi, ağ servisleri gibi çok daha geniş bir bilgi teknolojisi kaynaklarının servis olarak sunulmasını mümkün kılmaktadır [20].

3. Bulut Bilişim Ağ Topolojileri (Cloud Computing Network Topologies)

Bulut bilişim ağ topolojileri, veri merkezleri ve dağıtık sistemler arasında yüksek bant genişliği, düşük gecikme ve ölçeklenebilirlik sağlayarak, bulut hizmetlerinin güvenilir ve verimli bir şekilde sunulmasını mümkün kılan mimari yapılardır [21]. Bu bağlamda, farklı performans ihtiyaçlarına ve sistem

*Sorumlu Yazar/Corresponding Author: huseyingokmen@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 58 – 69, 2025
DOI: 10.62301/usmtd.1703696

gereksinimlerine göre çeşitli topoloji modelleri geliştirilmiştir. Bu modeller arasında donanım seviyesinde yüksek kapasiteli anahtarların omurgayı oluşturduğu anahtar (switch) merkezli topolojiler (örneğin Spine-Leaf [22], Fat-Tree [23] ve RUFT-PL [24]) ile her sunucunun birden çok ağ arabirimi üzerinden paket iletimi ve yönlendirme işlevine katıldığı sunucu-merkezli topolojiler (BCube [25], DCell [26], Jellyfish [27] vb.) öne çıkar. Switch-merkezli mimariler, ASIC tabanlı switch donanımları sayesinde uçtan uca düşük gecikme ve yüksek bant genişliği sunmaktadır. Sunucu-merkezli yaklaşımlar her bir sunucuyu hem hesaplama hem de yönlendirme düğümü haline getirerek ağın ölçeklenebilirliğini ve hata toleransını artırmaktadır. Her iki model de omurga-yaprak (spine-leaf) veya çok katmanlı Clos benzeri düzenler temelinde şekillenir. Böylece artan trafik yükü ve dağıtık uygulama gereksinimleri karşılanırken, kaynak kullanımı ve yedeklilik dengesi optimize edilir. Fat Tree, Z-Fat Tree, RUFT, RUFT-PL, VL2, DCell, BCube ve Clos gibi topolojiler [28], ağ yapısının esnekliği, hata toleransı, maliyet etkinliği ve trafik yönetimi gibi farklı avantajlar sunmaktadır. Bu topolojiler, özellikle bulut bilişim ortamlarında büyük veri trafiğini dengelemek, ağ darboğazlarını önlemek ve hizmet sürekliliğini sağlamak amacıyla tercih edilmektedir. Bu modeller, hem klasik hem de gelişmiş yapıları temsil ederek bulut bilişim altyapılarının güçlü ve zayıf yönlerinin analizine olanak sağlamaktadır.

Bulut bilişim ağ topolojileri, genellikle çok katmanlı bir mimari yapı üzerine kuruludur. Bu yapı veri merkezlerinde ağ trafiğinin verimli, güvenli ve ölçeklenebilir bir şekilde yönetilmesini sağlar. Özellikle switch merkezli topoloji olan Fat Tree, Z-Fat Tree ve RUFT-PL gibi topolojilerde yaygın olarak kullanılan bu yapı, Uç (edge) katmanı, Toplama (aggregation) katmanı ve Çekirdek (core) katmanı olmak üzere üç temel katmandan oluşur [28]. Edge katmanı, ağın en alt seviyesinde yer alır ve sunucuların doğrudan bağlandığı anahtarlardan oluşur. Yerel veri trafiği bu katmanda başlar ve üst katmanlara iletilir. Toplama katmanı, uç ile çekirdek katmanları arasında köprü görevi görür ve verilerin yönlendirilmesi, filtrelenmesi gibi işlemleri gerçekleştirir. En üstte yer alan core katmanı ise ağın omurgasını oluşturur. Veri merkezi bölümleri arasında yüksek bant genişliğine sahip veri iletimini sağlamaktadır. Bu katmanlı yapı, ağın modülerliğini ve ölçeklenebilirliğini artırırken, aynı zamanda yük dengeleme ve yedeklilik gibi özelliklerle ağın güvenilirliğini güçlendirir. Bulut bilişim altyapılarında binlerce sunucunun etkin ve esnek biçimde birbirine bağlanabilmesi için bu hiyerarşik yapı kritik bir rol oynamaktadır [28].

3.1. Fat tree topolojisi (Fat tree topology)

Fat Tree topolojisi, veri merkezleri ve büyük ölçekli ağ altyapıları için tasarlanmış, çok katmanlı, simetrik ve ölçeklenebilir bir ağ mimarisidir. Bu yapı, geleneksel ağaç (tree) topolojisinin tıkanma (congestion) sorununu çözmek amacıyla geliştirilmiştir. [29]

Fat Tree topolojisi, her biri k portlu anahtarlardan (switch) oluşan simetrik ve çok katmanlı bir ağ mimarisi olarak tanımlanır. k değeri, topolojinin boyutunu belirleyen temel parametredir. Verilen $k \in N$ (tek sayı olmamak üzere), aşağıdaki sayısal özellikler elde edilir;

Her biri $\frac{k}{2}$ edge ve $\frac{k}{2}$ aggregation anahtarı içeren k adet pod bulunur. Toplam edge ve aggregation anahtar sayısı Denklem (1)' e göre tespit edilir.

$$S_{edge} = S_{agg} = kx\frac{k}{2} = \frac{k^2}{2} \quad (1)$$

Toplam çekirdek anahtar sayısı Denklem (2)' ye göre;

$$S_{core} = \left(\frac{k}{2}\right)^2 \quad (2)$$

Toplam sunucu sayısı Denklem (3)' e göre;

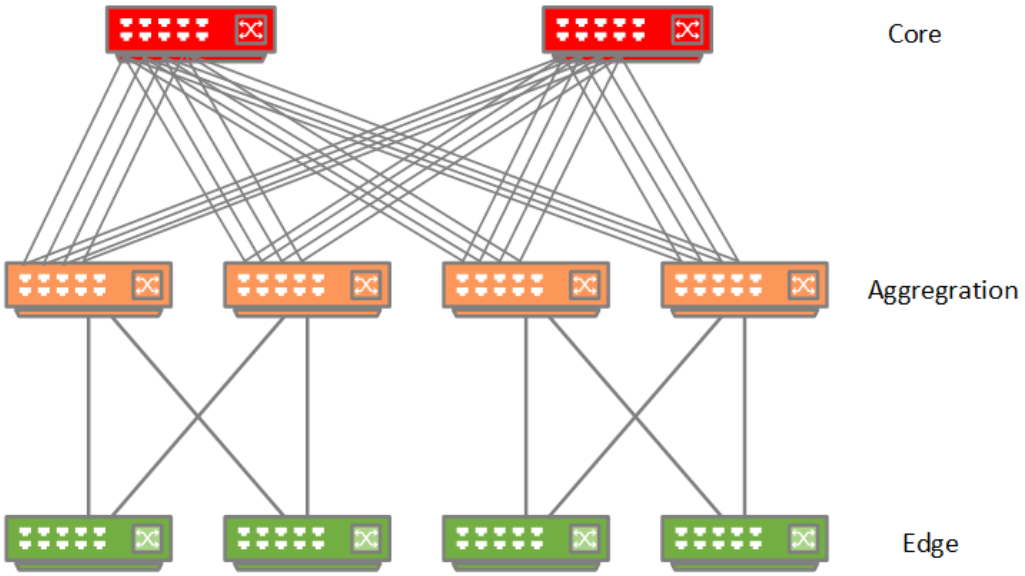
$$N_{host} = S_{edge}.\frac{k}{2} = \frac{k^2}{2}x\frac{k}{2} = \frac{k^3}{2} \quad (3)$$

Elde edilir.

*Sorumlu Yazar/Corresponding Author: huseyingokmen@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 58 – 69, 2025
DOI: 10.62301/usmtd.1703696

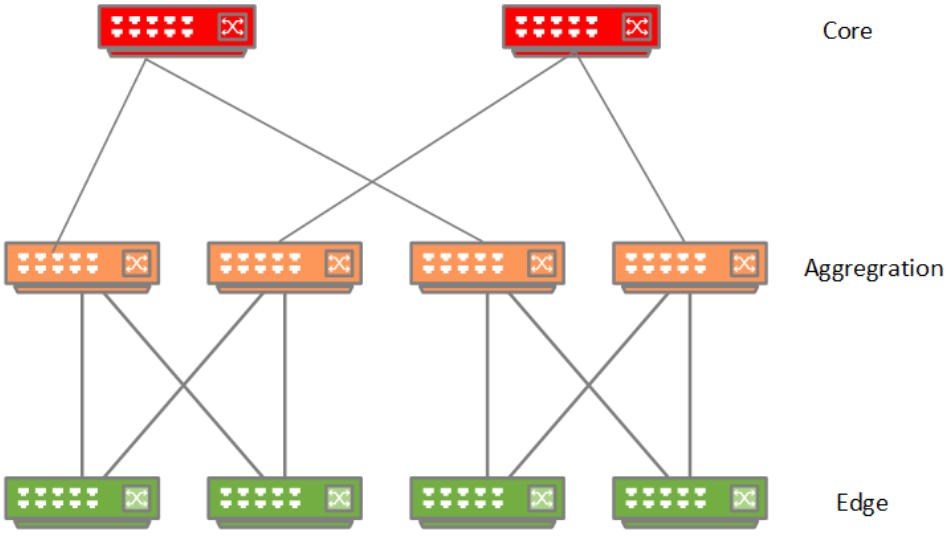
Şekil 3'te k=2 ile elde edilen Fat Tree topolojisi görülmektedir.



Şekil 3. $k = 2$ için Fat Tree Topolojisi

3.2. Z- Fat tree topolojisi (Z- fat tree topology)

Z-Fat Tree topolojisi, klasik Fat Tree mimarisinin bağlantı yoğunluğunu azaltmak ve sistem kaynaklarını daha verimli kullanmak amacıyla geliştirilmiştir. Desen tabanlı (pattern-based) bir ağ mimarisidir. Fat Tree' den farklı olarak edge ile aggregation katmanları arasındaki bağlantılar, belirli bağlantı desenlerine göre sınırlandırılmıştır. Bu sayede ağdaki toplam bağlantı sayısı azaltılmıştır. Bu nedenle donanım maliyetleri ve enerji tüketimi optimize edilmiştir. Full, Cross, Diagonal ve Upper/Lower üçgen Z – Fat Tree topolojisne ait desenlerdir. Full deseninde edge ve aggregation kombinasyonlarının tümü bağlanır. Cross desende yalnızca $(i + j)\ mod\ 2\ = 0$ koşulunu sağlayan edge i ve aggregation j anahtarları bağlanır. Diagonal desende ise sadece $i = j$ veya $i + j = k - 1$ olan çiftler bağlanır[30]. Upper/Lower üçgen desenlerinde ise edge anahtarları sadece üst/alt üçgen bölgedeki aggregation anahtarları ile bağlantılıdır. Şekil 4'te k=2 ile elde edilen Full desenli Z-Fat Tree topolojisi görülmektedir.



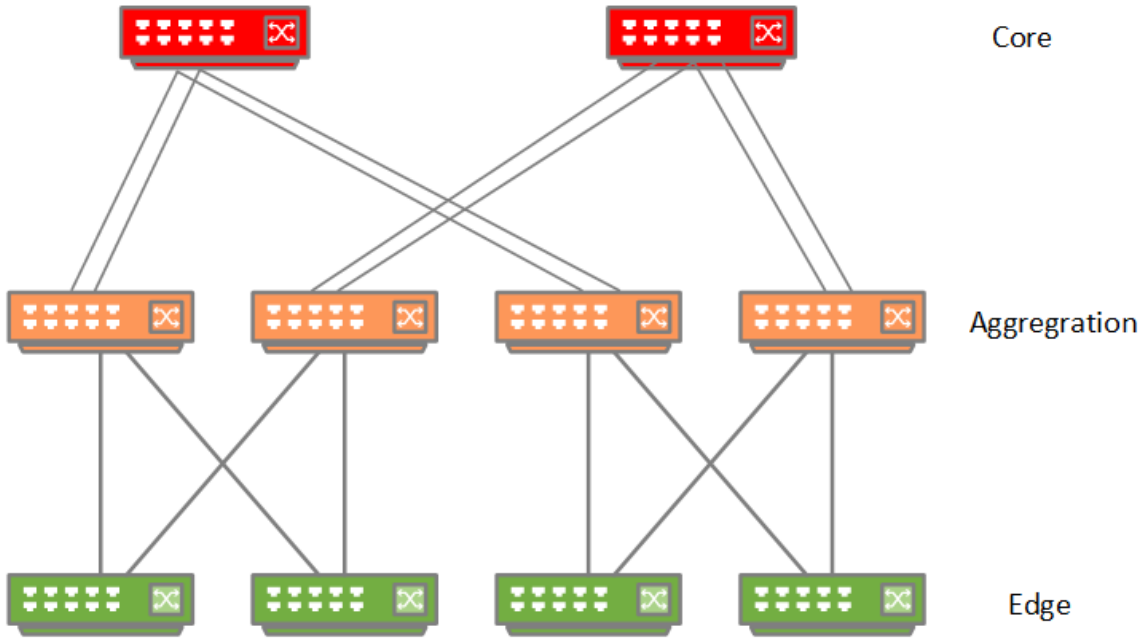
Şekil 4. $k = 2$ için Z-Fat Tree Cross Topolojisi

*Sorumlu Yazar/Corresponding Author: huseyingokmen@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 58 – 69, 2025
DOI: 10.62301/usmtd.1703696

3.3. RUFT-PL topolojisi (RUFT-PL topology)

RUFT-PL (Reduced Unidirectional Fat Tree with Parallel Links) topolojisi, klasik Fat Tree mimarisinin basitleştirilmiş ve iyileştirilmiş bir türevidir. "Reduced" ifadesi, gereksiz bağlantıların ortadan kaldırılmasını; "Unidirectional", yönlendirilmiş (tek yönlü) veri akışını; "Parallel Links" ise iki düğüm arasında birden fazla bağlantının (paralel linkin) kullanılmasını ifade etmektedir. Edge katmanındaki anahtarlar, doğrudan aggregation katmanındaki anahtarlara ve core anahtarlarına bağlanmaktadır. Her bir edge–aggregation veya aggregation–core bağlantısı, birden fazla paralel fiziksel hatla desteklenmektedir. Bu yapı sayesinde, aynı hedefe birden fazla alternatif yol oluşmaktadır. Bu alternatif yollar ağ tıkanıklıkları önlemektedir. Aynı zamanda bağlantı arızalarına karşı dayanıklılığı artırmaktadır. [31] Şekil 5'te k=2 ile elde edilen Full desenli RUFT PL topolojisi görülmektedir



Şekil 5. $k = 2$ için RUFT-PL Cross Topolojisi

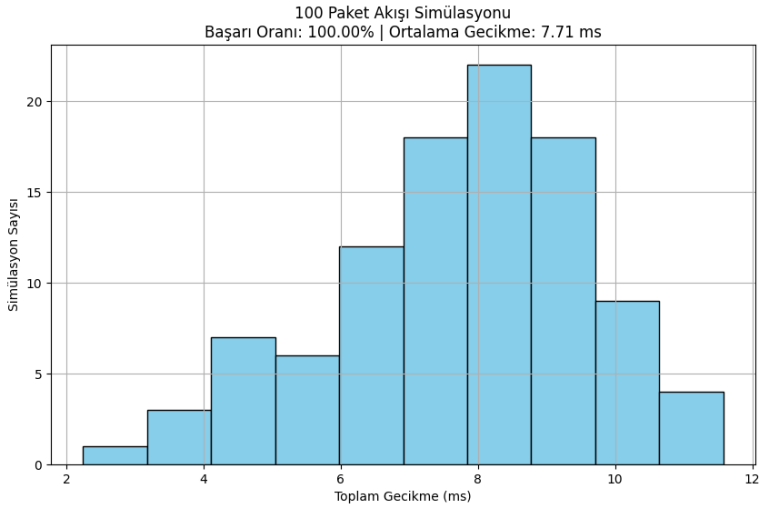
4. Araştırma Bulguları (Research Findings)

Bu çalışmada, Fat Tree, Z-Fat Tree ve RUFT-PL topolojilerinin ağ performansı ve hata toleransı açısından karşılaştırmalı analizi yapılmıştır. Simülasyon süreci Python programlama dili kullanılarak gerçekleştirilmiş ve ağ topolojileri NetworkX [32] kütüphanesi aracılığıyla dinamik olarak modellenmiştir. Her topoloji için belirli bir k=8 parametresi kullanılarak core, aggregation ve edge katmanları ile sunucular oluşturulmuş; bağlantılar ise ilgili topolojinin bağlantı kurallarına göre otomatik olarak atanmıştır. Paket akışı simülasyonu kapsamında, edge katmanından rastgele seçilen bir kaynak ve hedef düğüm arasında en kısa yol belirlenmiş ve bu yol üzerinden iletimin toplam gecikme süresi hesaplanmıştır. Ayrıca, ağ bağlantılarının belirli bir oranı (%0–%60 arası) rastgele devre dışı bırakılarak her topoloji için link arızası senaryoları uygulanmış ve bu durumda paket iletiminin başarı durumu ölçülmüştür. Her senaryo için simülasyon 100 veya 1000 tekrar ile çalıştırılmış, böylece gecikme dağılımı, başarı oranı ve ağ dayanıklılığı istatistiksel olarak değerlendirilmiştir.

Simülasyon sonuçlarına göre, her üç topolojide de paketlerin hedefe başarıyla ulaştığı ve başarı oranının %100 olduğu gözlemlenmiştir. Ancak, bu benzer başarı oranlarına rağmen ağların toplam gecikme süreleri ve dağılım yapıları önemli farklılıklar göstermektedir.

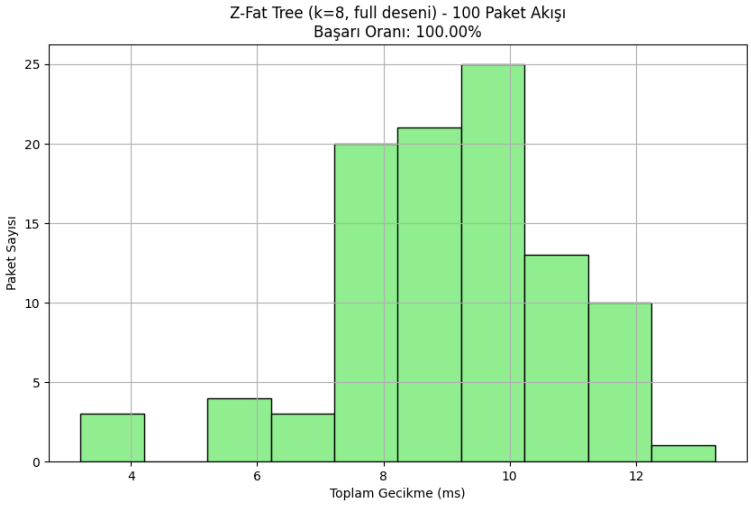
*Sorumlu Yazar/Corresponding Author: huseyingokmen@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 58 – 69, 2025
DOI: 10.62301/usmtd.1703696



Şekil 6. Fat-Tree topolojisine göre gecikme dağılımı

Şekil 6'ya göre, Fat Tree topolojisinde, gecikme süreleri oldukça dengeli bir dağılım sergilemiştir. Ortalama gecikme 7.71 ms olarak ölçülmüş, simülasyonların büyük çoğunluğu 6–9 ms aralığında yoğunlaşmıştır. Bu dağılım, Fat Tree'nin düzenli ve eşit sayıda bağlantıya sahip mimarisinin, veri iletiminde tutarlı ve kararlı bir performans sunduğunu göstermektedir.

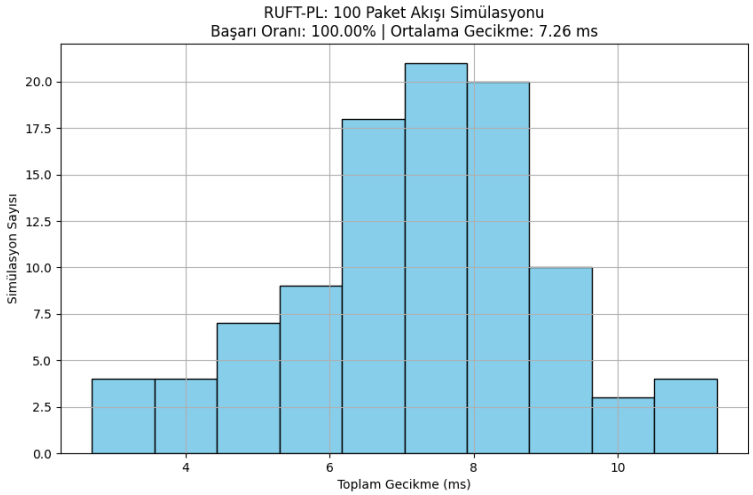


Şekil 7. Z-Fat-Tree topolojisine göre gecikme dağılımı

Şekil 7'ye göre, Z-Fat Tree (full bağlantı desenli) topolojide ise gecikme süreleri biraz daha geniş bir aralıkta dağılmıştır. Başarı oranı yine %100 olmasına rağmen, gecikme sürelerinin sağa doğru kaydığı, yani bazı paketlerin daha yüksek gecikmelere maruz kaldığı görülmektedir. Bu durum, bağlantı sayısının fazla olmasının sağladığı yedekliliğe rağmen, bazı rotaların daha uzun olabileceğini ve bu nedenle gecikme sürelerinin artabileceğini ortaya koymaktadır. Bu yapı, performans açısından esneklik sağlasa da, gecikme açısından daha az öngörülebilir bir karakteristik sergilemektedir.

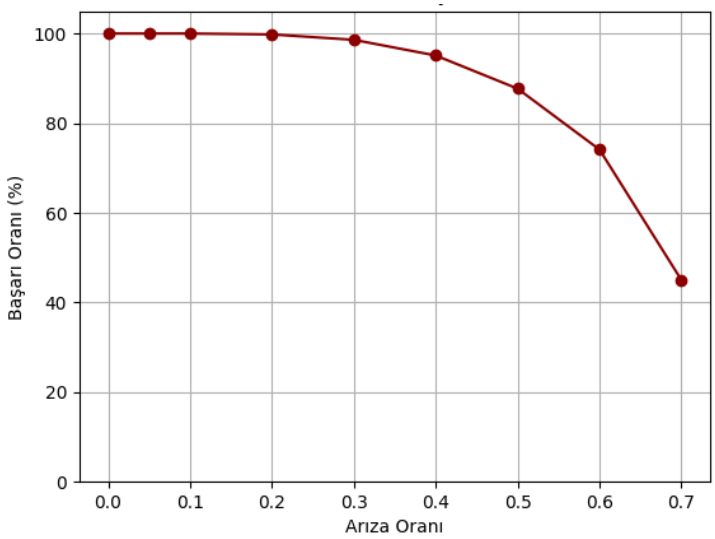
*Sorumlu Yazar/Corresponding Author: huseyingokmen@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 58 – 69, 2025
DOI: 10.62301/usmtd.1703696



Şekil 8. RUFT-PL topolojisine göre gecikme dağılımı

Şekil 8'e göre, RUFT-PL topolojisi, en düşük ortalama gecikmeye (7.26 ms) sahip olmasıyla öne çıkmaktadır. Yapısında bulunan paralel bağlantılar, verinin farklı ve daha kısa yollar üzerinden aktarılabilmesine olanak tanımakta ve bu sayede gecikme sürelerinin düşmesini sağlamaktadır. Gecikme dağılımı, Fat Tree'ye benzer şekilde simetrik ve kararlı bir görünüm sunarken, gelişmiş bağlantı yedekliliği sayesinde hem düşük gecikme hem de yüksek güvenilirlik sağlamaktadır.

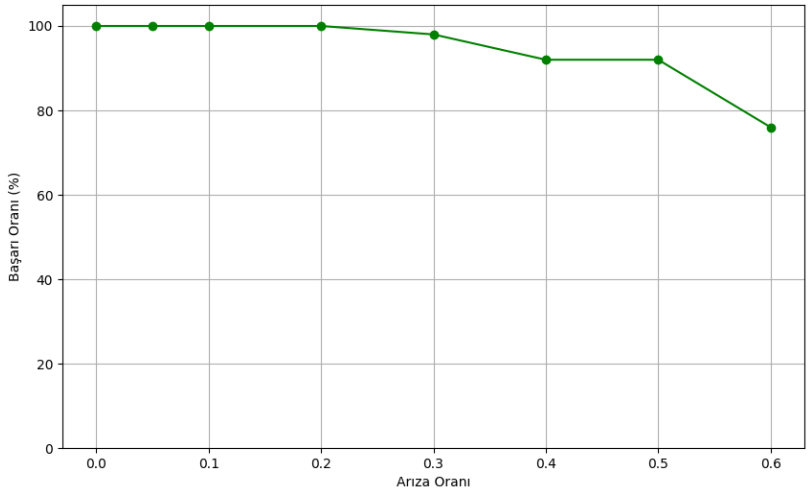


Şekil 9. Fat-Tree topolojisine göre bağlantı arızası ve iletilen paketlerin başarı oranı

Şekil 9'a göre, Fat Tree topolojisi, %0 ile %30 arası arıza oranlarında yüksek başarı oranını korumaktadır. Ancak %40 seviyesinden itibaren başarının hızla düşmeye başladığı, %70 arıza oranında ise başarı oranının %45'in altına indiği görülmektedir. Bu durum, Fat Tree mimarisinin belirli bir yedek bağlantı kapasitesine sahip olduğunu, fakat bu kapasitenin ötesinde kritik bağlantıların kaybıyla iletişimin sekteye uğradığını göstermektedir. Başka bir deyişle, Fat Tree yapısı sınırlı arıza toleransına sahiptir ve yüksek oranda bağlantı kaybı sistemin genel performansını ciddi şekilde etkileyebilmektedir.

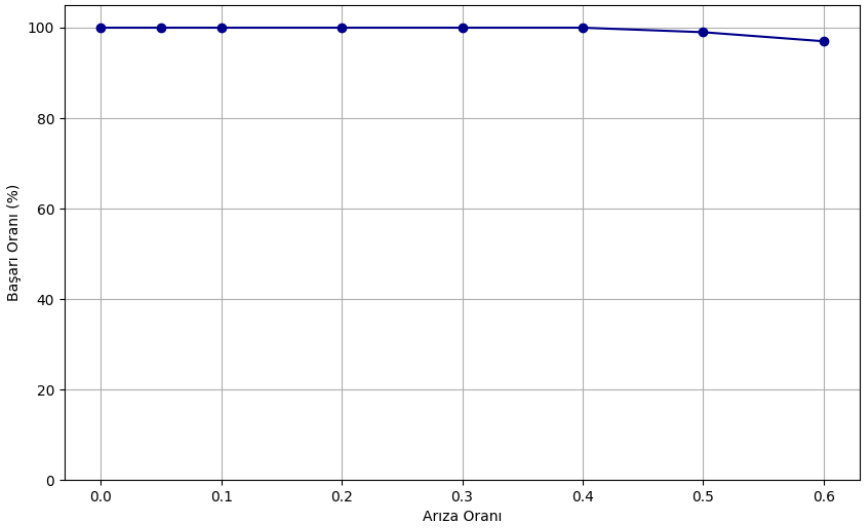
*Sorumlu Yazar/Corresponding Author: huseyingokmen@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 58 – 69, 2025
DOI: 10.62301/usmtd.1703696



Şekil 10. Z-Fat-Tree topolojisine göre bağlantı arızası ve iletilen paketlerin başarı oranı

Şekil 10'a göre, Z-Fat Tree topolojisi, full bağlantı deseni kullanılarak tasarlanmış olmasına rağmen, arıza oranı %40'a ulaştığında başarı oranında belirgin bir düşüş göstermektedir. İlk %30'luk arıza oranında sistem oldukça kararlı şekilde çalışmaya devam ederken, %40 ve üzeri oranlarda paket iletiminin sekteye uğradığı anlaşılmaktadır. Z-Fat Tree yapısının daha az bağlantıyla daha hafif bir ağ yapısı sunması, sistem kaynakları açısından avantaj oluştursa da, bu durum aynı zamanda yüksek arıza oranlarında performans kaybına daha açık hale gelmesine neden olmaktadır.



Şekil 11. RUFT-PL topolojisine göre bağlantı arızası ve iletilen paketlerin başarı oranı

Şekil 11'e göre, RUFT-PL topolojisi ise test edilen üç yapı arasında en yüksek dayanıklılığı sergilemiştir. %60 arıza oranına kadar başarı oranı %95'in üzerinde korunmuştur. Bu yapı, edge ve aggregation katmanları ile aggregation ve core katmanları arasında birden fazla paralel bağlantı içerdiğinden, arızalı bağlantılar aktif iletişimi büyük ölçüde etkilememektedir. Bu durum, RUFT-PL topolojisini özellikle yüksek güvenilirlik, yedeklilik ve servis sürekliliği gerektiren sistemler için son derece uygun kılmaktadır. Ağın önemli bir kısmı devre dışı kalsa dahi, alternatif yollar sayesinde iletişim devam edebilmektedir.

5. Tartışma ve Sonuç (Discussion and Conclusion)

Bu çalışmada, üç farklı veri merkezi ağ topolojisi olan Fat Tree, Z-Fat Tree (cross bağlantı desenli) ve RUFT-PL yapıları, paket iletimi başarısı, gecikme performansı ve bağlantı arızalarına karşı dayanıklılık açısından karşılaştırılmıştır. Gerçekleştirilen simülasyonlar sonucunda, tüm topolojilerin düşük arıza oranlarında yüksek başarı oranı sunduğu, ancak arıza oranı yükseldikçe bu başarı oranının topolojinin yapısal özelliklerine bağlı olarak farklılaştığı gözlemlenmiştir. Fat Tree topolojisi düzenli yapısıyla

*Sorumlu Yazar/Corresponding Author: huseyingokmen@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 58 – 69, 2025
DOI: 10.62301/usmtd.1703696

dengeli bir performans sunarken, Z-Fat Tree özellikle belirli bağlantı desenleriyle gecikme açısından daha değişken sonuçlar üretmiştir. RUFT-PL topolojisi ise paralel bağlantıların sağladığı yedeklilik sayesinde hem düşük gecikme süreleri hem de yüksek arıza toleransı ile en iyi genel performansı göstermiştir.

Literatüre katkı açısından, bu çalışma üç farklı topolojiyi hem başarı oranı hem de gecikme dağılımı açısından karşılaştırmalı olarak analiz ederek, arıza toleransı perspektifinden farklarını sistematik biçimde ortaya koymaktadır. Önceki çalışmaların çoğunda bu tür karşılaştırmalar yalnızca bant genişliği veya teorik kapasite üzerinden yapılırken, bu çalışma bağlantı arızaları gibi gerçek dünya koşullarını simüle ederek daha kapsamlı bir değerlendirme sunmaktadır.

Elde edilen bulgular, ağ topolojisi seçiminde yalnızca nominal performans değil, aynı zamanda beklenmedik bağlantı kayıplarına karşı sistemin tepkisinin de dikkate alınması gerektiğini ortaya koymaktadır. Kritik uygulamalar ve servis sürekliliği gerektiren sistemlerde, RUFT-PL gibi yüksek esneklik ve hata toleransı sunan yapılara yönelmek, veri iletiminin güvenliğini ve sistemin sürdürülebilirliğini artıracaktır. Gelecek çalışmalarda, farklı anahtar port değerlerinin (örn. 12, 16, 20) performansa etkisi incelenerek ölçeklenebilirlik analizleri yapılabilir. Ayrıca, bu çalışmanın kapsamını genişletmek amacıyla farklı topolojilerin de deneysel analizlere dâhil edilmesi planlanmakta; böylece daha geniş bir ağ mimarisi havuzunda performans, gecikme ve hata toleransı kriterlerinin nasıl değiştiği sistematik olarak değerlendirilebilecektir.

6. Kaynaklar (References)

[1] R. Buyya, C.S. Yeo, S. Venugopal, J. Broberg, I. Brandic, Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility, Future Gener. Comput. Syst. 25 (6) (2009) 599–616.

[2] P. Kumari, P. Kaur, A survey of fault tolerance in cloud computing, J. King Saud Univ. Comput. Inf. Sci. 33 (10) (2021) 1159–1176.

[3] P.M. Mell, T. Grance, The NIST definition of cloud computing, Recommendations of the National Institute of Standards and Technology (2011).

[4] S. Patidar, D. Rane, P. Jain, A survey paper on cloud computing, Proc. 2012 2nd Int. Conf. Adv. Comput. Commun. Technol. (ACCT) (2012) 394–398.

[5] B. Varghese, R. Buyya, Next Generation Cloud Computing: New Trends and Research Directions, arXiv preprint arXiv:1707.07452 (2017). Available from: http://arxiv.org/abs/1707.07452.

[6] M.K. Gokhroo, M.C. Govil, E.S. Pilli, Detecting and Mitigating Faults in Cloud Computing Environment, Proc. 3rd Int. Conf. Comput. Intell. Commun. Technol. (CICT), IEEE (2017).

[7] L.P. Saikia, Y.L. Devi, Fault Tolerance Techniques and Algorithms in Cloud Computing, Int. J. Comput. Sci. Commun. Netw. 4 (1) (2014) 1–8.

[8] Z. Amin, N. Sethi, H. Singh, Review on Fault Tolerance Techniques in Cloud Computing, Int. J. Comput. Appl. 116 (18) (2015).

[9] M. Nazari Cheraghlou, A. Khadem-Zadeh, M. Haghparast, A survey of fault tolerance architecture in cloud computing, J. Netw. Comput. Appl. 61 (2016) 81–92.

[10] Z. Amin, N. Sethi, H. Singh, Review on Fault Tolerance Techniques in Cloud Computing, Int. J. Comput. Appl. 116 (18) (2015).

[11] X. Gu, H. Wang, Online anomaly prediction for robust cluster systems, Proc. Int. Conf. Data Eng. (2009) 1000–1011.

[12] I. Odun-Ayo, R. Goddy-Worlu, V. Samuel, V. Geteloma, Cloud designs and deployment models: A systematic mapping study, BMC Res. Notes 12 (1) (2019).

*Sorumlu Yazar/Corresponding Author: huseyingokmen@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 58 – 69, 2025
DOI: 10.62301/usmtd.1703696

[13] R. Gohil, H. Patel, Comparative Analysis of Cloud Platform: Amazon Web Service, Microsoft Azure, and Google Cloud Provider: A Review, Proc. 15th Int. Conf. Comput. Commun. Netw. Technol. (ICCCNT) (2024).

[14] E.N. Ekwonwune, C.D. Anyiam, O.E. Osuagwu, Modelling Conceptual Framework for Private Cloud Infrastructure Deployment in the ICT Centre of Tertiary Institutions, Commun. Netw. 10 (3) (2018) 117–125.

[15] A. Marinos, G. Briscoe, Community Cloud Computing (2009).

[16] V. Khadilkar, M. Kantarcioglu, B. Thuraisingham, S. Mehrotra, Secure Data Processing in a Hybrid Cloud, arXiv preprint arXiv:1105.1982 (2011). Available from: http://arxiv.org/abs/1105.1982.

[17] S.H.H. Madni, M.S.A. Latiff, Y. Coulibaly, S.M. Abdulhamid, Resource scheduling for infrastructure as a service (IaaS) in cloud computing: Challenges and opportunities, J. Netw. Comput. Appl. 68 (2016) 173–200.

[18] O. Krancher, P. Luther, M. Jost, Key Affordances of Platform-as-a-Service, J. Manag. Inf. Syst. 35 (3) (2018) 776–812.

[19] P.K. Chouhan, F. Yao, S.Y. Yerima, S. Sezer, Software as a Service: Analyzing Security Issues (2014).

[20] M. Howard, Cloud Computing – Everything As A Service, arXiv preprint arXiv:2206.07094 (2022). Available from: http://arxiv.org/abs/2206.07094.

[21] B. Wang, Z. Qi, R. Ma, H. Guan, A.V. Vasilakos, A survey on data center networking for cloud computing, Comput. Netw. 91 (2015) 528–547.

[22] A. Singh, J. Ong, A. Agarwal, G. Anderson, A. Armistead, R. Bannon, S. Boving, G. Desai, B. Felderman, P. Germano, A. Kanagala, J. Provost, J. Simmons, E. Tanda, J. Wanderer, U. Hölzle, S. Stuart, A. Vahdat, Jupiter Rising: A Decade of Clos Topologies and Centralized Control in Google's Datacenter Network, Proc. ACM SIGCOMM (2015) 183–197.

[23] M. Al-Fares, A. Loukissas, A. Vahdat, A scalable commodity data center network architecture, Comput. Commun. Rev., ACM (2008).

[24] X. Li, S. Kumar, RUFT-PL: A Fault-Tolerant Fat-Tree with Parallel Links, Proc. IEEE ICCCN (2013) 1–7.

[25] C. Guo, G. Lu, D. Li, H. Wu, X. Zhang, Y. Shi, C. Tian, Y. Zhang, S. Lu, BCube: A high performance, server-centric network architecture for modular data centers, SIGCOMM Comput. Commun. Rev. 39 (4) (2009) 63–74.

[26] C. Guo, H. Wu, K. Tan, L. Shi, Y. Zhang, S. Lu, DCell: A scalable and fault-tolerant network structure for data centers, SIGCOMM Comput. Commun. Rev. 38 (4) (2008) 75–86.

[27] A. Singla, C.-Y. Hong, L. Popa, P.B. Godfrey, Jellyfish: Networking Data Centers Randomly (2012).

[28] B. Lebiednik, A. Mangal, N. Tiwari, A Survey and Evaluation of Data Center Network Topologies, arXiv preprint arXiv:1605.01701 (2016). Available from: http://arxiv.org/abs/1605.01701.

[29] H. Emesowum, A. Paraskelidis, M. Adda, Fault tolerance capability of cloud data center, Proc. ICCP (2017) 495–502. https://doi.org/10.1109/iccp.2017.8117053

[30] H. Emesowum, A. Paraskelidis, M. Adda, Management of fault tolerance and traffic congestion in cloud data center, Proc. ICOIACT (2018) 10–15. https://doi.org/10.1109/icoiact.2018.8350692

[31] D.F. Bermúdez Garzón, C.G. Requena, M.E. Gómez, P. López, J. Duato, A family of fault-tolerant efficient indirect topologies, IEEE Trans. Parallel Distrib. Syst. 27 (4) (2016) 927–940. https://doi.org/10.1109/TPDS.2015.2430863

[32] J. Calle-Cancho, C. Cruz-Carrasco, D. Cortés-Polo, J. Galeano-Brajones, J. Carmona-Murillo, Enhancing Programmability in Next-Generation Networks: An Innovative Simulation Approach, Electron. (Switzerland) 13 (3) (2024).

*Sorumlu Yazar/Corresponding Author: huseyingokmen@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 70 – 79, 2025
DOI: 10.62301/usmtd.1702079

Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi
International Journal of Sustainable Engineering and Technology
ISSN: 2618-6055 / Cilt:9, Sayı:1, (2025), 70 – 79
Araştırma Makalesi – Research Article

TARIMSAL VERİMLİLİĞİ ARTIRMAK İÇİN AGRO-KLİMATİK KOŞULLARA BAĞLI OLARAK MAKİNE ÖĞRENMESİ TEMELLİ MAHSUL TAVSİYE SİSTEMİ

*Onur SEVLİ[1]

[1]Burdur Mehmet Akif Ersoy Üniversitesi, Mühendislik-Mimarlık Fakültesi, Bilgisayar Mühendisliği Burdur

(Geliş/Received: 19.05.2025, Kabul/Accepted: 17.06.2025, Yayınlanma/Published: 30.06.2025)

ÖZ

Tarım, dünya genelinde ekonomik ve toplumsal açıdan stratejik bir sektör olup, özellikle gıda güvenliği ve kırsal kalkınma açısından kritik bir role sahiptir. Bu sektörün sürdürülebilirliği, iklim değişikliği ve toprak özelliklerindeki farklılıklar gibi çok sayıda çevresel değişkenden etkilenmektedir. Çiftçilerin karşılaştığı en temel sorunlardan biri, mevcut agro-klimatik koşullara en uygun mahsulü belirleyebilmek ve bu doğrultuda üretim yapabilmektir. Bu çalışmanın temel amacı, makine öğrenmesi algoritmalarını kullanarak toprak ve çevre koşullarına göre en uygun mahsulü tahmin edebilen bir model geliştirmektir. Modelleme sürecinde, toprağın azot (N), fosfor (P), potasyum (K) içeriği ile sıcaklık, nem, pH değeri ve yağış miktarı gibi çevresel parametreler girdi olarak kullanılmıştır. Çalışmada Rastgele Orman, Lojistik Regresyon, Destek Vektör Makineleri, K-En Yakın Komşu ve Gradyan Artırma olmak üzere beş farklı makine öğrenmesi algoritması uygulanmıştır. Her bir modelin hiperparametrelerini optimize etmek için ızgara arama yaklaşımı ve modellerin doğruluğunu değerlendirmek için 5-kat çapraz doğrulama kullanılmıştır. Sonuçlara göre, %99.32'lik doğruluk oranı ile en başarılı model Rastgele Orman olmuştur. Diğer algoritmalar da yüksek doğruluk sağlamış olsa da Rastgele Orman'ın performansı belirgin şekilde öne çıkmıştır. Elde edilen sonuçlar, literatürde yer alan güncel çalışmalarla karşılaştırıldığında yöntemlerin etkinliğini ve uygulanabilirliğini ortaya koymaktadır. Bu çalışma, çiftçilerin veri destekli kararlar almasına olanak sağlayarak, tarımsal verimliliği ve kaynak kullanım etkinliğini artırabilecek akıllı bir mahsul tavsiye sistemi geliştirilmesine katkı sunmaktadır.

Anahtar kelimeler: Mahsul tavsiye sistemi, Makine öğrenmesi, Sınıflandırma

A MACHINE LEARNING-BASED CROP RECOMMENDATION SYSTEM BASED ON AGRO-CLIMATIC CONDITIONS TO ENHANCE AGRICULTURAL PRODUCTIVITY

ABSTRACT

Agriculture is a strategic sector worldwide from both economic and societal perspectives, playing a critical role especially in food security and rural development. The sustainability of this sector is influenced by numerous environmental variables such as climate change and variations in soil properties. One of the main challenges faced by farmers is identifying the most suitable crop for the prevailing agro-climatic conditions and producing accordingly. The primary objective of this study is to develop a model capable of predicting the most appropriate crop based on soil and environmental conditions using machine learning algorithms. In the modeling process, environmental parameters such as nitrogen (N), phosphorus (P), and potassium (K) content of the soil, along with temperature, humidity, pH level, and rainfall amount, were used as inputs. Five different machine learning algorithms were applied in this study, namely Random Forest, Logistic Regression, Support Vector Machines, K-Nearest Neighbor and Gradient Boosting. The grid search approach was used to optimize each model's

*Sorumlu Yazar/Corresponding Author: onursevli@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 70 – 79, 2025
DOI: 10.62301/usmtd.1702079

hyperparameters, and 5-fold cross-validation was used to assess the models' accuracy. According to the results, the most successful model was Random Forest with an accuracy of 99.32%. Although the other algorithms also achieved high accuracy, the performance of the Random Forest model stood out significantly. When compared with recent studies in the literature, the findings demonstrate the effectiveness and applicability of the proposed methods. This study contributes to the development of an intelligent crop recommendation system that can enhance agricultural productivity and resource use efficiency by enabling farmers to make data-driven decisions.

Keywords: Crop recommendation system, Machine learning, Classification

1. Giriş (Introduction)

Tarım sektörü, dünya nüfusunun beslenme ihtiyacını karşılamanın yanı sıra birçok ülke için önemli bir gelir kaynağıdır. Küresel nüfusun sürekli artışı, gıda güvenliğini sağlamak amacıyla tarımsal üretimde düzenli bir artışı zorunlu kılmaktadır [1]. Geleneksel tarım yöntemleri, iklim değişikliğinin ve artan çevresel baskıların etkisiyle verimlilikte düşüş yaşanma risklerine açıktır. İklim değişikliği, toprak verimliliğindeki azalmalar ve kaynakların sınırlı olması gibi faktörler tarımsal üretimi olumsuz etkilemektedir. Bu zorlukların üstesinden gelmek ve sürdürülebilir tarım uygulamalarını teşvik etmek için teknolojik yeniliklere ihtiyaç duyulmaktadır [2]. Bu bağlamda, sürdürülebilir ve verimli tarım uygulamaları, artan talebi karşılarken çevreye olumsuz etkiyi en aza indirmek için hayati öneme sahiptir. Hassas tarım, tarımsal girdileri ve uygulamalarını optimize etmek için veri odaklı bir yaklaşım sunarak bu ihtiyaca cevap vermektedir. Bu uygulamalar veri odaklı kararlar alarak kaynak kullanımını optimize etmeyi ve verimliliği artırmayı hedefler. Bu bağlamda, makine öğrenmesi teknikleri, tarımsal verilerden anlamlı bilgiler çıkararak çiftçilere rehberlik etme potansiyeline sahiptir [3], [4].

Çiftçilerin en kritik kararlarından biri, arazilerinin ve iklim koşullarının özelliklerine en uygun mahsulü seçmektir. Yanlış mahsul seçimi, düşük verim, kaynak israfı ve ekonomik kayıplara neden olur. Geleneksel yöntemlerle mahsul seçimi genellikle deneyime veya sınırlı bilgiye dayanır. Ancak, toprak özellikleri (mineraller, pH), iklimsel faktörler (sıcaklık, nem, yağış) ve mahsulün özel gereksinimleri gibi birçok karmaşık faktörün bir arada değerlendirilmesi gerekmektedir. Makine öğrenmesi algoritmaları, bu tür çok değişkenli verileri analiz ederek belirli koşullar için en uygun mahsulü tahmin edebilen modeller geliştirmeyi sağlar [5].

Teknolojinin yaygınlaşması ve yapay zekâ tekniklerinin gelişmesi ile birlikte son dönemde tarım alanında makine öğrenmesi algoritmaları kullanılarak yapılan çalışmaların sayısı giderek artmaktadır. Toprak besin maddeleri (Azot, Fosfor, Potasyum), sıcaklık, nem, pH ve yağış gibi agro-klimatik parametreler, mahsulün büyümesini ve verimini önemli ölçüde etkileyen temel faktörlerdir. Mohapatra ve Kale [6] agro-klimatik ve toprakla ilgili bir dizi parametre kullanarak çeşitli makine öğrenmesi teknikleri ile tahminlemeler gerçekleştirmiştir Yağış, nem, toprak tipi, pH ve sıcaklık parametreleri üzerinde K-En Yakın Komşu (KNN), Yapay Sinir Ağları (YSA), Rastgele Orman (RO) ve Destek Vektör Makineleri (DVM) algoritmaları kullanılarak gerçekleştirilen sınıflandırmalarda KNN algoritması ile %97.95 doğruluğa ulaşmışlardır.

Madhuri ve Indiramma [7]; minimum, maksimum ve ortalama sıcaklık, güneşlenme süresi, toprak dokusu, toprak pH'ı, çakıl durumu, erozyon, eğim, derinlik ve toprak drenajı parametreleri kullanarak dört mahsul türünü (mısır, darı, pirinç, şeker kamışı) tahminlemeye çalışmışlardır. YSA kullanarak gerçekleştirdikleri analizlerde %96 doğruluk elde etmişlerdir. Bu çalışmalarında, çeşitli toprak ve iklim faktörlerini mahsul tavsiyesi için entegre etmede YSA'nın etkinliğini vurgulamaktadırlar.

Acharya vd. [8], üre, fosfor, potasyum, pH, sıcaklık gibi verilere dayalı olarak çiftçilere en uygun mahsulü öneren bir sistem sunulmuştur. 10 farklı mahsul içeren 10 bin örnekli veri seti üzerinde Naive Bayes (NB) , DVM , Karar Ağacı (KA), ve RO algoritmaları ile sınıflandırmalar gerçekleştirmişlerdir. Her bir modeli bağımsız olarak değerlendirdikten sonra tümünün birleşimi ile ortaya koydukları hibrit modeli test etmişler ve %98.98 doğruluk elde etmişlerdir.

Deepika ve Andrew [9]; azot, fosfor, potasyum oranları, sıcaklık, nem, pH ve yağış miktarı parametrelerini kullanarak ortamda yetişmesi muhtemel 22 farklı mahsul türünü tahmin etmeye çalışmışlardır. Çok katmanlı algılayıcılarla (ÇKA) gerçekleştirdikleri sınıflandırmalarda %97 doğruluk değerine ulaşmışlardır. ÇKA'nın doğrusal olmayan ilişkileri modelleme ve çeşitli veri türlerinden

*Sorumlu Yazar/Corresponding Author: onursevli@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 70 – 79, 2025
DOI: 10.62301/usmtd.1702079

öğrenme yeteneği bu modeli geleneksel makine öğrenmesi modellerine nazaran daha ideal bir çözüm haline getirmektedir.

Ramzan vd. [10], nesnelerin interneti teknolojisi kullanarak gerçekleştirdikleri çalışmalarında sensörlerden gerçek zamanlı olarak toplanan sıcaklık, iletkenlik, toprak su ihtiyacı ve dokusu gibi parametrelere ek olarak azot, fosfor, yağış bilgisi gibi parametreleri içeren ikincil bir veri seti kullanarak uygun mahsulün seçimine yönelik bir tahminleme çalışması gerçekleştirmişlerdir. Beş farklı algoritma kullanarak gerçekleştirdikleri analizlerde YSA ile %93.63 doğruluğa ulaşmışlardır.

Bhavani vd. [11], toprak mineral değerleri, pH, yağış, sıcaklık ve nem ölçümlerinden oluşan 7 özellik ve 2200 adet örnekli veri seti üzerinde Lojistik Regresyon (LR), DVM, KNN, Karar ağacı, RO, Naive Bayes (NB) algoritmaları yanında iki gizli katmanlı bir derin sinir ağı kullanmışlardır. En düşük doğruluğu %75 ile NB modeli ile en yüksek doğruluğu ise %97 ile derin sinir ağı ile elde etmişlerdir. Elde edilen bulgular derin sinir ağlarının toprak ve çevre arasındaki karmaşık ilişkileri geleneksel makine öğrenmesi yöntemlerine göre daha iyi ortaya koyduğunu göstermektedir.

Bu çalışmada çeşitli agro-klimatik parametreleri girdi olarak kullanarak en uygun mahsulü tahmin edebilen makine öğrenimi tabanlı bir sınıflandırma modelinin geliştirilmesi amaçlanmıştır. Bu doğrultuda, yaygın olarak kullanılan beş farklı sınıflandırma algoritması – Rastgele Orman, Lojistik Regresyon, Destek Vektör Makineleri, K-En Yakın Komşu ve Gradyan Artırma – uygulanmış, hiperparametreleri optimize edilmiş ve performansları kapsamlı şekilde değerlendirilmiştir. Çalışmanın sonuçları, benzer literatür çalışmalarıyla karşılaştırılarak mevcut araştırmanın tarım alanındaki makine öğrenimi uygulamalarına katkısı tartışılmıştır.

2. Materyal ve Metot (Material and Method)

2.1. Veri seti (Dataset)

Bu çalışmada kullanılan veri seti, çeşitli agro-klimatik parametrelere dayanarak en uygun mahsulün tahmin edilmesini amaçlayan bir yapıya sahiptir. Kamuya açık bir veri deposu olan Kaggle'dan elde edilmiştir. Veri seti, hassas tarım ve çiftçilere, tarım danışmanlarına ve politika yapıcılara destek olmayı amaçlayan makine öğrenimi uygulamalarında tipik olarak kullanılmaktadır. Veri setinde yer alan öznitelikler Tablo 1'de verilmiştir.

Tablo 1. Veri setinin öznitelikleri (Features of the dataset)

Öznitelik Adı	Açıklama	Birimi
N	Topraktaki azot içeriği	mg/kg
P	Topraktaki fosfor içeriği	mg/kg
K	Topraktaki potasyum içeriği	mg/kg
temperature	Ortalama sıcaklık	°C
humidity	Ortalama bağıl nem	%
ph	Toprak pH değeri	sayısal
rainfall	Yağış miktarı	Mm
label	Hedef değişken	22 mahsul türü

Özniteliklerden ilk 7 adedi girdi ve label özniteliği tahminlenecek hedef özellik olmak üzere toplam 8 öznitelik yer almaktadır. Girdi özniteliklerin istatistiki özeti Tablo 2'de yer almaktadır.

Tablo 2. Özniteliklerin istatistiki karakteristiği (Statistical characteristics of the features)

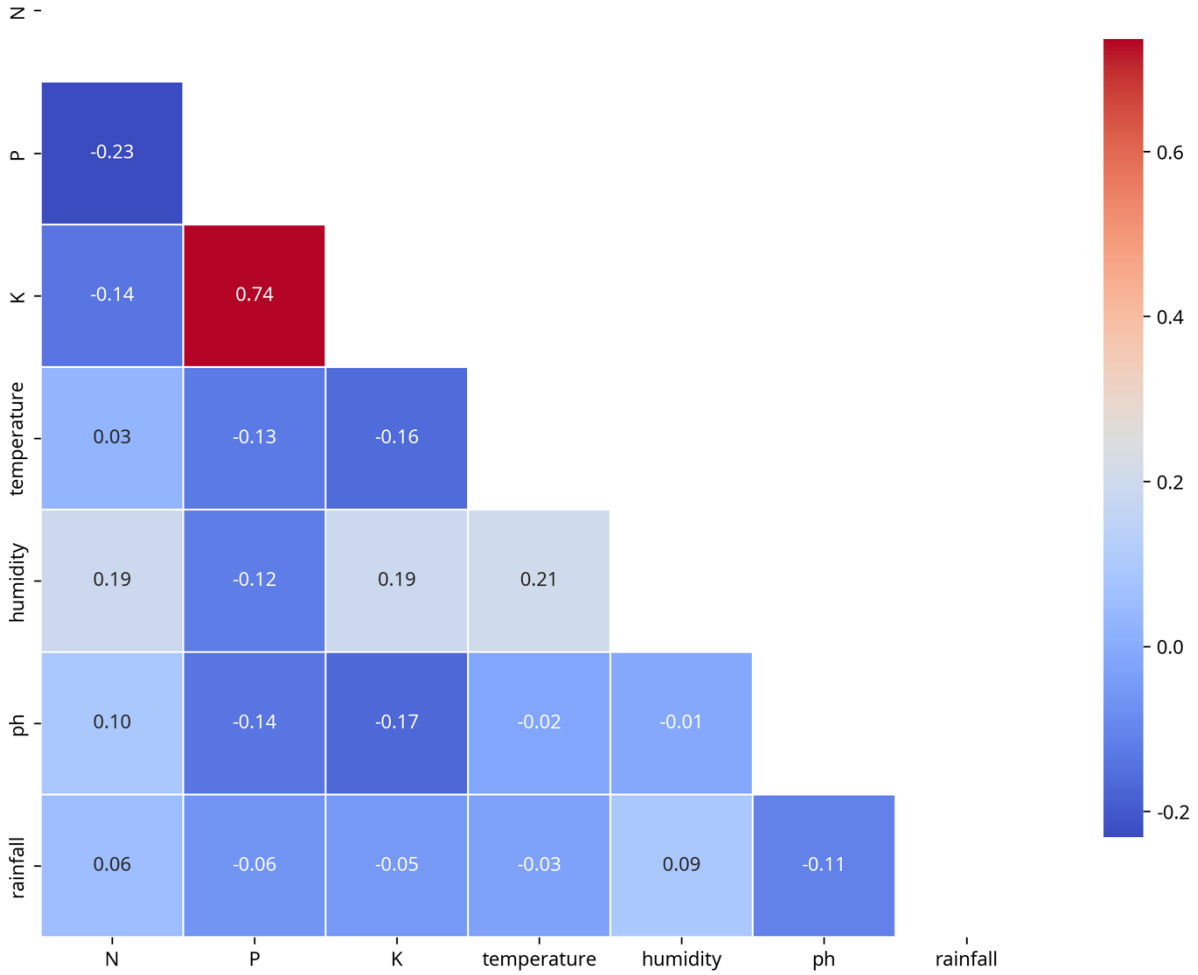
Öznitelik Adı	Minimum	Maximum	Ortalama	Standart Sapma
N	0	140	50.55	36.92
P	5	145	53.36	32.99
K	5	205	48.15	50.65
temperature	8.83	43.68	25.62	5.06
humidity	14.26	99.98	71.48	22.26
ph	3.5	9.94	6.47	0.77

*Sorumlu Yazar/Corresponding Author: onursevli@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 70 – 79, 2025
DOI: 10.62301/usmtd.1702079

rainfall	20.21	298.56	103.46	54.96

Veri setinde yer alan sayısal girdi öznitelikler arasındaki korelasyon Şekil 1'de ısı haritası olarak verilmiştir.



Şekil 1. Korelasyon matrisi (Correlation matrix)

Şekil 1'de verilen korelasyon matrisi incelendiğinde öznitelikler arasında genel olarak yüksek bir korelasyon söz konusu değildir. Yalnız toprakta bulunan fosfor (P) ve potasyum (K) oranları arasında 0.74'lük bir korelasyon olduğu görülmektedir.

Hedef özellik 22 farklı mahsul türünü içermektedir. Veri setinde yer alan mahsuller Tablo 3'te listelenmiştir.

Tablo 3. Veri setinde yer alan mahsul adları (Crop names in the dataset)

Mahsul adı	Mahsul adı
Apple (elma)	Maize (mısır)
Banana (muz)	Mango (mango)
Blackgram (siyah nohut)	Mothbeans (güve fasulyesi)
Chickpea (nohut)	Mungbean (maş fasulyesi)
Coconut (Hindistan cevizi)	Muskmelon (kavun)
Coffee (kahve)	Orange (portakal)
Cotton (pamuk)	Papaya (papaya)
Grapes (üzüm)	Pigeonpeas (güvercin bezelyesi)
Jute (jüt bitkisi)	Pomegranate (nar)
Kidneybeans (barbunya)	Rice (pirinç)
Lentil (mercimek)	Watermelon (karpuz)

*Sorumlu Yazar/Corresponding Author: onursevli@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 70 – 79, 2025
DOI: 10.62301/usmtd.1702079

Veri seti içerisinde 22 farklı mahsul yer almakta ve her bir mahsule ait 100 adet örnek bulunmakta olup, veri seti toplam 2200 örnekten oluşmaktadır. Hedef sınıfların her birine ait eşit sayıda örnek bulunmasından dolayı veri seti dengeli bir dağılıma sahiptir.

2.2. Veri ön işleme (Data preprocessing)

Veri ön işleme, makine öğrenmesi modellerinin performansını doğrudan etkileyen kritik bir adımdır. Bu çalışmada uygulanan ön işleme adımları şunlardır:

1. Eksik Veri Kontrolü: Veri seti, eksik değerler açısından incelenmiş ve herhangi bir eksik değere rastlanmamıştır.
2. Kategorik Veri Kodlaması: Hedef değişken olan 'label' (mahsul türleri) kategorik yapıda olduğundan, algoritmalarının işleyebilmesi için sayısal formata dönüştürülmüştür. Bu amaçla etiket kodlama (label encoding) kullanılmıştır.
3. Öznitelik Ölçeklendirme: Farklı ölçeklerdeki sayısal özniteliklerin model performansı üzerindeki olumsuz etkisini gidermek ve özellikle KNN ve DVM gibi mesafeye duyarlı algoritmaların performansını artırmak için öznitelik ölçeklendirme uygulanmıştır.

2.3. Kullanılan makine öğrenmesi modelleri (The machine learning models employed)

Bu çalışmadaki mahsul sınıflandırma problemi için literatürde de yaygın olarak kullanılan aşağıdaki beş makine öğrenmesi modeli kullanılmıştır:

Rastgele Orman (RO), birden fazla karar ağacının bir araya gelerek oluşturduğu bir topluluk (ensemble) öğrenme yöntemidir ve sınıflandırma ile regresyon problemlerinde sıklıkla kullanılır. Bu yöntem, her ağaç için farklı rastgele örnekler ve özellik alt kümeleri seçerek ağaçlar arasında çeşitlilik sağlar; bu da modelin genelleme yeteneğini artırır [12]. RO, özellikle yüksek boyutlu veri setlerinde, çok sayıda özellik içeren durumlarda etkili sonuçlar sunar ve genellikle aşırı öğrenmeye karşı dirençli yapısıyla bilinir. Ayrıca, modelin her bir özelliğe verdiği önemi belirleyebilmesi, özellikle açıklanabilirlik açısından değerlidir. Bu nedenle sağlık, finans ve biyoinformatik gibi pek çok alanda tercih edilmektedir [13].

Lojistik Regresyon (LR), bağımlı değişkenin kategorik olduğu durumlarda kullanılan temel istatistiksel modelleme yöntemlerinden biridir. Sıklıkla ikili sınıflandırma problemlerinde kullanılsa da, çok sınıflı genişletmeleri ile çoklu sınıflandırma problemlerinde de etkili şekilde uygulanabilir [14]. Modelin en önemli avantajı, tahmin edilen olasılıkların yorumlanabilir olmasıdır; bu, özellikle tıp ve sosyal bilimler gibi karar destek sistemlerinin önem taşıdığı alanlarda kullanışlıdır. LR, yapısı gereği doğrusal sınırlar çizse de uygun dönüşümlerle doğrusal olmayan problemler için de genişletilebilir. Basitliği ve hesaplama verimliliği sayesinde büyük veri setlerinde bile hızlı sonuçlar verebilir.

Destek Vektör Makineleri (DVM), örnekler arasındaki maksimum marjini sağlayacak hiperdüzlemi bulmayı amaçlayan, denetimli bir öğrenme algoritmasıdır. Bu özelliği sayesinde genelleme yeteneği oldukça yüksektir ve özellikle az sayıda örnekle yüksek boyutlu uzaylarda etkili sonuçlar verir [15]. DVM, doğrusal ayırıcıların yetersiz kaldığı durumlarda çekirdek (kernel) fonksiyonları sayesinde doğrusal olmayan ayırımlar da yapabilir. DVM'nin dezavantajlarından biri, büyük veri setlerinde eğitim süresinin uzun olabilmesidir; ancak doğruluk açısından çoğu zaman tatmin edici performans sağlar.

K-En Yakın Komşu (KNN), sınıflandırma ve regresyon problemlerinde kullanılan, örneklerin doğrudan özellik uzayındaki yakınlığına göre sınıflandırıldığı parametrik olmayan bir algoritmadır. Herhangi bir öğrenme süreci gerektirmemesi, yani "lazy learning" olması nedeniyle basit ve uygulaması kolaydır [16]. KNN, veri noktaları arasındaki mesafeye dayanarak karar verir ve bu nedenle özelliklerin uygun şekilde ölçeklendirilmesi oldukça kritiktir. Gürültülü verilere karşı duyarlıdır ve özellikle yüksek boyutlu veri setlerinde performansı düşebilir, bu da "boyutsallık laneti" olarak bilinir [17]. Bununla birlikte, doğru parametre (K değeri) seçildiğinde etkili ve güvenilir sonuçlar verebilir.

Gradyan Artırma (GA), zayıf öğrenicileri — genellikle karar ağaçlarını — ardışık olarak eğiterek hata değerini en aza indirmeye çalışan güçlü bir topluluk öğrenme yöntemidir. Her yeni model, bir öncekinin hata oranını azaltacak şekilde eğitilir ve bu süreç, modelin genel doğruluğunu artırır [18]. Bu teknik, sınıflandırma ve regresyon gibi birçok denetimli öğrenme probleminde yüksek performans sergiler. GA,

*Sorumlu Yazar/Corresponding Author: onursevli@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 70 – 79, 2025
DOI: 10.62301/usmtd.1702079

overfitting riskine karşı dikkatli hiperparametre ayarları gerektirir ancak düzenleme teknikleriyle bu risk azaltılabilir. Günümüzde XGBoost, LightGBM ve CatBoost gibi modern GA tabanlı kütüphaneler, makine öğrenmesi yarışmalarında sıklıkla birincilik alan modellerin temelini oluşturur [19].

2.4. Hiperparametre optimizasyonu (Hyperparameter optimization)

Modellerin performansını en üst düzeye çıkarmak için hiperparametre optimizasyonu uygulanmıştır. Her bir model için hiperparametre aralıkları tanımlanmış ve ızgara arama yöntemi kullanılarak en iyi hiperparametre kombinasyonları belirlenmiştir. Optimizasyon sürecinde 5 kat çapraz doğrulama ve doğruluk metriği esas alınmıştır.

3. Deneysel Çalışma ve Bulgular (Experimental Study and Findings)

Çeşitli agro-klimatik değerlere bağlı olarak yetiştirilmeye uygun mahsul tahminlemesine yönelik 7 girdi ve 1 hedef olmak üzere 8 adet öznitelik ve 2200 örnekten oluşan veri seti üzerinde literatürde yaygın kullanılan Rastgele Orman (RO), Lojistik Regresyon (LR), Destek Vektör Makineleri (DVM) , K-En Yakın Komşu (KNN) ve Gradyan Artırma (GA) algoritmaları kullanılarak sınıflandırmalar gerçekleştirilmiştir. Her bir algoritmanın ideal parametrelerine ulaşmak için gerçekleştirilen optimizasyon süreci sonrası tespit edilen hiperparametre değerleri Tablo 4'te verilmiştir.

Tablo 4. Modellerin optimize edilmiş hiperparametreleri (Optimized hyperparameters of the models)

Model	Optimize hiperparametreler
RO	'max_depth': 10, 'min_samples_leaf': 1, 'min_samples_split': 2, ' n_estimators': 200
LR	'C': 100, 'max_iter': 100, 'solver': 'lbfgs'
DVM	'C': 10, 'gamma': 'scale', 'kernel': 'linear'
KNN	'metric': 'manhattan', 'n_neighbors': 5, 'weights': 'distance'
GA	'learning_rate': 0.2, 'max_depth': 3, 'n_estimators': 200

Hiperparametre optimizasyonundan sonra elde edilen en iyi modellerin genelleme performansını daha güvenilir bir şekilde değerlendirmek amacıyla 5 kat çapraz doğrulama uygulanmıştır. Optimize edilmiş hiperparametrelerle yapılandırılan modellerin performansları farklı metrikler açısından değerlendirilmiş ve ortalama skorlar raporlanmıştır.

Model performanslarının değerlendirilmesinde temel olarak karışıklık matrislerinden (confusion matrix) yararlanılmıştır. Bir karışıklık matrisi kullanılan sınıflandırma modelinin sınıfları birbirinden ayırt etmedeki başarısını ölçmeye yardımcı olur. Genel olarak bir karışıklık matrisinin yapısı Şekil 2'de görülmektedir.

		Gerçek sınıf	
		Pozitif	*Negatif*
Tahmin edilen	*Pozitif*	Doğru Pozitif (DP)	Yanlış Pozitif (YP)
	Negatif	Yanlış Negatif (YN)	Doğru Negatif (DN)

Şekil 2. Karışıklık matrisi

Sınıflandırma modelinin Pozitif olarak tahmin ettiği değer gerçekte de Pozitif ise Doğru Pozitif (DP), gerçekte Negatif ise Yanlış Pozitif (YP) olarak adlandırılır. Benzer şekilde modelin Negatif olarak tahmin ettiği değer gerçekte de Negatif ise Doğru Negatif (DN), gerçekte Pozitif ise Yanlış Negatif (YN) olarak adlandırılır. Bu değerlere bağlı olarak üretilen doğruluk (accuracy), kesinlik (precision), duyarlılık (recall) ve F1 Skor değerleri Tablo 5'te verilen şekilde hesaplanır.

*Sorumlu Yazar/Corresponding Author: onursevli@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 70 – 79, 2025
DOI: 10.62301/usmtd.1702079

Tablo 5. Model performans metrikleri (Model performance metrics)

Metrik	Açıklama	Formül
Doğruluk (Accuracy)	Doğru sınıflandırılan örneklerin toplam örneklere oranıdır.	(DP + DN) / (DP + DN + YP +YN)
Kesinlik (Precision)	Pozitif olarak tahmin edilen örneklerden kaçının gerçekten pozitif olduğunun oranıdır.	DP / (DP + YP)
Duyarlılık (Recall)	Gerçekten pozitif olan örneklerden kaçının doğru bir şekilde pozitif olarak tahmin edildiğinin oranıdır.	DP / (DP + YN)
F1 Skoru (F1-Score)	Kesinlik ve duyarlılığın harmonik ortalamasıdır, iki metrik arası dengeyi gösterir.	2 * (Kesinlik * Duyarlılık) / (Kesinlik + Duyarlılık)

Optimize edilmiş RO modeli için 5 kat çapraz doğrulama işlemi ile elde edilen ortalama değerler Tablo 6'da verilmiştir.

Tablo 6. Rastgele Orman ile elde edilen sonuçlar (Results obtained using Random Forest)

Metrik	Ortalama Değer (%)
Doğruluk	99.32
Kesinlik	99.35
Duyarlılık	99.32
F1 Skoru	99.33

Optimize edilmiş LR modeli için 5 kat çapraz doğrulama işlemi ile elde edilen ortalama değerler Tablo 7'de verilmiştir.

Tablo 7. Lojistik Regresyon ile elde edilen sonuçlar (Results obtained using Logistic Regression)

Metrik	Ortalama Değer (%)
Doğruluk	98.41
Kesinlik	98.52
Duyarlılık	98.41
F1 Skoru	98.46

Optimize edilmiş DVM modeli için 5 kat çapraz doğrulama işlemi ile elde edilen ortalama değerler Tablo 8'de verilmiştir.

Tablo 8. Destek Vektör Makineleri ile elde edilen sonuçlar (Results obtained using Support Vector Machines)

Metrik	Ortalama Değer (%)
Doğruluk	98.86
Kesinlik	98.97
Duyarlılık	98.86
F1 Skoru	98.91

Optimize edilmiş KNN modeli için 5 kat çapraz doğrulama işlemi ile elde edilen ortalama değerler Tablo 9'da verilmiştir.

Tablo 9. K-En Yakın Komşu ile elde edilen sonuçlar (Results obtained using K-Nearest Neighbors)

Metrik	Ortalama Değer (%)

*Sorumlu Yazar/Corresponding Author: onursevli@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 70 – 79, 2025
DOI: 10.62301/usmtd.1702079

Doğruluk	98.18
Kesinlik	98.23
Duyarlılık	98.18
F1 Skoru	98.20

Optimize edilmiş GA modeli için 5 kat çapraz doğrulama işlemi ile elde edilen ortalama değerler Tablo 10'da verilmiştir.

Tablo 10. Gradyan Artırma ile elde edilen sonuçlar (Results obtained using Gradient Boosting)

Metrik	Ortalama Değer (%)
Doğruluk	98.64
Kesinlik	98.67
Duyarlılık	98.64
F1 Skoru	98.65

Kullanılan 5 farklı sınıflandırma modeli için elde edilen skorlar özet olarak Tablo 11'de verilmiştir.

Tablo 11. Tüm modeller ile elde edilen sonuçlar (Results obtained from all models)

Model	Doğruluk	Kesinlik	Duyarlılık	F1 Skoru
RO	99.32	99.35	99.32	99.33
LR	98.41	98.52	98.41	98.46
DVM	98.86	98.97	98.86	98.91
KNN	98.18	98.23	98.18	98.20
GA	98.64	98.67	98.64	98.65

Bu çalışmada test edilen beş farklı modelin, mahsul tavsiyesi probleminde dikkate değer bir performans gösterdiği görülmektedir. Tüm metrikler açısından en yüksek skorları veren RO modeli, %99.32 doğruluk sağlamıştır. Bu durum, RO algoritmasının topluluk öğrenme yapısı sayesinde varyansı azaltma ve aşırı öğrenmeyi engelleme yeteneği ile açıklanabilir. Karar ağaçlarının birleşimi, farklı özellik etkileşimlerini yakalayarak karmaşık ilişkileri modellemede etkili olmuştur. DVM ve GA modelleri sırayla %98.86 ve %98.64 doğruluk sağlamışlardır. Bu modeller tüm metrikler açısından %98.5'in üzerinde oranlarla RO modeline yakın bir performans sergilemişlerdir. DVM'nin 'linear' çekirdek ile yüksek performans göstermesi, veri setindeki sınıfların doğrusal olarak ayırt edilebilirliğinin yüksek olduğuna işaret etmektedir. LR ve KNN modelleri diğerlerine nazaran daha az başarı gösterse de, %98'in üzerindeki skorları ile oldukça etkili oldukları söylenebilir. Özellikle LR'nin basitliği ve yorumlanabilirliği, bazı pratik uygulamalarda tercih sebebi olabilir.

Bu çalışmada elde edilen bulgular literatürdeki benzer güncel çalışmalarla karşılaştırılarak Tablo 12'de verilmiştir.

Tablo 12. Literatür karşılaştırması (Comparison with the literature)

Çalışma	Veri seti özellikleri	Kullanılan Model	Doğruluk
Mohapatra ve Kale (2021) [6]	Yağış, nem, toprak tipi, pH ve sıcaklık	KNN	%97.95
Madhuri ve Indiramma (2021) [7]	Sıcaklık, güneşlenme süresi, pH, çakıl durumu, erozyon, eğim, derinlik ve toprak drenajı	YSA	%96
Musanese vd. (2023) [20]	Azot, fosfor, potasyum, pH, yağış, sıcaklık, nem	YSA	%97
Acharya vd. (2024) [8]	Üre, fosfor, potasyum, pH, sıcaklık	Hibrit(NB+DVM+KA+RO)	%98.98
Deepika ve Andrew (2024) [9]	Azot, fosfor, potasyum, pH, yağış, sıcaklık, nem	Çok katmanlı algılayıcı	%97
Ramzan vd. (2024) [10]	Sıcaklık, iletkenlik, toprak su ihtiyacı ve dokusu, azot, fosfor, yağış bilgisi	YSA	%93.63
Bhavani vd. (2024) [11]	Toprak mineral değerleri, pH, yağış, sıcaklık ve nem	Naive Bayes	%75

*Sorumlu Yazar/Corresponding Author: onursevli@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 70 – 79, 2025
DOI: 10.62301/usmtd.1702079

Afzal vd. (2025) [2]	Azot, fosfor, potasyum, pH, yağış, sıcaklık, nem	RO+XGBoost	%98
Bu çalışma	Azot, fosfor, potasyum, pH, yağış, sıcaklık, nem	Optimize RO	%99.32

Tablo 12 incelendiğinde literatürdeki çalışmaların hemen hemen benzer özelliklere sahip veri setleri üzerinde çalıştığı görülmektedir. Literatürdeki çalışmalarda genel olarak hatrı sayılır başarıların elde edildiği görülmektedir. Bu çalışmaya en yakın doğruluk oranı %98 ile Afzal vd. [2] tarafından yapılan çalışmada elde edilmiştir. Afzal vd. [2] çalışmalarında RO ve XGBoost birleşimi hibrit bir algoritma kullanmışlardır. Bu çalışmada da RO Modeli ile en yüksek doğruluk değerine ulaşılmıştır. Hibrit bir model kullanılmamasına karşın model parametrelerinin optimize edilmesi geleneksel modellere nazaran daha yüksek bir başarı elde edilmesini sağlamıştır.<Literatürdeki çalışmaların geneli ile karşılaştırıldığında bu çalışmada elde edilen %99.32 oranındaki doğruluk diğer çalışmalarla rekabet edebilecek düzeyde ve görece daha yüksektir.

Elde edilen yüksek başarının yanında bu çalışmanın sınırlılık olarak kabul edilecek bazı yanları da bulunmaktadır. Kullanılan veri seti belirli bir coğrafi bölgeye ait olduğu için daha geniş bir çeşitliliğe sahip toprak ve iklim koşullarına uygun olmayabilir. Bu nedenle modellerin farklı bölgelerdeki ölçümlerle test edilerek genelleme kabiliyetinin sınanması gerekebilir. Modellerin performansı mevcut veri setine bağlıdır ve veri setindeki olası gürültü veya yanlılıklar sonuçları etkileyebilir. Bu çalışma statik bir veri seti üzerinde gerçekleştirilmiştir; dinamik olarak değişen iklim koşulları ve toprak özellikleri için modellerin zamanla güncellenmesi gerekebilir.

4. Sonuç (Conclusion)

Bu çalışmada, yedi farklı agro-klimatik özelliğe (N, P, K, sıcaklık, nem, pH, yağış) dayanarak 22 farklı mahsul türü için en uygun olanı tahmin etmek amacıyla beş farklı makine öğrenmesi sınıflandırma modeli (Rastgele Orman, Lojistik Regresyon, Destek Vektör Makineleri, K-En Yakın Komşu ve Gradyan Artırma) başarıyla uygulanmış ve karşılaştırılmıştır. Modellerin hiperparametreleri optimize edilmiş ve performansları doğruluk, kesinlik, duyarlılık ve F1 skoru metrikleri kullanılarak değerlendirilmiştir. Elde edilen sonuçlar, tüm modellerin yüksek performans gösterdiğini, özellikle Rastgele Orman modelinin %99.32 doğruluk oranı ile en iyi sonucu verdiğini ortaya koymuştur. Bu bulgular, makine öğrenmesinin tarımda karar verme süreçlerini desteklemek ve mahsul verimliliğini artırmak için güçlü bir araç olabileceğini göstermektedir. Çalışmanın sonuçları, literatürdeki benzer çalışmalarla da uyumludur ve Rastgele Orman gibi topluluk öğrenme yöntemlerinin bu tür sınıflandırma problemlerinde etkinliğini bir kez daha teyit etmektedir.

Geliştirilen modeller, çiftçilere kendi arazileri ve iklim koşulları için en uygun mahsulü seçme konusunda bilimsel temelli tavsiyeler sunarak, kaynakların daha verimli kullanılmasına ve tarımsal üretimin sürdürülebilirliğine katkıda bulunma potansiyeline sahiptir. Gelecekteki çalışmalar, daha kapsamlı veri setleri ve ek özellikler kullanarak bu modelleri daha da geliştirmeye odaklanabilir.

5. Kaynaklar (References)

[1] H.C.J. Godfray, J.R. Beddington, I.R. Crute, L. Haddad, D. Lawrence, J.F. Muir, et al., Food security: the challenge of feeding 9 billion people, Science 327 (5967) (2010) 812–818, doi:10.1126/science.1185383.

[2] H. Afzal, M. Ahmad, A. Aftab, et al., Incorporating soil information with machine learning for crop recommendation to improve agricultural output, Sci. Rep. 15 (1) (2025) 8560, doi:10.1038/s41598-025-88676-z.

[3] R. Parvathi, Crop recommendation system by artificial neural network, (2021).

[4] D.N. Varshitha, S. Choudhary, An artificial intelligence solution for crop recommendation, Indones. J. Electr. Eng. Comput. Sci. 25 (3) (2022) 1688–1695.

[5] F.S. Prity, M.A. Islam, T. Rahman, et al., Enhancing agricultural productivity: a machine learning approach to crop recommendations, Hum.-Centric Intell. Syst. 4 (1) (2024) 497–510, doi:10.1007/s44230-024-00081-3.

*Sorumlu Yazar/Corresponding Author: onursevli@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 70 – 79, 2025
DOI: 10.62301/usmtd.1702079

[6] B.N. Mohapatra, V. Kale, Crop recommendation system using machine learning, ITEGAM-JETIA 10 (48) (2024) 63–68.

[7] J. Madhuri, M. Indiramma, Artificial neural networks based integrated crop recommendation system using soil and climatic parameters, Indian J. Sci. Technol. 14 (19) (2021) 1587–1597.

[8] N. Acharya, P. Khatiwada, R. Pandey, S. Niroula, P. Chapagain, Crop recommendation system using machine learning: a comparative study, in: Proc. KEC Conf. 2024, (2024) 302–311, doi:10.3126/injet.v1i2.66708.

[9] M.P. Deepika, G. Andrew, Artificial neural network for crop recommendation, Grenze Int. J. Eng. Technol. 10 (2024) 3933–3936.

[10] S. Ramzan, A. Khan, A. Hussain, et al., An innovative artificial neural network model for smart crop prediction using sensory network based soil data, PeerJ Comput. Sci. 10 (2024) e2478.

[11] R. Bhavani, V. Agalya, N. Kiruthika, C. Sugapriya, Deep neural network based crop recommendation system, Int. Res. J. Eng. Technol. (IRJET) 11 (4) (2024) 278–281.

[12] L. Breiman, Random forests, Mach. Learn. 45 (1) (2001) 5–32, doi:10.1023/A:1010933404324.

[13] X. Chen, H. Ishwaran, Random forests for genomic data analysis, Genomics 99 (6) (2012) 323–329, doi:10.1016/j.ygeno.2012.04.003.

[14] D.W. Hosmer, S. Lemeshow, R.X. Sturdivant, Applied Logistic Regression, third ed., Wiley, 2013, doi:10.1002/9781118548387.

[15] C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. 20 (3) (1995) 273–297, doi:10.1007/BF00994018.

[16] T. Cover, P. Hart, Nearest neighbor pattern classification, IEEE Trans. Inf. Theory 13 (1) (1967) 21–27, doi:10.1109/TIT.1967.1053964.

[17] K. Beyer, J. Goldstein, R. Ramakrishnan, U. Shaft, When is 'nearest neighbor' meaningful?, in: Int. Conf. Database Theory, (1999) 217–235, doi:10.1007/3-540-49257-7_15.

[18] J.H. Friedman, Greedy function approximation: a gradient boosting machine, Ann. Stat. 29 (5) (2001) 1189–1232, doi:10.1214/aos/1013203451.

[19] T. Chen, C. Guestrin, XGBoost: a scalable tree boosting system, in: Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., (2016) 785–794, doi:10.1145/2939672.2939785.

[20] C. Musanase, A. Vodacek, D. Hanyurwimfura, A. Uwitonze, I. Kabandana, Data-driven analysis and machine learning-based crop and fertilizer recommendation system for revolutionizing farming practices, Agriculture 13 (11) (2023) 2141.

***Sorumlu Yazar/Corresponding Author:** onursevli@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 80 – 93, 2025
DOI: 10.62301/usmtd.1716034

Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi
International Journal of Sustainable Engineering and Technology
ISSN: 2618-6055 / Cilt:9, Sayı:1, (2025), 80 – 93
Araştırma Makalesi – Research Article

COMPARATIVE ANALYSIS OF CLASSICAL AND QUANTUM SVM MODELS ON MEDICAL DIAGNOSIS DATASETS

*Gamzepelin Aksoy[1], Zeynep Özpolat[2]

[1]Isparta University of Applied Sciences, Faculty of Technology, Department of Computer Engineering, Isparta
[2]Mus Alparslan University, Faculty of Engineering-Architecture, Department of Software Engineering, Mus

(Geliş/Received: 08.06.2025, Kabul/Accepted: 19.06.2025, Yayınlanma/Published: 30.06.2025)

ABSTRACT

Quantum-assisted machine learning approaches have become a significant area of research in the healthcare domain by offering alternative solutions to classical methods, particularly when dealing with high-dimensional and complex datasets. This study presents a comparative evaluation of the classification performance of classical Support Vector Machines (SVM) and quantum-based algorithms Quantum Support Vector Machine (QSVM) and Pegasos-QSVM on healthcare data.
Experimental analyses were conducted using three distinct medical datasets related to liver disease, breast cancer, and heart failure. The results demonstrate that the QSVM model consistently achieved the highest and most stable classification accuracy. Although the Pegasos-QSVM model achieved comparable accuracy rates in certain configurations, its performance was generally more variable. Nevertheless, thanks to its lower computational cost and faster processing time, Pegasos-QSVM emerges as a promising alternative, particularly in resource-constrained environments. The findings suggest that quantum-assisted models can deliver performance levels competitive with classical approaches, particularly highlighting the effectiveness of QSVM on small- to medium-sized datasets.

Keywords: Pegasos-QSVM, QSVM, Medical Datasets, Quantum Machine Learning

TIBBİ TANI VERİ SETLERİ ÜZERİNDE KLASİK VE KUANTUM SVM MODELLERİNİN KARŞILAŞTIRMALI ANALİZİ

ÖZ

Kuantum destekli makine öğrenimi yaklaşımları, özellikle yüksek boyutlu ve karmaşık veri kümeleriyle çalışırken klasik yöntemlere alternatif çözümler sunarak sağlık alanında önemli bir araştırma konusu hâline gelmiştir. Bu çalışma, sağlık verileri üzerinde klasik Destek Vektör Makineleri (SVM) ile kuantum tabanlı algoritmalar olan Kuantum Destek Vektör Makinesi (QSVM) ve Pegasos-QSVM'nin sınıflandırma performanslarının karşılaştırmalı bir değerlendirmesini sunmaktadır.
Karaciğer hastalığı, meme kanseri ve kalp yetmezliğiyle ilgili üç farklı tıbbi veri seti kullanılarak deneysel analizler gerçekleştirilmiştir. Elde edilen sonuçlar, QSVM modelinin tutarlı bir şekilde en yüksek ve en istikrarlı sınıflandırma doğruluğunu sağladığını göstermektedir. Pegasos-QSVM modeli belirli konfigürasyonlarda benzer doğruluk oranlarına ulaşsa da, genel olarak daha değişken bir performans sergilemiştir. Bununla birlikte, daha

*Sorumlu Yazar/Corresponding Author: gamzepelinaksoy@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 80 – 93, 2025
DOI: 10.62301/usmtd.1716034

düşük hesaplama maliyeti ve daha hızlı işlem süresi sayesinde Pegasos-QSVM, özellikle kaynakların kısıtlı olduğu ortamlarda umut vadeden bir alternatif olarak öne çıkmaktadır. Bulgular, kuantum destekli modellerin klasik yaklaşımlarla rekabet edebilecek düzeyde performans gösterebildiğini ve özellikle QSVM'nin küçük ve orta ölçekli veri kümelerinde etkili olduğunu ortaya koymaktadır.

Anahtar kelimeler: Pegasos-QSVM, QSVM, Tıbbi veri setleri, Kuantum makine öğrenme

1. Introduction

Advances in technology have significantly improved the early detection and accurate diagnosis of diseases through the use of high-performance analytical methods. Machine Learning (ML) and Deep Learning (DL) algorithms increasingly serve as digital tools that support clinical decision-making processes [1,2]. As a result, modern technologies have become essential for the efficient collection, storage, and analysis of medical data. However, traditional ML approaches often fall short in handling high-dimensional and complex healthcare datasets, particularly in terms of classification accuracy, computational efficiency, and generalizability.

The exponential growth of the global population has led to a substantial increase in healthcare-related data, rendering conventional computational methods inadequate for managing and analyzing large-scale datasets. Although supercomputers were initially employed to overcome these challenges, the increasing complexity of data has begun to exceed their processing capacity, thereby exposing the limitations of classical computing. Consequently, quantum computing based on fundamentally different algorithmic principles has emerged as a promising paradigm for data intensive applications in healthcare analytics [3,4].

Unlike classical computers, which rely on binary logic, quantum computers are grounded in the principles of quantum physics and operate using phenomena such as superposition and entanglement concepts that transcend classical mechanics [5]. These quantum properties enable systems to process multiple computational paths simultaneously, allowing them to encode and manipulate complex probability distributions. This parallelism proves especially advantageous in analyzing high-dimensional and large-scale healthcare datasets, as it facilitates the exploration of numerous possible outcomes in a single computational cycle [6,7].

Within this framework, Quantum Machine Learning (QML), which represents the integration of quantum computing with machine learning techniques, has emerged as a promising paradigm for enhancing data processing efficiency, accelerating computation, and enabling more robust predictive modeling [8]. QML algorithms are increasingly utilized in the literature to tackle the limitations of classical methods, particularly in managing complex and high-dimensional healthcare datasets [9]. A noteworthy example is the Quantum Support Vector Machine (QSVM), an extension of classical SVMs, which has demonstrated potential in reducing computational complexity while improving classification performance in healthcare applications [10].

In recent years, quantum-based algorithms that offer promising alternatives to classical methods have gained traction in healthcare-related research. Consequently, recent studies that utilize classical SVM, QSVM, and Pegasos-QSVM techniques for medical classification tasks have been examined and reviewed.

Guido et al. [11], in their review on the application of SVM in healthcare, highlighted its strong performance on high-dimensional and imbalanced datasets. Nevertheless, they pointed out persistent limitations, particularly in interpretability and handling of missing values. Kodipalli and Devi [12], compared the Fuzzy TOPSIS and SVM methods for the joint prediction of PCOS and mental health issues, concluding that Fuzzy TOPSIS yielded a higher accuracy of 98.20%. They also noted the superiority of fuzzy approaches in processing linguistic data. Kareem et al. [13], employed the IQ-OTH/NCCD CT lung cancer dataset and achieved 89.88% accuracy through an image processing and SVM-based approach. The best results were obtained by combining a polynomial kernel SVM with Gabor filters and Gray-Level Co-occurrence Matrix (GLCM) features. Öztekin et al. [14], compared six machine learning algorithms, including SVM, for breast cancer diagnosis using the Breast-XD dataset. The highest accuracy (94.34%) was obtained using the explainable XGBoost model (X²GAI). Finally, Ramu et al. [15], proposed a hybrid model integrating SVM and convolutional neural networks (CNN)

*Sorumlu Yazar/Corresponding Author: gamzepelinaksoy@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 80 – 93, 2025
DOI: 10.62301/usmtd.1716034

for early prediction of chronic kidney disease. This model surpassed the performance of classical SVM (94.8%) and Random Forest, reaching an accuracy of 96.8%.

Tudisco et al. [16], compared quantum models, including QSVM and QNN, with classical algorithms such as SVM on datasets related to heart failure, diabetes, and prostate cancer. Their results indicated that QSVM outperformed classical models, especially in imbalanced datasets, by achieving higher recall rates. Similarly, Khushal and Fatima [17], integrated quantum models such as QSVM and Q-KNN with fuzzy logic (FL) to improve both computational efficiency and classification accuracy. Their proposed Fuzzy Quantum Machine Learning (FQML) approach achieved superior results compared to standard QML techniques on chronic disease datasets. Maheshwari et al. [18], employed optimized QSVM (OQSVM) and hybrid quantum multilayer perceptron (HQMLP) models for the classification of cardiovascular diseases. They highlighted that the use of PauliFeatureMap and fine-tuned hyperparameters led to enhanced accuracy in their models. In the field of breast cancer diagnosis, Jose P. et al. [19], demonstrated that the QSVM model optimized using the Elitist Non-Dominated Sorting Genetic Algorithm (ENSGA) significantly outperformed traditional SVM variants, underscoring the algorithm's potential in medical classification tasks.

Chatterjee ve Das [20], proposed a Staged Pegasos Quantum Support Vector Classifier for breast cancer diagnosis, which stratifies patients based on the severity of their condition. The model's performance was further enhanced through the incorporation of fuzzy logic and expert knowledge-based weighted algorithms. Nasir et al. [21], conducted a comparative study of classical and quantum machine learning algorithms for the classification of dermatological diseases. While classical models such as Bagging and Decision Trees achieved the highest accuracy of 98%, among quantum-based approaches, the Pegasos-QSVM model stood out with an accuracy of 84%, surpassing classical SVM in terms of precision. Singh and Pokhrel [22], assessed several quantum machine learning algorithms using various feature mapping techniques for the binary classification of genomic sequence data. Their analysis revealed that Pegasos-QSVM yielded particularly high recall rates. Munshi et al. [23], compared the performance of Pegasos-QSVM and Variational Quantum Classifier (VQC) on a lung cancer dataset and reported that Pegasos-QSVM achieved a superior accuracy of 85%, outperforming VQC.

As a result of the literature review, it is evident that studies based on Quantum Machine Learning (QML) are frequently employed in the healthcare domain, at least as commonly as classical algorithms. However, research specifically focused on the Pegasos-QSVM algorithm remains relatively scarce. This study aims to fill the existing gap in the literature by comprehensively evaluating the performance of both QSVM and Pegasos-QSVM models in real-world healthcare classification tasks, providing a comparative analysis against classical SVM models.

As evidenced by the literature review, Quantum Machine Learning (QML) techniques are increasingly employed in healthcare applications, offering performance levels that often rival those of classical machine learning algorithms. However, studies focusing specifically on the Pegasos-QSVM model remain limited. This study seeks to address this gap by comprehensively evaluating the performance of QSVM and Pegasos-QSVM models in real-world healthcare classification tasks and providing a comparative analysis against classical SVM methods. Through this investigation, the potential advantages of quantum-based approaches in processing complex medical data are highlighted.

The structure of the remainder of this paper is organized as follows: Section 2 introduces the datasets utilized in the study, details the preprocessing procedures, and outlines the classification algorithms applied. Section 3 presents the experimental results obtained from the implementation of the algorithms. Section 4 discusses the study's contributions to the existing literature, provides an evaluation of the findings, and outlines directions for future research.

2. Material and Method

In this study, three publicly available medical datasets were utilized to evaluate and compare the performances of classical and quantum-based classification algorithms. The datasets employed were the Breast Cancer Coimbra Dataset, Heart Failure Clinical Records Dataset, and the Indian Liver Patient Dataset. These datasets were selected due to their binary classification structure, which aligns with the requirements of the Pegasos-QSVM algorithm.

*Sorumlu Yazar/Corresponding Author: gamzepelinaksoy@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 80 – 93, 2025
DOI: 10.62301/usmtd.1716034

Three distinct classification methods were implemented: classical Support Vector Machines (SVM), Quantum Support Vector Machines (QSVM), and Pegasos-QSVM. Prior to classification, multiple preprocessing steps were conducted to ensure data consistency and enhance model performance. Initially, all features were standardized using the StandardScaler, transforming each feature to have zero mean and unit variance. Subsequently, Principal Component Analysis (PCA) was applied for dimensionality reduction, reducing the number of features to between 3 and 9 components to meet the input dimensionality constraints of quantum circuits. Finally, the PCA-transformed data were rescaled using MinMaxScaler to the range [0, π], which is commonly preferred for quantum state encoding due to the periodic nature of quantum gates. This rescaling step enabled the conversion of classical data into a qubit-compatible format, facilitating its integration into quantum circuits.

For quantum classification, the PCA-reduced datasets were mapped from classical bit format to quantum qubit format using four distinct quantum feature maps: X, Y, Z, and ZZ. These feature maps function as encoding mechanisms that embed classical data into high-dimensional Hilbert spaces via parameterized quantum circuits. To evaluate the effect of circuit depth on model performance, each feature map was tested under two repetition settings: reps = 1 and reps = 2. This allowed for a comparative analysis of how entanglement complexity influences the expressivity of the quantum kernels.

The quantum models were trained using the QSVM and Pegasos-QSVM algorithms within a simulated quantum environment. All simulations were carried out using the Qiskit framework with the statevector_simulator backend. In the QSVM implementation, the FidelityQuantumKernel class was employed to compute the quantum kernel matrix by estimating the similarity (fidelity) between quantum states encoded via selected feature maps [24]. Model performance was evaluated using 5-fold stratified cross-validation to ensure balanced representation of class distributions. For the Pegasos-QSVM model, the regularization parameter was set to C = 1000, and the number of optimization steps was set to τ = 100, based on empirical tuning. Across all models and experiments, two key evaluation metrics were considered: classification accuracy and execution time, both averaged over the five folds.

A schematic representation of the complete methodology is presented in Figure 1, illustrating the entire workflow from data preprocessing to classification using both classical and quantum SVM algorithms.

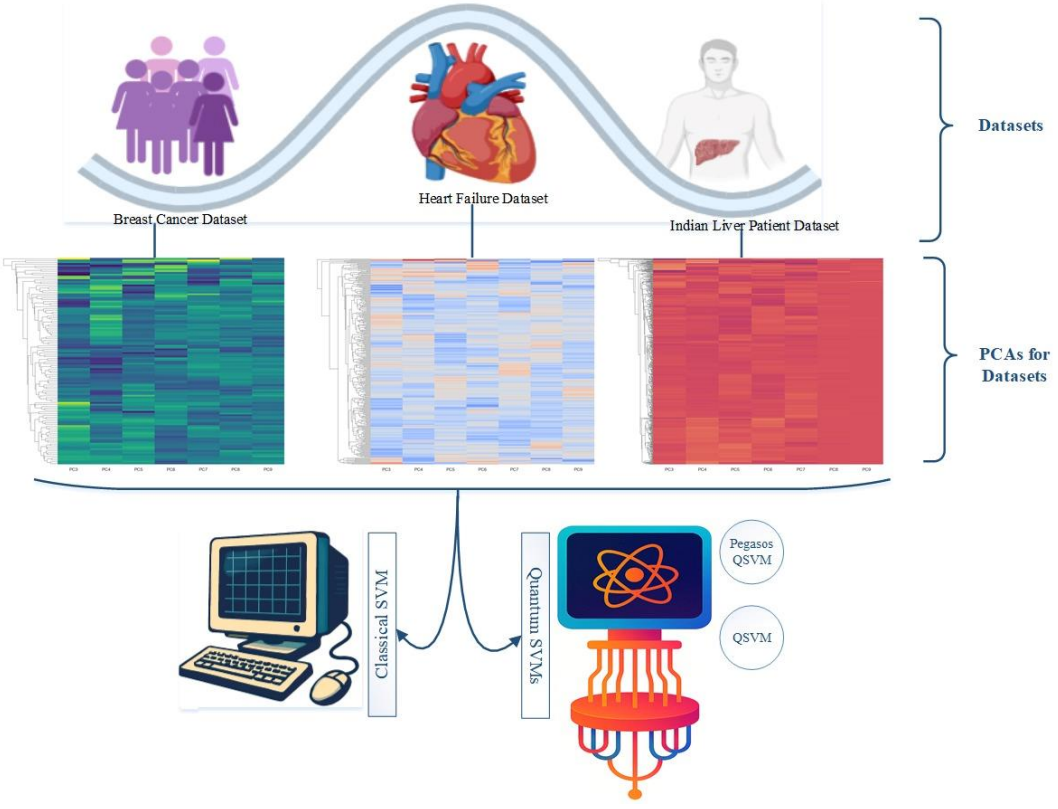


Figure 1. Classification workflow using classical and quantum SVM algorithms

*Sorumlu Yazar/Corresponding Author: gamzepelinaksoy@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 80 – 93, 2025
DOI: 10.62301/usmtd.1716034

2.1.Dataset

In this study, three distinct clinical datasets were utilized to develop and evaluate predictive models for different medical conditions. All datasets were obtained from the UCI Machine Learning Repository, a widely used open-access platform for benchmarking machine learning algorithms [25–27]. Each dataset includes structured, domain-specific features that enable classification tasks related to liver disease, breast cancer, and heart failure, respectively.

2.1.1. Indian liver patient dataset

A publicly available dataset containing the biochemical profiles of 584 individuals from the northeastern region of Andhra Pradesh, India, is employed in this study. The dataset comprises demographic variables such as age and gender, along with clinically significant biochemical indicators including total and direct bilirubin, alkaline phosphatase (ALP), alanine aminotransferase (ALT), aspartate aminotransferase (AST), total protein, albumin, and the albumin-to-globulin ratio. Of the participants, 441 are male and 142 are female. These features are widely recognized as metabolic biomarkers commonly used in the diagnosis of liver dysfunction. The dataset was constructed with the primary objective of supporting the development of predictive models capable of determining whether an individual is affected by liver disease based on these biomarkers [28].

2.1.2. Breast cancer coimbra dataset

An open-access dataset comprising 116 clinical records exclusively from female participants in the Coimbra region of Portugal is employed for breast cancer prediction. The dataset includes nine attributes, encompassing both basic anthropometric measurements such as age and body mass index (BMI) and metabolic or inflammatory biomarkers, including glucose, insulin, HOMA (Homeostatic Model Assessment), leptin, adiponectin, resistin, and MCP-1 (Monocyte Chemoattractant Protein-1). These variables are considered potential indicators associated with the biological mechanisms of breast cancer. Each instance is labeled based on the subject's diagnosis status, making the dataset suitable for supervised classification tasks. The dataset was curated to support the development of machine learning–based early detection systems for breast cancer [29].

2.1.3. Heart failure clinical records dataset

Collected from 299 individuals diagnosed with heart failure, this clinical dataset provides detailed information on demographic, physiological, and laboratory parameters relevant to cardiovascular risk assessment. It includes 12 attributes such as age, sex, ejection fraction, serum creatinine, serum sodium, platelet count, creatinine phosphokinase levels, and the presence of conditions like anemia, diabetes, high blood pressure, and smoking habits. The dataset also contains a follow-up duration feature (in days) and a binary outcome indicating whether a death event occurred during the observation period. These variables are widely recognized as important predictors in the progression and prognosis of heart failure. The dataset serves as a resource for developing predictive models aimed at identifying high-risk patients and improving early intervention strategies in clinical practice [30].

2.2.Data preparation

In this study, the preprocessing pipeline consisted of three main stages. First, all features were standardized using the StandardScaler to ensure equal contribution by transforming them to have zero mean and unit variance. Next, dimensionality reduction was performed via Principal Component Analysis (PCA), aligning the number of features with the number of qubits used in the quantum classification models. Lastly, the reduced feature set was rescaled into the range $[0,\pi]$ using MinMaxScaler, making it compatible with quantum state encoding. This sequence of transformations prepared the data appropriately for both classical and quantum machine learning algorithms.

Principal Component Analysis is a widely adopted statistical technique for reducing the dimensionality of datasets, especially those with a high number of features. It aims to preserve the most relevant information by generating new components that capture the majority of the variance in the original data, with minimal information loss. These components are constructed to be mutually uncorrelated, allowing complex datasets to be restructured into more compact and interpretable forms [31].

*Sorumlu Yazar/Corresponding Author: gamzepelinaksoy@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 80 – 93, 2025
DOI: 10.62301/usmtd.1716034

By applying orthogonal transformations, PCA converts correlated variables into independent ones, which helps suppress noise and reveal meaningful relationships between features [32]. This simplified structure improves the performance of various learning models, including those based on quantum computation. Thanks to its effectiveness and interpretability, PCA remains one of the most widely preferred methods for dimensionality reduction in the literature [33].

2.3. Methods

2.3.1. Support vector machine

Support vector machines are robust and theoretically grounded supervised learning methods widely used for solving classification problems. The fundamental principle of SVM is to separate data instances belonging to different classes by constructing an optimal hyperplane within the feature space. The algorithm aims to maximize the margin between the classes while simultaneously minimizing classification errors [34].

After training the model on labeled data, the resulting decision function is applied to new instances to perform classification. The data points closest to the decision boundary are referred to as “support vectors” and they play a crucial role in determining the performance of the classifier. In linearly separable datasets, SVM constructs a hyperplane that is equidistant from both classes. For datasets that are not linearly separable, kernel functions are employed to transform the data into a higher-dimensional space where a linear separation becomes feasible. Commonly used kernels include linear, polynomial, sigmoid, and Radial Basis Function (RBF) kernels. The hyperplane is positioned perpendicular to the shortest distance between data points from each class, thereby enhancing the separability and improving the model's generalization capability. SVM is a powerful method rooted in statistical learning theory, capable of producing effective results in both binary and multiclass classification tasks [35,36].

2.3.2. Quantum support vector machine

Quantum support vector machines represent a quantum-enhanced adaptation of the classical Support Vector Machine (SVM) algorithm. While classical SVMs aim to find an optimal hyperplane that separates different classes by maximizing the margin between them, QSVM utilizes quantum state space as the feature domain. This allows for more expressive data representation and the potential to solve complex classification problems more efficiently.

The QSVM algorithm integrates Grover’s search algorithm as a quantum subroutine to optimize non-convex cost functions, enabling it to reach global optima [37]. Classical data are encoded into quantum states, and kernel functions are employed to transform the data into a higher-dimensional space where separation becomes feasible. Using quantum superposition and parallelism, QSVM evaluates multiple states simultaneously, accelerating the training process.

In addition to improving classification accuracy, QSVM reduces time complexity from linear to sublinear, making it particularly promising for large-scale learning problems. However, its performance is highly dependent on quantum hardware quality, as quantum noise may impact reliability. Despite this, QSVM stands out for its potential to enhance machine learning efficiency by combining statistical learning theory with quantum computing principles [38].

2.3.3. Pegasos quantum support vector machine

Pegasos-QSVM (Primal Estimated sub-GrAdient SOlver for SVM-Classification) is an efficient algorithm developed to address nonlinear binary classification problems using SVM. Unlike traditional SVM solvers that rely on dual formulation, Pegasos-QSVM directly operates on the primal objective and employs stochastic sub-gradient descent to approximate a solution. Its optimization strategy aims to minimize hinge loss while applying regularization, ensuring both accuracy and generalizability against overfitting [39].

In quantum implementations, constructing accurate kernel matrices often requires a large number of quantum circuit executions, significantly increasing computational costs. Pegasos-QSVM is well-suited for such quantum settings, as it preserves the use of the kernel trick: instead of explicitly mapping data into high-dimensional feature space, it relies on computing inner products. This approach allows for the

*Sorumlu Yazar/Corresponding Author: gamzepelinaksoy@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 80 – 93, 2025
DOI: 10.62301/usmtd.1716034

construction of the separating hyperplane through kernel evaluations, and the decision function is expressed as a weighted sum of kernel values associated with the support vectors. Additionally, the bias term is implicitly handled by embedding a constant into the kernel function, enhancing model flexibility. The quantum Pegasos algorithm requires fewer kernel evaluations than dual QSVM, offering a significant advantage in settings with large datasets and limited quantum resources [40].

Furthermore, in QSVMs, classical data is embedded into the quantum feature space using parameterized quantum circuits, resulting in quantum kernel functions derived from Hilbert–Schmidt inner products. These quantum kernels are only approximately computable due to the probabilistic nature of quantum measurements governed by Born's rule. This inherent statistical noise in quantum settings makes the Pegasos approach even more appealing, as it avoids constructing the full kernel matrix and yields a training complexity that is independent of dataset size. Therefore, Pegasos-QSVM is expected to outperform standard QSVMs in terms of scalability and efficiency when applied to large training sets under quantum constraints [10].

2.3.4. FeatureMap

Quantum feature maps are fundamental components that allow classical data to be encoded into quantum circuits in a form suitable for quantum processing. These maps project classical input vectors into quantum states by embedding them in a high-dimensional Hilbert space [41]. This process involves building quantum circuits composed of specific quantum gates to represent the data as qubits, making it usable by quantum machine learning algorithms.

The choice of feature map significantly influences the expressivity and entanglement structure of the circuit. Key factors include the circuit depth, encoding method, gate types, and the entanglement pattern. Commonly used feature maps include Z, ZZ, and Pauli-based variants, each with unique characteristics. The Z FeatureMap applies Hadamard gates followed by single-qubit unitary operations, encoding data without creating entanglement, thus resulting in a simpler and faster circuit design. In contrast, the ZZ feature map introduces entanglement through Controlled-Z gates, enhancing representational capacity at the cost of increased computational complexity [39]. The PauliFeatureMap offers even greater flexibility by combining Pauli X, Pauli Y, and Z operations, enabling the circuit to adapt to the structure of the input data and the specific classification task [42].

In this study, Pauli X, Pauli Y, Z, and ZZ feature maps were employed to encode the PCA-reduced classical data into quantum form. These maps were selected to assess the impact of different encoding strategies on the classification performance of QSVM and Pegasos-QSVM. Each map introduces different structural and computational complexities, allowing for comparative analysis. For all feature maps, experiments were conducted with both reps=1 and reps=2 to evaluate the influence of circuit depth. Furthermore, the entanglement parameter was set to "full" in the Pauli X, Pauli Y, and ZZ feature maps to investigate how various entanglement configurations affect the models' performance. The circuit structures of the Z, ZZ, Pauli X, and Pauli Y feature maps used in this study are illustrated in Figure 2. These visualizations provide a clear comparison of the different entanglement schemes and circuit complexities introduced by each mapping strategy.

*Sorumlu Yazar/Corresponding Author: gamzepelinaksoy@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 80 – 93, 2025
DOI: 10.62301/usmtd.1716034

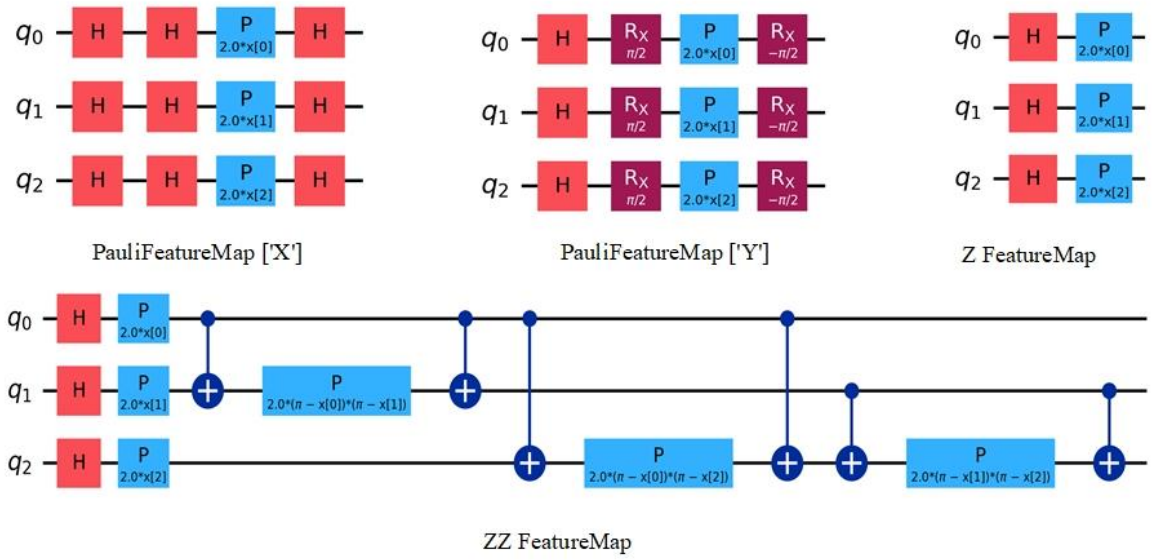


Figure 2. Quantum circuits of the Z, ZZ, Pauli X, and Pauli Y feature maps constructed with 3 qubits and reps = 1

3. Research Findings

In this section, the results of analyses conducted on three different healthcare datasets are presented. The performance of classical SVM and quantum-based models QSVM and Pegasos-QSVM was evaluated based on classification accuracy. To ensure the reliability of the models on both training and testing sets, a 5-fold cross-validation strategy was employed. For the sake of reproducibility, the random seed was fixed at 42 during data splitting.

To enable a comprehensive comparison of quantum algorithms, the experiments were carried out using seven different qubit sizes and four distinct quantum feature maps. Prior to classification, all datasets underwent a series of preprocessing steps. PCA was applied for dimensionality reduction, resulting in datasets with varying numbers of features suitable for quantum circuit input constraints. All algorithms were developed using the Python programming language and implemented with the Qiskit library. Both classical and quantum machine learning models were executed in the Google Colab cloud environment. For the execution of quantum algorithms, the statevector simulator environment was utilized.

The findings for each dataset are presented in the following subsections, starting with the Indian Liver Patient Dataset. The results obtained using the Pegasos-QSVM and QSVM algorithms on the "Indian Liver Patient Dataset," based on different qubit values and feature maps with circuit repetition counts set to 1 and 2, are presented in Table 1.

Table 1. Accuracy analysis for the ILPD using Pegasos-QSVM and QSVM algorithms

	Pegasos-QSVM				**QSVM**			
	Accuracy (%) Reps=1							
Qubits	***X***	***Y***	***Z***	***ZZ***	***X***	***Y***	***Z***	***ZZ***
3	62.98	66.91	66.91	65.53	71.36	71.36	71.36	71.36
4	62.98	64.51	64.51	62.12	71.36	71.36	71.36	70.33
5	71.36	70.67	70.67	63.29	71.36	71.36	71.36	71.18
6	71.36	66.03	66.03	62.74	71.36	70.84	70.84	68.61
7	62.98	65.19	65.19	66.40	71.36	71.36	71.36	71.18
8	62.74	55.09	55.09	57.09	71.36	71.53	71.53	71.01
9	62.74	62.75	62.75	66.87	71.36	71.70	71.70	71.01

*Sorumlu Yazar/Corresponding Author: gamzepelinaksoy@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 80 – 93, 2025
DOI: 10.62301/usmtd.1716034

					Accuracy (%) Reps=2			
Qubits	*X*	*Y*	*Z*	*ZZ*	*X*	*Y*	*Z*	*ZZ*
3	66.91	66.74	69.30	66.39	71.36	71.36	71.36	71.36
4	64.51	65.54	65.87	68.79	71.36	71.36	71.36	71.01
5	70.67	71.36	68.62	71.36	71.36	71.36	71.18	72.38
6	66.03	61.53	64.23	57.56	71.36	71.36	70.67	70.67
7	65.19	62.98	66.56	64.66	71.36	71.36	71.70	71.53
8	55.09	61.90	60.06	62.61	71.53	71.36	71.70	71.70
9	62.75	62.56	67.22	50.09	71.70	71.36	71.70	71.70

The QSVM algorithm demonstrated consistently high and stable performance across nearly all qubit and feature map combinations, achieving an accuracy rate of approximately 71%. The highest recorded accuracy was 72.38%, obtained with 5 qubits, the ZZ feature map, and Reps=2. These results indicate that the QSVM algorithm is relatively insensitive to variations in feature map types or qubit numbers and provides robust and consistent outcomes, especially when circuit repetitions are increased. In contrast, the performance of the Pegasos-QSVM algorithm appeared to be more variable across different configurations. Although it achieved accuracy levels comparable to QSVM in certain scenarios such as 5 and 6 qubits with Reps=1 (both reaching 71.36%) its overall sensitivity to the choice of feature map and qubit count was notably higher. Particularly with the Pauli X and Pauli Y feature maps and at 8 or 9 qubits, accuracy values dropped below 60%. These findings suggest that while Pegasos-QSVM can deliver results similar to QSVM under specific settings, QSVM generally yields more stable and reliable performance across a broader range of parameters.

Table 2 presents the accuracy results obtained from the Breast Cancer Coimbra dataset using Pegasos-QSVM and QSVM algorithms across varying qubit numbers (from 3 to 9), different feature maps (Pauli X, Pauli Y, Z, ZZ), and circuit repetition counts (Reps=1 and Reps=2).

Table 2. Accuracy analysis for the Breast Cancer dataset using Pegasos-QSVM and QSVM algorithms

	Pegasos-QSVM				**QSVM**			
					Accuracy (%) Reps=1			
Qubits	*X*	*Y*	*Z*	*ZZ*	*X*	*Y*	*Z*	*ZZ*
3	52.64	53.41	54.42	51.78	55.17	61.21	61.21	65.52
4	51.70	65.54	66.41	48.95	55.17	67.24	67.24	50.00
5	52.57	67.32	67.32	59.42	55.17	72.41	72.41	56.03
6	53.51	65.51	65.51	53.44	55.17	72.41	74.14	59.48
7	54.28	65.51	65.51	54.28	55.17	75.86	75.86	59.48
8	55.18	62.10	63.01	61.27	55.17	73.28	73.28	58.62
9	59.42	61.20	66.38	59.42	55.17	75.00	75.00	57.76
					Accuracy (%) Reps=2			
Qubits	*X*	*Y*	*Z*	*ZZ*	*X*	*Y*	*Z*	*ZZ*
3	53.41	55.11	55.98	51.67	61.21	55.17	54.31	54.31
4	65.54	49.96	58.62	49.09	67.24	55.17	67.24	51.72
5	67.32	52.72	60.22	52.64	72.41	55.17	70.69	56.90
6	65.51	53.55	65.58	55.22	74.14	55.17	68.97	55.17
7	65.51	50.04	58.51	49.13	75.86	55.17	70.69	56.90
8	62.10	57.57	63.77	56.88	73.28	55.17	73.28	55.17
9	61.20	51.70	64.57	58.51	75.00	55.17	74.14	55.17

The results show that the QSVM algorithm consistently outperformed Pegasos-QSVM across nearly all configurations, especially in settings with 5 to 7 qubits and when using the Z feature map. The highest accuracy value for QSVM was 75.86%, achieved with 7 qubits, the Pauli X or Y feature map, and Reps=1. Furthermore, even with low repetition counts, QSVM maintained relatively stable and high

*Sorumlu Yazar/Corresponding Author: gamzepelinaksoy@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 80 – 93, 2025
DOI: 10.62301/usmtd.1716034

accuracy, generally remaining above 70% in its best configurations. In contrast, the Pegasos-QSVM algorithm exhibited more fluctuating and less predictable performance. While it reached competitive accuracy values in certain configurations (e.g., 67.32% with 5 qubits and the Pauli Y feature map, Reps=1), its performance tended to vary significantly across different qubit and feature map combinations. For instance, accuracy dropped below 55% in multiple cases, especially for ZZ feature maps and low qubit counts. Table 3 presents the accuracy rates obtained using the Pegasos-QSVM and QSVM algorithms on the Heart Failure dataset.

Table 3. Accuracy analysis for the Heart Failure dataset using Pegasos-QSVM and QSVM algorithms

	Pegasos-QSVM				**QSVM**			
	Accuracy (%) Reps=1							
Qubits	***X***	***Y***	***Z***	***ZZ***	***X***	***Y***	***Z***	***ZZ***
3	67.89	63.58	67.59	52.52	67.89	74.58	74.58	69.23
4	39.44	61.59	62.58	53.53	67.89	77.26	77.26	68.90
5	60.77	62.89	57.54	60.56	67.89	74.92	74.92	67.89
6	60.56	67.89	68.56	58.86	67.89	76.92	76.92	69.90
7	60.56	70.57	69.55	65.23	67.89	76.59	76.59	67.56
8	67.89	69.59	62.28	58.54	67.89	74.92	74.92	67.89
9	67.89	61.56	65.56	59.77	67.89	75.59	75.59	67.89
	Accuracy (%) Reps=2							
Qubits	***X***	***Y***	***Z***	***ZZ***	***X***	***Y***	***Z***	***ZZ***
3	65.59	62.56	68.89	54.20	74.58	67.89	73.91	68.23
4	63.23	67.89	58.47	52.54	77.26	67.89	75.25	66.22
5	66.56	67.89	65.56	51.77	74.92	67.89	73.91	67.22
6	64.46	60.44	65.21	52.55	76.92	67.89	75.59	67.56
7	63.59	59.56	64.24	57.56	76.59	67.89	73.24	67.89
8	55.11	65.89	65.88	55.00	74.92	67.89	70.90	67.89
9	61.23	58.45	60.55	59.53	75.59	67.89	68.23	68.23

The QSVM algorithm consistently delivered higher accuracy rates compared to Pegasos-QSVM across both repetition scenarios. For Reps=1, the highest accuracy achieved by QSVM was 77.26%, obtained using 4 qubits with either the Pauli X or Y feature map. In the Reps=2 scenario, QSVM maintained similarly strong performance, particularly with qubit counts between 5 and 7 and the Z feature map, achieving accuracy rates ranging from 73% to 75%. In contrast, the accuracy rates of the Pegasos-QSVM algorithm exhibited a wider variability. For instance, a notably low accuracy of 39.44% was recorded with 4 qubits and the Pauli X feature map, whereas higher accuracies up to 67.89% were achieved in certain scenarios with 8 and 9 qubits. This indicates that Pegasos-QSVM is more sensitive to configuration choices and struggles to maintain consistent performance. In conclusion, the QSVM algorithm demonstrated more stable and higher-accuracy classifications on this dataset, particularly when using fewer repetitions (Reps=1) and a moderate number of qubits (4 to 7).

Table 4 presents the classification accuracy results of the classical SVM algorithm applied to the Indian Liver Patient, Breast Cancer Coimbra, and Heart Failure datasets. The performances are reported based on varying numbers of features, following dimensionality reduction through PCA. This allows for a direct comparison of how the number of features impacts the accuracy of classical SVM across different medical datasets.

Table 4. Accuracy results of SVM across three datasets with varying feature counts

	Indian Liver Patient	***Breast Cancer Coimbra***	***Heart Failure***
Features	**SVM Accuracy (%)**		
3	71.36	62.93	71.58
4	71.18	74.06	71.25

***Sorumlu Yazar/Corresponding Author:** gamzepelinaksoy@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 80 – 93, 2025
DOI: 10.62301/usmtd.1716034

5	71.18	72.39	72.24
6	70.84	71.56	72.56
7	70.84	75.00	73.57
8	71.19	73.26	74.58
9	71.19	72.39	74.58

When comparing the results in Table 4 with those presented in the previous tables for QSVM and Pegasos-QSVM models, several noteworthy observations emerge. Firstly, the QSVM algorithm demonstrates consistent and high performance across all datasets and qubit configurations. Particularly in the Indian Liver Patient and Heart Failure datasets, QSVM frequently achieves equal or higher accuracy values compared to classical SVM. This suggests that QSVM is capable of operating effectively on datasets with limited feature dimensions. On the other hand, the performance of the Pegasos-QSVM algorithm exhibits greater variability. Significant changes in accuracy are observed depending on the feature map and qubit configuration. Although in certain scenarios (e.g., the Breast Cancer dataset with 5 qubits and Reps=2), Pegasos-QSVM achieves accuracy values comparable to or even surpassing QSVM, its overall performance appears to be less consistent. This indicates that Pegasos-QSVM is more sensitive to hyperparameter selection and requires careful tuning to achieve optimal results.

In addition to classification accuracy, the training and prediction times of the models were also evaluated to analyze their practical applicability. All algorithms were executed on the same hardware. Time measurements were conducted using the time module, covering the duration from the start of training to the completion of prediction. For quantum algorithms, the statevector_simulator was used, allowing for the assessment of processing times under ideal conditions.

When considering the execution times of the applied models, it was observed that the classical SVM algorithm produced results in the shortest amount of time. This is attributed to the lower computational complexity of classical machine learning methods. On the other hand, while the QSVM algorithm often achieved higher accuracy, it exhibited the longest execution time due to the additional computational burden of quantum kernel evaluations. The Pegasos-QSVM algorithm was positioned between these two in terms of runtime. Unlike QSVM, this method does not compute the full quantum kernel matrix; instead, it performs calculations over selected subsets, significantly reducing the overall execution time. This finding indicates that although quantum-based methods have notable potential, there remains considerable room for improvement in terms of computational efficiency. Furthermore, with respect to the feature maps used, it was observed that the Pauli X, Pauli Y, and Z maps required similar execution times, whereas the ZZ feature map resulted in longer processing durations due to its more complex structure. For instance, when the number of circuit repetitions (reps) was set to 1 in the Pegasos-QSVM algorithm, it was observed that the Pauli X, Pauli Y, Z, and ZZ feature maps completed their analyses in 775, 807, 614, and 1933 seconds, respectively. Similarly, in the QSVM algorithm under the same conditions, the execution times for the Pauli X, Pauli Y, Z, and ZZ feature maps were recorded as 36290, 39853, 38588, and 73954 seconds, respectively.

4. Results and Discussion

In this study, classical and quantum support vector machines (SVM, QSVM, Pegasos-QSVM) were compared using the Indian Liver Patient, Breast Cancer Coimbra, and Heart Failure datasets. The evaluations were based on classification accuracy, considering various feature sizes (3–9), different quantum feature maps (Pauli X, Pauli Y, Z, ZZ), and circuit repetition counts (Reps=1 and Reps=2).

The results demonstrate that the QSVM algorithm generally provides more consistent and higher accuracy rates compared to classical SVM and Pegasos-QSVM. Especially in the Indian Liver Patient and Heart Failure datasets, QSVM often achieved accuracy values equal to or higher than those of classical SVM. This indicates that QSVM performs effectively on datasets with limited feature dimensions. On the other hand, the performance of the Pegasos-QSVM algorithm showed more variability and appeared to be more sensitive to the choice of feature map and qubit configuration. Although high accuracy values were obtained in some combinations (e.g., 5 qubits with Reps=2 in the Breast Cancer dataset), the overall performance was less consistent compared to QSVM. This suggests that Pegasos-QSVM requires more careful tuning of hyperparameters to achieve optimal results. The

*Sorumlu Yazar/Corresponding Author: gamzepelinaksoy@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 80 – 93, 2025
DOI: 10.62301/usmtd.1716034

classical SVM model, meanwhile, remained competitive particularly with a higher number of features. However, QSVM, benefiting from the quantum kernel advantage, generally outperformed classical SVM across different settings.

In conclusion, the QSVM algorithm stands out as an effective quantum machine learning approach for small to mid-sized datasets, while the Pegasos-QSVM algorithm may require further optimization to fully realize its potential. Additionally, beyond classification accuracy, the runtime analysis revealed significant insights into the practical applicability of the models. While QSVM consistently delivered higher accuracy, it incurred the longest training and prediction times due to the computational burden of full quantum kernel evaluations. In contrast, Pegasos-QSVM demonstrated a notable advantage in terms of speed, requiring substantially less computation time. This efficiency stems from its ability to avoid full kernel matrix construction by selectively computing only necessary kernel evaluations. Therefore, Pegasos-QSVM may be considered a more practical choice in scenarios where computational resources or execution time are limited, despite its relatively lower and less stable accuracy.

In future work, hyperparameter optimization techniques (such as grid search or Bayesian optimization) can be employed to enhance the classification accuracy of the Pegasos-QSVM algorithm. Moreover, experimenting with combinations of different quantum feature maps or implementing hybrid classical-quantum kernel strategies may help improve the balance between accuracy and stability. These advancements could make Pegasos-QSVM a more competitive alternative in terms of both speed and predictive performance.

5. References

[1] T.B. Alakus, M. Baykara, Comparison of Monkeypox and Wart DNA Sequences with Deep Learning Model, Applied Sciences 12 (2022) 10216. https://doi.org/10.3390/app122010216.

[2] Ö. Yildirim, A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification, Computers in Biology and Medicine 96 (2018) 189–202. https://doi.org/10.1016/j.compbiomed.2018.03.016.

[3] N. Jeyaraman, M. Jeyaraman, S. Yadav, S. Ramasubramanian, S. Balaji, Revolutionizing Healthcare: The Emerging Role of Quantum Computing in Enhancing Medical Technology and Treatment, Cureus (2024). https://doi.org/10.7759/cureus.67486.

[4] R. Ur Rasool, H.F. Ahmad, W. Rafique, A. Qayyum, J. Qadir, Z. Anwar, Quantum Computing for Healthcare: A Review, Future Internet 15 (2023) 94. https://doi.org/10.3390/fi15030094.

[5] Z. Li, Analysis of the Principles of Quantum Computing and State-of-the-Art Applications, Theoretical and Natural Science 41 (2024) 65–71. https://doi.org/10.54254/2753-8818/41/2024CH0155.

[6] D. Dhinakaran, L. Srinivasan, S.M. Udhaya Sankar, D. Selvaraj, Quantum-based privacy-preserving techniques for secure and trustworthy internet of medical things an extensive analysis, QIC 24 (2024) 227–266. https://doi.org/10.26421/QIC24.3-4-3.

[7] A.M. Dalzell, S. McArdle, M. Berta, P. Bienias, C.-F. Chen, A. Gilyén, C.T. Hann, M.J. Kastoryano, E.T. Khabiboulline, A. Kubica, G. Salton, S. Wang, F.G.S.L. Brandão, Quantum algorithms: A survey of applications and end-to-end complexities, (2023). https://doi.org/10.48550/arXiv.2310.03011.

[8] T.M. Khan, A. Robles-Kelly, Machine Learning: Quantum vs Classical, IEEE Access 8 (2020) 219275–219294. https://doi.org/10.1109/ACCESS.2020.3041719.

[9] P. Lamichhane, D.B. Rawat, Quantum Machine Learning: Recent Advances, Challenges, and Perspectives, IEEE Access 13 (2025) 94057–94105. https://doi.org/10.1109/ACCESS.2025.3573244.

[10] V. Havlíček, A.D. Córcoles, K. Temme, A.W. Harrow, A. Kandala, J.M. Chow, J.M. Gambetta, Supervised learning with quantum-enhanced feature spaces, Nature 567 (2019) 209–212. https://doi.org/10.1038/s41586-019-0980-2.

*Sorumlu Yazar/Corresponding Author: gamzepelinaksoy@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 80 – 93, 2025
DOI: 10.62301/usmtd.1716034

[11] R. Guido, S. Ferrisi, D. Lofaro, D. Conforti, An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review, Information 15 (2024) 235. https://doi.org/10.3390/info15040235.

[12] A. Kodipalli, S. Devi, Prediction of PCOS and Mental Health Using Fuzzy Inference and SVM, Front. Public Health 9 (2021). https://doi.org/10.3389/fpubh.2021.789569.

[13] H.F. Kareem, M.S. AL-Huseiny, F.Y. Mohsen, E.A. Khalil, Z.S. Hassan, Evaluation of SVM performance in the detection of lung cancer in marked CT scan dataset, Indonesian Journal of Electrical Engineering and Computer Science 21 (2021) 1731–1738. https://doi.org/10.11591/ijeecs.v21.i3.pp1731-1738.

[14] P.S. Oztekin, O. Katar, T. Omma, S. Erel, O. Tokur, D. Avci, M. Aydogan, O. Yildirim, E. Avci, U.R. Acharya, Comparison of Explainable Artificial Intelligence Model and Radiologist Review Performances to Detect Breast Cancer in 752 Patients, Journal of Ultrasound in Medicine 43 (2024) 2051–2068. https://doi.org/10.1002/jum.16535.

[15] K. Ramu, S. Patthi, Y.N. Prajapati, J.V.N. Ramesh, S. Banerjee, K.B.V.B. Rao, S.I. Alzahrani, R. ayyasamy, Hybrid CNN-SVM model for enhanced early detection of Chronic kidney disease, Biomedical Signal Processing and Control 100 (2025) 107084. https://doi.org/10.1016/j.bspc.2024.107084.

[16] A. Tudisco, D. Volpe, G. Turvani, Quantum Machine Learning in Healthcare: Evaluating QNN and QSVM Models, (2025). https://doi.org/10.48550/arXiv.2505.20804.

[17] R. Khushal, D.U. Fatima, Fuzzy quantum machine learning (FQML) logic for optimized disease prediction, Computers in Biology and Medicine 192 (2025) 110315. https://doi.org/10.1016/j.compbiomed.2025.110315.

[18] D. Maheshwari, U. Ullah, P.A.O. Marulanda, A.G.-O. Jurado, I.D. Gonzalez, J.M.O. Merodio, B. Garcia-Zapirain, Quantum Machine Learning Applied to Electronic Healthcare Records for Ischemic Heart Disease Classification, Human-Centric Computing and Information Sciences 13 (2023) 1–15. https://doi.org/10.22967/HCIS.2023.13.006.

[19] J. P, S. Hariharan, V. Madhivanan, S. N, M. Krisnamoorthy, A.K. Cherukuri, Enhanced QSVM with elitist non-dominated sorting genetic optimisation algorithm for breast cancer diagnosis, IET Quantum Communication 5 (2024) 384–398. https://doi.org/10.1049/qtc2.12113.

[20] S. Chatterjee, A. Das, An ensemble algorithm using quantum evolutionary optimization of weighted type-II fuzzy system and staged Pegasos Quantum Support Vector Classifier with multi-criteria decision making system for diagnosis and grading of breast cancer, Soft Comput 27 (2023) 7147–7178. https://doi.org/10.1007/s00500-023-07939-x.

[21] Y. Nasir, K. Kadian, V. Kumar, A. Wary, Harnessing Quantum Computing: A Comparative Study in Skin Disease Detection with Traditional ML, in: T. Senjyu, C. So–In, A. Joshi (Eds.), Smart Trends in Computing and Communications, Springer Nature, Singapore, 2024: pp. 361–370. https://doi.org/10.1007/978-981-97-1323-3_30.

[22] N. Singh, S.R. Pokhrel, Modeling Quantum Machine Learning for Genomic Data Analysis, (2025). https://doi.org/10.48550/arXiv.2501.08193.

[23] M. Munshi, R. Gupta, N.K. Jadav, S. Tanwar, A. Nair, D. Garg, Quantum Machine Learning-based Lung Cancer Prediction Framework for Healthcare 4.0, in: 2024 Asia Pacific Conference on Innovation in Technology (APCIT), 2024: pp. 1–6. https://doi.org/10.1109/APCIT62007.2024.10673456.

[24] FidelityQuantumKernel - Qiskit Machine Learning 0.8.2, (n.d.). <https://qiskit-community.github.io/qiskit-machine-learning/stubs/qiskit_machine_learning.kernels.FidelityQuantumKernel.html#> (accessed 15.06.2025).

[25] J.P. Miguel Patrcio, Breast Cancer Coimbra, (2018). https://doi.org/10.24432/C52P59.

[26] N.V. Bendi Ramana, ILPD (Indian Liver Patient Dataset), (2022). https://doi.org/10.24432/C5D02C.

[27] Unknown, Heart Failure Clinical Records, (2020). https://doi.org/10.24432/C5Z89R.

***Sorumlu Yazar/Corresponding Author:** gamzepelinaksoy@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 80 – 93, 2025
DOI: 10.62301/usmtd.1716034

[28] I. Straw, H. Wu, Investigating for bias in healthcare algorithms: a sex-stratified analysis of supervised machine learning models in liver disease prediction, BMJ Health Care Inform 29 (2022) e100457. https://doi.org/10.1136/bmjhci-2021-100457.

[29] M. Patrício, J. Pereira, J. Crisóstomo, P. Matafome, M. Gomes, R. Seiça, F. Caramelo, Using Resistin, glucose, age and BMI to predict the presence of breast cancer, BMC Cancer 18 (2018) 29. https://doi.org/10.1186/s12885-017-3877-1.

[30] D. Chicco, G. Jurman, Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone, BMC Med Inform Decis Mak 20 (2020) 16. https://doi.org/10.1186/s12911-020-1023-5.

[31] I.T. Jolliffe, J. Cadima, Principal component analysis: a review and recent developments, Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences 374 (2016) 20150202. https://doi.org/10.1098/rsta.2015.0202.

[32] G.T. Reddy, M.P.K. Reddy, K. Lakshmanna, R. Kaluri, D.S. Rajput, G. Srivastava, T. Baker, Analysis of Dimensionality Reduction Techniques on Big Data, IEEE Access 8 (2020) 54776–54788. https://doi.org/10.1109/ACCESS.2020.2980942.

[33] Z. Özpolat, Kuantum Tabanli Boyut İndirgeme ve Siniflandirici Gerçekleştirilmesi, MSc Thesis, Firat University, Elazig, Turkey, 2023.

[34] M. Kaur, K. Jain, A. Singla, K. Kadian, Quantum Exploration in Ransomware Detection with Conventional Machine Learning Approaches, in: 2024 IEEE International Conference on Contemporary Computing and Communications (InC4), IEEE, 2024, pp. 1–8. https://doi.org/10.1109/InC460750.2024.10649082.

[35] V.N. Vapnik, The Nature of Statistical Learning Theory, Springer, New York, NY, 2000. https://doi.org/10.1007/978-1-4757-3264-1.

[36] K.C. Chua, V. Chandran, U.R. Acharya, C.M. Lim, Application of Higher Order Spectra to Identify Epileptic EEG, Journal of Medical Systems 35 (2011) 1563–1571. https://doi.org/10.1007/s10916-010-9433-z.

[37] D. Anguita, S. Ridella, F. Rivieccio, R. Zunino, Quantum optimization for training support vector machines, Neural Networks 16 (2003) 763–770. https://doi.org/10.1016/S0893-6080(03)00087-X.

[38] A. Zeguendry, Z. Jarir, M. Quafafou, Quantum Machine Learning: A Review and Case Studies, Entropy 25 (2023) 287. https://doi.org/10.3390/e25020287.

[39] M. Aly, S. Fadaaq, O.A. Warga, Q. Nasir, M.A. Talib, Experimental Benchmarking of Quantum Machine Learning Classifiers, in: 2023 6th International Conference on Signal Processing and Information Security (ICSPIS), 2023: pp. 240–245. https://doi.org/10.1109/ICSPIS60075.2023.10343811.

[40] A. Thomsen, Comparing Quantum Neural Networks and Quantum Support Vector Machines, MSc Thesis, ETH Zurich, 2021, 97 p. https://doi.org/10.3929/ETHZ-B-000527559.

[41] S. Altares-López, A. Ribeiro, J.J. García-Ripoll, Automatic design of quantum feature maps, Quantum Sci. Technol. 6 (2021) 045015. https://doi.org/10.1088/2058-9565/ac1ab1.

[42] A. Daspal, OptiPauli: An algorithm to find a near-optimal Pauli Feature Map for Quantum Support Vector Classifiers, in: 2022 IEEE International Conference on Quantum Computing and Engineering (QCE), 2022: pp. 828–830. https://doi.org/10.1109/QCE53715.2022.00133.

***Sorumlu Yazar/Corresponding Author:** gamzepelinaksoy@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 94 – 106, 2025
DOI: 10.62301/usmtd.1707321

Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi
International Journal of Sustainable Engineering and Technology
ISSN: 2618-6055 / Cilt: 9, Sayı: 1, (2025), 94 – 106
Araştırma Makalesi – Research Article

INTERACTIVE X-RAY TUBE SIMULATION: A TOOL FOR BIOMEDICAL EDUCATION

Ali AKYÜZ[1,2*], Durmuş TEMİZ[1]

[1]Burdur Mehmet Akif Ersoy Üniversitesi, Bucak Emin Gülmez Teknik Bilimler Meslek Yüksekokulu, Burdur
[2]Burdur Mehmet Akif Ersoy Üniversitesi, Bucak Bilgisayar ve Bilişim Fakültesi, Burdur

(Geliş/Received: 27.05.2025, Kabul/Accepted: 20.06.2025, Yayınlanma/Published: 30.06.2025)

ABSTRACT

This paper presents an interactive X-ray Tube Simulator designed for students in biomedical engineering and biomedical device technology. Developed using HTML, CSS, and JavaScript and integrating the Chart.js library for dynamic graphical representations, this simulator offers dynamic visualizations and real-time graphical analysis to understand X-ray tube operation comprehensively. The simulator allows users to select tube voltage (kV), filament current (mA), exposure time (ms), anode angle (°), anode rotation speed (RPM), anode material (Tungsten, Molybdenum, Rhodium), cathode thickness (mm), filter thickness (mm), filter material (Aluminum, Copper, Molybdenum), collimator aperture (mm), chamber distance (cm), glass type (Lead, Borosilicate) and cooling type (Oil, Water), Air) and instantly observe their effects on X-ray spectra, radiation dose, heat accumulation and energy efficiency. This tool aims to bridge the gap between theoretical knowledge and practical application, promoting a deeper understanding of X-ray physics and safety considerations in a safe and cost-effective environment.

Keywords: Simulation, X-ray, Education, Biomedical

ETKİLEŞİMLİ X-IŞINI TÜPÜ SİMÜLASYONU: BİYOMEDİKAL EĞİTİM İÇİN BİR ARAÇ

ÖZ

Bu çalışma, biyomedikal mühendisliği ve biyomedikal cihaz teknolojisi öğrencileri için tasarlanmış etkileşimli bir X-Işını Tüpü Simülatörü sunmaktadır. HTML, CSS ve JavaScript kullanılarak geliştirilen ve dinamik grafiksel gösterimler için Chart.js kütüphanesini entegre eden bu simülatör, X-ışını tüpü operasyonunun kapsamlı bir şekilde anlaşılmasını sağlamak amacıyla dinamik görselleştirmeler ve gerçek zamanlı grafik analizleri sunar. Simülatör, kullanıcıların tüp voltajı (kV), flaman akımı (mA), maruz kalma süresi (ms), anot açısı (°), anot dönüş hızı (RPM), anot materyali (Tungsten, Molibden, Rodyum), katot kalınlığı (mm), filtre kalınlığı (mm), filtre materyali (Alüminyum, Bakır, Molibden), kolimatör açıklığı (mm), oda mesafesi (cm), cam türü (Kurşun, Borosilikat) ve soğutma türü (Yağ, Su, Hava) gibi tüm temel parametreleri manipüle etmelerine ve bunların X-ışını spektrumları, radyasyon dozu, ısı birikimi ve enerji verimliliği üzerindeki etkilerini anında gözlemlemelerine olanak tanır. Bu araç teorik bilgi ile pratik uygulama arasındaki boşluğu doldurmayı, güvenli ve uygun maliyetli bir ortamda X-ışını fiziği ve güvenlik hususlarının daha derinlemesine anlaşılmasını teşvik etmeyi amaçlamaktadır.

Anahtar kelimeler: Simülasyon, X-ışını, Eğitim, Biyomedikal

***Sorumlu Yazar/Corresponding Author:** aakyuz@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 94 – 106, 2025
DOI: 10.62301/usmtd.1707321

1. Introduction

X-ray technology revolutionized the world of medicine with a chance discovery by Wilhelm Conrad Röntgen in the late 19th century. These high-energy rays of the electromagnetic spectrum serve as a critical tool for non-invasive imaging of the human body. After Röntgen's wife took the first X-ray image, this technology became a unique tool for visualizing the internal structure of organs and tissues. Thanks to X-ray equipment that emits invisible rays, it has become possible to examine our bones and the anomalies in the internal organs in detail. X-rays are one of the most widely used tools for diagnosing diseases, from oncology to occupational health. However, it is important to remember both the potential dangers and the power of X-rays [1-4]. Long-term cell damage and increased risk of cancer can result from high doses of X-ray exposure. X-ray applications should always be carried out cautiously, and unnecessary exposure should be avoided [5]. The amazing discovery of X-rays has saved millions of lives and enabled new ways to diagnose and treat diseases. X-rays are absorbed at certain rates as they pass through solids. The density and atomic structure of the material affect this absorption rate. For example, elements with high atomic numbers absorb more X-rays than soft tissues. These differences create different gradations in X-ray images and give important details about, for example, a patient's internal anatomy. Each tissue appears in a different color, as if touched with a paintbrush; details cannot be missed [6]. X-rays are generated when high-energy electrons collide with a target material, typically tungsten. The electrons lose energy during this collision, and X-rays are generated. These rays are collected on a detector or film to form an image, much like the lens of a camera [7]. To optimize the imaging process, the detector system and X-ray instrument design significantly impact image quality. The X-rays ' wavelength can vary depending on the materials and the targeted area's nature [8]. X-ray technology is becoming increasingly more reliable and efficient thanks to advances in digital systems and image processing methods. With digital X-ray systems providing faster and clearer images at lower doses, patients are exposed to less radiation. Furthermore, by facilitating remote access and storage, image digitization is raising the standard of medical services [9].

In modern healthcare, trained personnel in the biomedical field are very important for device installation, use, repair, and technical service. This field requires careful transfer of scientific knowledge and technical skills in the healthcare sector. It is of great importance in patient safety, treatment effectiveness, and the training of future specialists. Innovations in technology, especially in medical devices, require training that will adapt over time. An important part of this learning process includes simulations [10]. The disadvantages of traditional educational techniques are overcome by simulations that reduce the uncertainty of the real clinical environment, reduce the possibility of error, and provide students with a safe learning environment [11]. In particular, X-ray machine simulations help students understand imaging methods, develop decision-making skills, and understand how medical devices work. Moreover, giving them the chance to practice possible challenges they may encounter during the learning process is a wise investment that will raise the standard of healthcare services in the future. Computer-based simulations are usually conducted through relevant software, allowing users to improve their ability to use X-ray equipment in a virtual environment. The major advantage of such simulations is the huge savings for users in terms of time and cost. Students who learn about the complex settings of X-ray equipment must learn without experiencing errors in the real environment.

In this study, we developed an interactive X-Ray Tube Simulator designed as an educational tool for biomedical engineering and biomedical device technology students. Built using HTML, CSS, and JavaScript, with the Chart.js library for dynamic graphical representations, the simulator integrates dynamic visualizations and real-time graphical analyses to understand X-ray tube operation comprehensively. By allowing users to manipulate key parameters and observe their effects on X-ray spectra, radiation dose, heat accumulation, and energy efficiency, this tool aims to bridge the gap between theoretical knowledge and practical application, fostering a deeper understanding of X-ray physics and safety considerations in a safe and cost-effective environment.

2. Materials and Methods

The simulator was designed with comprehensive and interactive features to provide biomedical device technology engineering and electronics technician students with information about the operating principles of the X-ray tube.

*Sorumlu Yazar/Corresponding Author: aakyuz@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 94 – 106, 2025
DOI: 10.62301/usmtd.1707321

The user interface was created using HTML5 and CSS3 technologies in a contemporary and interactive manner. While CSS3 controls the interface's visual design and layout, HTML5 ensures that the page structure is arranged meaningfully and systematically. Within the HTML structure, basic elements like canvas elements for graphics, parameter controls (sliders and selection elements), and information boxes for summarizing information are arranged logically in groups.

A responsive and flexible design that adjusts to the simulator's various screen sizes uses contemporary layout techniques like CSS, flexbox, and grid. The color scheme, typefaces, and animation effects have been carefully chosen to improve the user experience and present information more clearly. This makes it simple for users to monitor results and change parameters visually.

JavaScript's ability to dynamically control parameter elements is one of the main features that allows user interaction in the simulator. Sliders representing various settings such as tube voltage (kV), filament current (mA), and exposure time (ms) are retrieved from the HTML structure using the document.getElementById() method and grouped into a JavaScript object (e.g., sliders).

The tab system in the user interface is designed to provide easy access to different analysis and visualization sections - for example, X-ray spectrum, dose analysis, heat analysis, energy analysis, and statistics. In this system, created with JavaScript, when the user clicks on a tab, the content of that tab appears on the screen. To achieve this, the querySelectorAll() method is used to select and loop over all tabs and content sections. While the “active” class is added to the clicked tab, other tabs and content are stripped of this class. Thus, users can easily navigate complex data and quickly access the necessary information.

The HTML5 Canvas API is used to visually represent the physical processes and parts of the X-ray tube. The 2D drawing context (getContext(“2d”)), obtained with JavaScript, enables the animation of dynamic processes such as the movement of electrons from cathode to anode, X-ray generation, anode rotation, and heating. Each physical component (cathode, anode, glass tube, collimator, filtration, etc.) is drawn in specific coordinates and with appropriate colors; visual effects such as color change according to the temperature of the anode are applied.

A realistic animation is achieved by moving electrons and X-rays according to speed and direction parameters. Parameters such as the rotation speed of the anode and temperature are also integrated into the animation. This approach allows students to directly observe the physical phenomena inside the X-ray tube and the effects of parameter changes, and makes abstract concepts concrete.

For the simulator to produce accurate and realistic results, it is critical to accurately describe the physical and electrical properties of the materials used (anode, cathode, glass, filtration, cooling). For this purpose, each material's atomic number (Z), K_α and K_β characteristic energy values, density, heat capacity, melting point, work function, emission efficiency, etc., are stored in JavaScript objects. For example, different properties such as emission efficiency are defined for anode materials (Tungsten, Molybdenum, Rhodium) and similarly for cathode materials (Tungsten, Tungsten with Thorium). This data is used in various calculations such as X-ray spectrum calculations, dose analysis, heat accumulation simulations, and energy efficiency evaluations, ensuring that the simulation produces results close to physical reality.

The simulator applies physical models to analyze the X-ray spectrum. The relative intensity of the beam, electron energy, and filtration effects primarily modulate the Bremsstrahlung spectrum, which is calculated over the energy range by considering the tube voltage (kV) and other factors. The X-ray intensity I is proportional to the electron energy E and the maximum energy Emax and is expressed by $I(E) \propto E(_{Emax}\text{-}E)$. This formula implies that part of the kinetic energy of the electrons is converted into X-rays, and the spectrum is shaped depending on the energy [12]. As the X-ray beam passes through the filtration material, low-energy photons are absorbed. This effect is modeled by the Lambert-Beer law (Equation 1). $I_{filter}(E)$ is the intensity after filtration, $\mu(E)$ is the energy-dependent absorption coefficient that varies with the material, and x is the filtration thickness. In the simulator, the filtration coefficient is approximated by a constant coefficient concerning the material and thickness. The X-ray beam also passes through obstacles such as the glass tube and the cathode thickness. The absorption coefficients similarly model these effects in Equations (2) and (3), where T_{glass} is the transmission coefficient of the

*Sorumlu Yazar/Corresponding Author: aakyuz@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 94 – 106, 2025
DOI: 10.62301/usmtd.1707321

glass, the absorption coefficient depends on the cathode thickness, d_{cat} is the cathode thickness. E is the energy [13, 14].

$$I_{filter}(E) = I(E)e^{-\mu(E)x} \quad (1)$$

$$I_{glass}(E) = I_{filter}(E)T_{glass} \quad (2)$$

$$I_{cat}(E) = I_{glassr}(E)e^{-\alpha dcat\ E} \quad (3)$$

Focusing cup voltage and tube current fluctuation affect the amplitude of the spectrum. The simulator expresses these effects by coefficients (V_{focus}: Focusing cup voltage, ripple: Percentage of current ripple, k_{focus}: Focusing effect coefficient). It is modeled by Equation (4). Depending on the atomic number of the anode material, characteristic K_α and K_β lines appear at certain energy levels. These lines are added to the spectrum when the tube voltage exceeds these energy levels [15].

$$I_{final}(E) = I_{cat}(E)\ (1+ k_{focus}V_{focus})(1-(ripple/100) + random\ fluctuation) \quad (4)$$

These formulas are calculated in JavaScript functions in the energy range to generate spectrum data. The estimated values are exported to Chart.js graphs and updated in real time as user parameters change. Thus, students can numerically and visually experience how physical parameters shape the X-ray spectrum.

The dose calculation was developed to model the radiation effect of X-ray tube parameters on the patient and the environment. The room output dose is calculated by the interaction of variables such as tube voltage (kV), filament current (mA), exposure time (ms), distance between the patient and the tube (cm), and atomic number of the anode material (Z) (Equation (5)) [16].

$$D = \frac{kV.mA.ms.Z}{dist^2.1000} f_{filter}.f_{collimator} \quad (5)$$

$$H = \frac{kV.mA.ms}{1400} f_{dia}.f_{ref} \quad (6)$$

D is the dose (μGy), dist is the distance (cm), and f_{filter} and $f_{collimator}$ are the coefficients representing the filtration and collimator effects. Filtration reduces the dose by absorbing low-energy photons, while a collimator affects the dose distribution by shaping the beam. The patient dose was modeled as approximately 70% of the room exit dose.

The simulator records dose values as a time series and visualizes them with dynamic charts using the Chart.js library. These charts show in real time how the dose changes over the exposure time, and depending on parameter changes. In addition, statistical data such as average dose, maximum dose, and dose rate are calculated and presented to the user. In this way, students can numerically and visually analyze the relationship between dose and physical parameters.

The heat generated at the anode and the temperature variation are important parameters in the safe operation of the X-ray tube. Heat accumulation is calculated depending on factors such as tube voltage (kV), filament current (mA), exposure time (ms), anode material, and anode diameter (Equation 7) [17].

$$H = \frac{kV.mA.ms}{1400} f_{dia}.f_{ref} \quad (7)$$

H is the heat accumulation (kHU), f_{dia} is a coefficient representing the effect of the anode diameter, and fref represents the cooling type's effectiveness (oil, water, air). The anode temperature T = 25 + (100 H/C) (°C) is calculated considering the heat accumulation and the heat capacity of the anode material. T is the anode temperature, C is the heat capacity of the anode material, and 25 is the initial temperature. The simulator calculates the anode temperature in real time and visualizes the heat accumulation with a color scale.

*Sorumlu Yazar/Corresponding Author: aakyuz@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 94 – 106, 2025
DOI: 10.62301/usmtd.1707321

Heat accumulation and temperature values are recorded as time series and then displayed as graphs using the Chart.js library. This enables students to investigate in detail the thermal behavior of the anode and the effects of parameter changes on heat management. The simulator performs energy calculation and efficiency analysis to evaluate the X-ray tube's energy conversion processes and efficiency. Total energy (Joules) is calculated using tube voltage (kV), filament current (mA), and exposure time (ms) (Equation 8).

$$E_{total} = V.I.t \quad (8)$$

The X-ray efficiency (η) is proportional to the atomic number (Z) of the anode material and the tube voltage (kV) (Equation 9).

$$\eta = 9.10^{-10}.Z.\ kV.\ f_{filter}.\ f_{cat}.\ 100 \quad (9)$$

X-ray energy and heat are calculated using Equations (10) and (11).

$$E_{x\text{-}ray} = E_{total}.\ \eta/100 \quad (10)$$

$$E_q = E_{total} - E_{x\text{-}ray} \quad (11)$$

The code calculates the total energy by multiplying the tube voltage, current, and exposure time. The X-ray efficiency is determined by the formula using the atomic number of the anode material and the tube voltage; the effects of filtration and cathode material are also considered. The total energy is divided into X-ray energy and the remainder as heat energy according to the efficiency ratio. These values are visualized in pie and gauge charts using Chart.js and updated in the user interface. Thus, energy distribution and efficiency are tracked in real time.

The animation loop runs continuously with the requestAnimationFrame function and performs the following operations in each frame: static elements (glass tube, anode, cathode, etc.) are drawn, electrons and X-rays are moved and drawn, and the rotation angle of the anode is updated. Electron generation occurs randomly, depending on the filament current, and moving particles are cleaned up as they leave the screen. This cycle ensures that the simulation provides a real-time and fluid visual experience.

The interactive nature of the simulator is based on users changing parameters and instantly observing the results. This interaction is achieved through JavaScript event listeners. Event listeners, such as oninput or onchange, are assigned to each slider and select element. When the user changes a parameter, the corresponding event listener is triggered, and the refreshUI() function is called. This function updates all parameter values, redoes spectrum, dose, heat, and energy calculations, and the graphics. In addition, the trackParameterChanges() function is called to track the number of parameter changes and update the statistics. In this way, the effects of user changes on the simulation are instantly visualized, and an interactive learning experience is provided.

The function getHeatColor (temp, meltPoint) calculates a color scale based on the anode temperature and provides the color change of the anode in the animation, so that the temperature increase is visually expressed. The new Electron () function generates new electron particles emitted from the cathode, which are characterized by velocity, position, and phase, and are moved in the animation. The newXRay(x, y) function generates X-ray particles emitted from the anode surface, which move at random angles and velocities to increase the realism of the simulation. These functions enable the dynamic and interactive nature of the simulator, providing the user with a rich visual experience.

Table 1 lists some properties and constants in the code that are not directly adjustable parameters but affect the simulation's behavior.

*Sorumlu Yazar/Corresponding Author: aakyuz@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 94 – 106, 2025
DOI: 10.62301/usmtd.1707321

Table 1. Parameters and values used in the simulator [18]

Anode Materials	W (Tungsten)	Z (Atomic Number) = 74, K_α = 17.5 eV, K_β = 19.6 keV, Density = 10.2 g/cm^3, color = "#f1c40f" (Yellow), Heat Capacity = 300 j/kg^{0}C, Melting Point = 3422 ^{0}C
	Mo (Molybdenum)	Z (Atomic Number) =42, K_α = 59 eV, K_β = 67 keV, Density = 19.3 g/cm^3, color = "#2ecc71" (Green), Heat Capacity = 150 j/kg^{0}C, Melting Point = 2623 ^{0}C
	Rh (Rhodium)	Z (Atomic Number) = 45, K_α = 20.2 eV, K_β = 22.7 keV, Density = 12.4 g/cm^3, color = "#f1c40f" (Yellow), Heat Capacity = 180 j/kg^{0}C, Melting Point = 1964 ^{0}C
Cathode Materials	W (Tungsten)	Work Function = 4.5 eV Emission Efficiency = 1.0
	ThW (Tungsten with Thorium)	Work Function = 2.6 eV Emission Efficiency = 1.5
Glass Types	Lead	Transmittance = 0.85 Density = 4.5 g/cm^3
	Borosilicate Glass	Transmittance = 0.95 Density = 2.2 g/cm^3
Filter Materials	Al (Aluminum)	Linear Attenuation Coefficient = 0.15 Z (Atomic Number) = 13
	Cu (Copper)	Linear Attenuation Coefficient = 0.35 Z (Atomic Number) = 29
	Mo (Molybdenum)	Linear Attenuation Coefficient = 0.45 Z (Atomic Number) = 42
Cooling Types	Oil	Efficiency: 1.0 Maximum Cooling Rate = 10 kHU/s
	Water	Efficiency: 1.5 Maximum Cooling Rate = 15 kHU/s
	Air	Efficiency: 0.7 Maximum Cooling Rate = 5 kHU/s
Chart Settings	maxDataPoints	Maximum number of data points to be shown in time series charts = 20
Initial Anode Temperature	25 °C	

3. Results and Discussion

The simulator provides dynamic visualizations and illustrations that improve comprehension of the operation of the X-ray tube. The electron beam animation shows the trajectory of electrons from the cathode to the anode and precisely captures their movement, speed, and focusing effects. X-ray production and propagation are animated to show how electrons interacting with the anode generate rays emitted at various angles and speeds, allowing users to track the direction and intensity of the X-ray beam visually. The anode's rotation is simulated based on its RPM value, with heat accumulation on its surface represented through color changes, which helps comprehend the risk of overheating. Additionally, filtration thickness, material, and collimator aperture are physically represented in the animation, making the shaping and filtering of the X-ray beam visually intuitive.

*Sorumlu Yazar/Corresponding Author: aakyuz@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 94 – 106, 2025
DOI: 10.62301/usmtd.1707321

Regarding spectrum and dose analysis, the simulator computes and shows the Bremsstrahlung and characteristic X-ray spectra based on the anode material and tube voltage, allowing students to study the energy distribution and peak formation. Physical models calculate doses for patients and room exits in real time. Graphs show how dose changes with filtration, collimation, exposure time, and distance. Time series analysis tracks dose and heat accumulation over time, allowing users to analyze both immediate and long-term effects of parameter changes.

The heat and energy efficiency analysis includes calculating the anode's heat accumulation and temperature changes and visualizing how cooling type and anode diameter influence heat management. The distribution of electrical energy into X-ray production and heat loss is computed, providing insight into the device's efficiency. X-ray yield is calculated by considering anode material, filtration, and cathode properties and is displayed to aid in understanding efficiency factors.

Finally, the simulator incorporates statistical tracking and feedback features. It monitors how frequently students adjust each parameter, supporting evaluation of the experimental process. Information on simulation duration and total accumulated dose is provided to raise awareness about safety and performance. Graphical summaries consolidate the effects of all parameters, enabling students to interpret and reflect on their experimental results easily. The parameter settings and controls are detailed in Table 2.

The screen output of the simulation can be seen as a single page. However, it is given in parts here. Figure 1 shows the X-ray generation animation and the selection area for tube parameters. All parameters can be changed on the screen. The X-ray spectrum screen output is given in Figure 2. The X-ray spectrum graph details the energy distribution of X-rays, with the x-axis representing a continuous energy range from 0 keV up to the tube voltage (e.g., 120 keV), and the y-axis indicating the relative radiation intensity. The spectrum is calculated based on several factors: Bremsstrahlung, which is the continuous spectrum resulting from electrons colliding with the anode and is modeled by the formula $I(E) \propto E(E_0 - E)$, where E is the energy and E_0 is the maximum energy equal to the tube voltage; characteristic K_α and K_β lines, which are intense peaks at specific energy levels dependent on the atomic structure of the anode material; the filtration effect, where the filter material and thickness reduce the low-energy portion of the spectrum by absorbing low-energy photons; cathode thickness and focusing voltage, which fine-tune the overall intensity and shape of the spectrum; and current ripple, which introduces small fluctuations in the spectrum. Changes in parameters affect the spectrum as follows: increasing the tube voltage raises the maximum energy of the spectrum and makes the characteristic lines more pronounced; increasing filtration reduces low-energy photons, cutting off the lower end of the spectrum; and changing the anode material alters the energies and intensities of the characteristic lines.

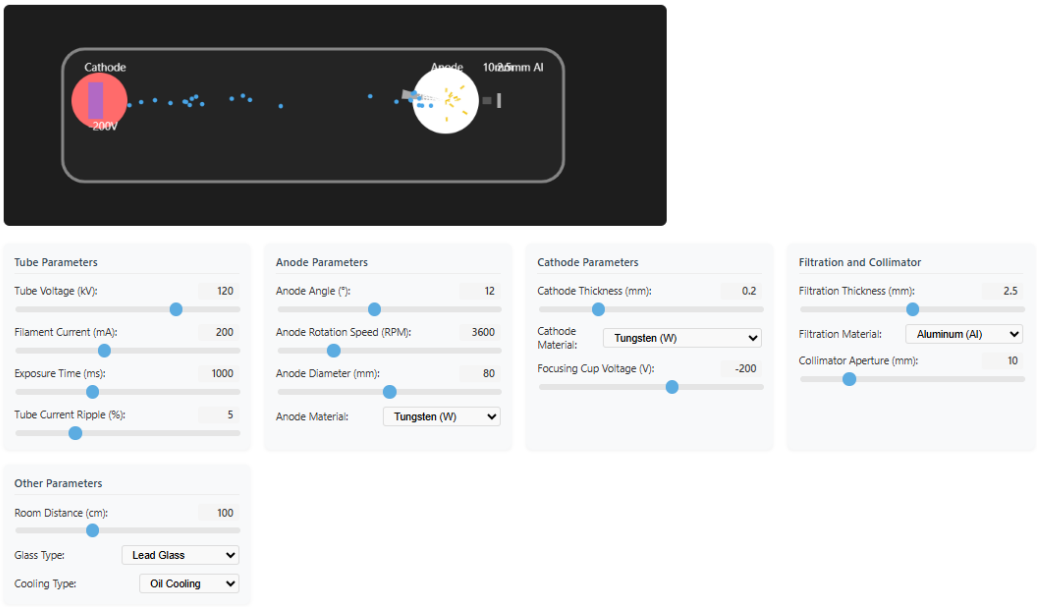


Figure 1. An animation showing the process of X-ray generation, along with a display of the parameter settings used

*Sorumlu Yazar/Corresponding Author: aakyuz@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 94 – 106, 2025
DOI: 10.62301/usmtd.1707321

Table 2. Parameter settings, controls, and description

Parameter	Min	Max	Step	Description
Tube Voltage (kV)	40	150	5	Determines the energy at which electrons strike the anode. Students can observe the effect of voltage variation on the X-ray spectrum and dose.
Filament Current (mA)	10	500	10	This parameter controls the number of electrons leaving the cathode. As the current increases, the electron density and, therefore, the X-ray production increase. With this parameter, the effect of current on dose and heat accumulation can be analyzed.
Exposure Time (ms)	10	3000	10	Determines the X-ray generation time. A long exposure time leads to an increase in the total dose. Students can experience the effect of exposure time on dose and energy consumption.
Anode Angle (°)	6	20	0.5	It is the inclination angle of the anode surface. The anode angle affects the focusing and heat dissipation of the X-ray beam. The simulator visualizes the effect of anode angle on instrument performance.
Anode Rotation Speed (RPM)	1000	1200	100	Determines the speed of the rotating anode. High speed allows heat to spread over the anode surface and prevents overheating. Students can understand the role of rotation speed on heat accumulation.
Anode Diameter (mm)	40	120	1	Refers to the physical size of the anode. Anode diameter is related to heat capacity and durability. With this parameter, the thermal behavior of the anode can be studied.
Cathode Thickness (mm)	0.1	0.5	0.01	These parameters are effective on the structural properties and electron emission efficiency of the cathode. The choice of cathode material is important in terms of operating temperature and efficiency.
Filtration Thickness (mm)	0	5	0.1	It allows filtering of low energy photons from the X-ray beam. Filtration maintains image quality while reducing patient dose. Students can experience the filtration effect of different materials and thicknesses.
Collimator Aperture (mm)	5	30	1	Allows shaping of the X-ray beam. Collimator aperture affects dose distribution and image area.
Room Distance (cm)	50	200	5	It is the distance between the X-ray tube and the patient. As the distance increases, the dose decreases. With this parameter, the change of the dose depending on the distance can be analyzed.
Focusing Cup Voltage	-500	0	10	It affects the focusing of the electron beam. Focusing voltage variation affects the intensity of the electron beam and thus X-ray production.
Glass Type and Refrigeration Type				The permeability of the tube glass and the cooling method of the anode are simulated to analyze the durability and performance of the device.

The dose analysis graph (screen output is shown in Figure 3) displays dose values over simulation time, with the x-axis representing time in seconds and the y-axis showing dose measured in microgray (μGy). Two lines are plotted: the room exit dose and the patient dose. Dose calculation depends on several parameters, including tube voltage, current, exposure time, distance, and filtration. The patient dose is modeled as approximately 70% of the room exit dose. The collimator aperture directly influences the dose, with larger apertures increasing the dose. Additionally, the dose decreases with increasing distance according to the inverse square law. Changes in parameters affect the dose as follows: increasing

*Sorumlu Yazar/Corresponding Author: aakyuz@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 94 – 106, 2025
DOI: 10.62301/usmtd.1707321

voltage, current, or exposure time raises the dose; increasing filtration thickness or distance reduces the dose; and widening the collimator aperture increases the dose. The dose graph responds in real time to parameter adjustments, showing dose accumulation over time.

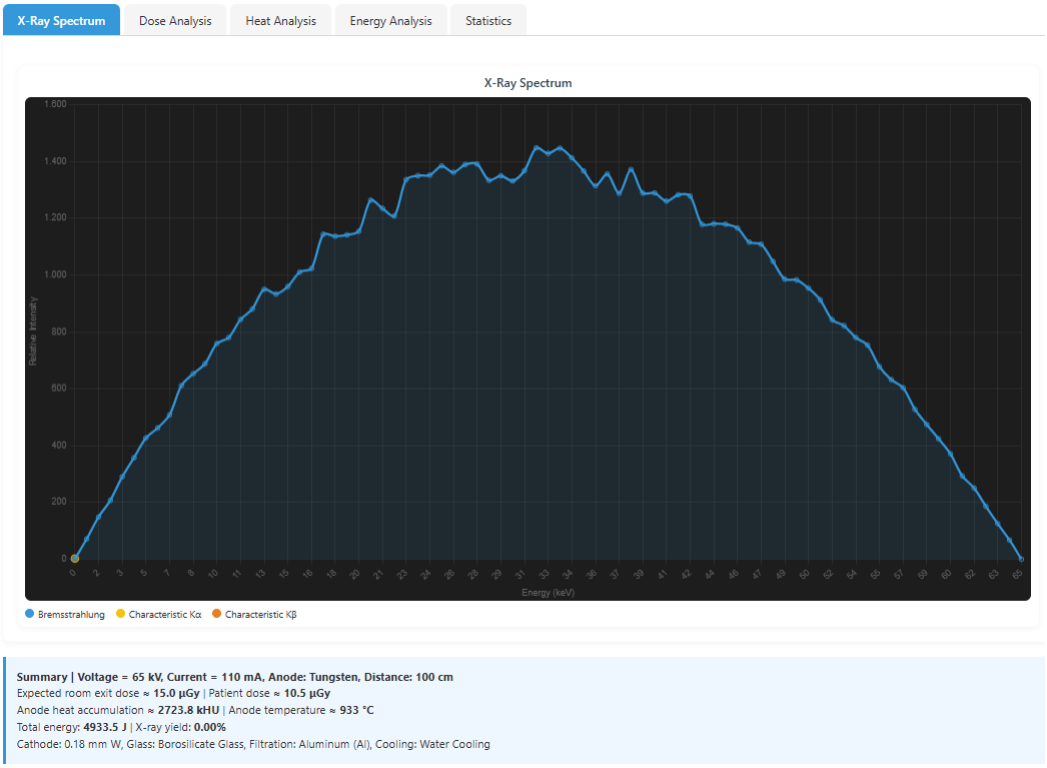


Figure 2. A screen output displaying the X-ray spectrum and its characteristics

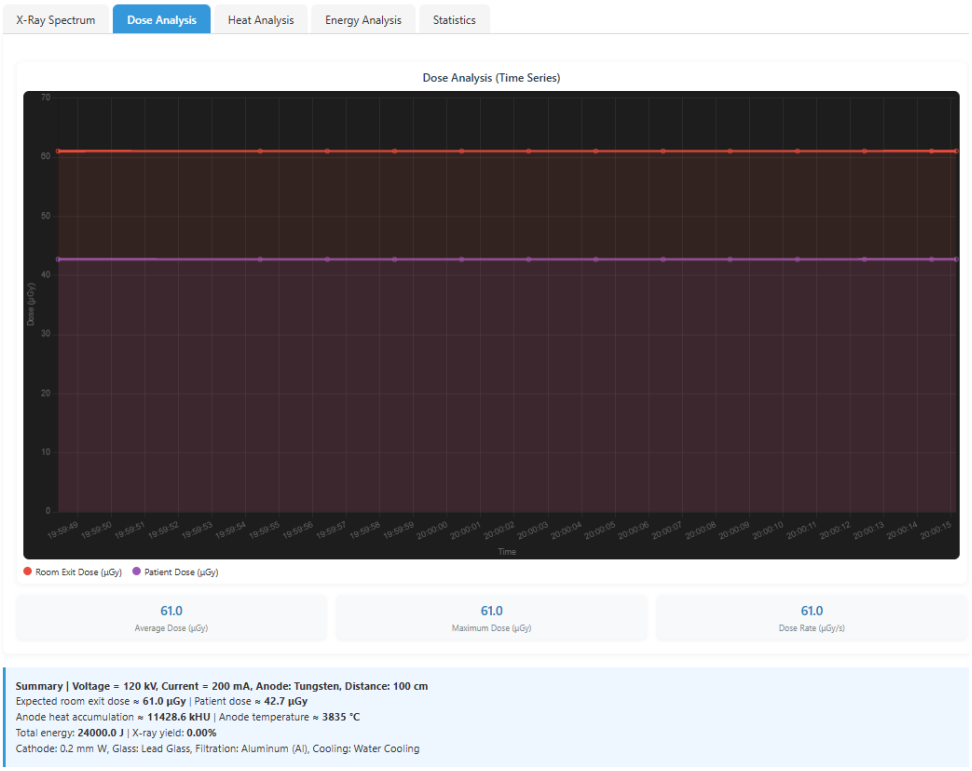


Figure 3. A screen output showing a graph of the dose analysis results, illustrating the distribution and levels of radiation dose

The heat analysis graph (screen output is shown in Figure 4) presents simulation time on the x-axis. It features two y-axes: the left y-axis shows heat accumulation measured in kilo heat units (kHU), while

*Sorumlu Yazar/Corresponding Author: aakyuz@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 94 – 106, 2025
DOI: 10.62301/usmtd.1707321

the right y-axis displays the anode temperature in degrees Celsius (°C). Heat accumulation depends on tube voltage, current, exposure time, anode diameter, and the type of cooling used. The anode temperature is calculated based on the accumulated heat and the anode's heat capacity. The anode's rotation speed and diameter influence how heat spreads and dissipates. The cooling method, whether oil, water, or air, determines the rate at which the anode cools as voltage, current, and exposure time increase, heat accumulation, and anode temperature rise. Conversely, increasing the anode's rotation speed and diameter reduces heat buildup and slows temperature increase. Changes in cooling type affect the maximum anode temperature and cooling rate. This graph visually enables monitoring of the anode's risk of overheating during operation.

Figure 4. A screen output displaying a graph of the heat analysis results, showing temperature distribution and thermal behavior.

The energy analysis (screen output is shown in Figure 5) includes two main visualizations: a pie chart showing the distribution of total energy into X-ray energy and heat energy, and a gauge chart indicating the percentage of X-ray efficiency. The total energy is calculated using the formula E=VIt. The X-ray efficiency η is computed by the formula $\eta = 9.10^{-10}.Z.\ kV$, where Z is the atomic number of the anode material and kV is the tube voltage; this value is adjusted based on filtration and cathode material effects. Heat energy is determined by subtracting the X-ray energy from the total energy. Parameter changes affect the analysis such that increasing voltage and selecting different anode materials raise X-ray efficiency, while filtration and cathode material also influence efficiency. Both the pie and gauge charts update in real time to reflect these changes.

*Sorumlu Yazar/Corresponding Author: aakyuz@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 94 – 106, 2025
DOI: 10.62301/usmtd.1707321

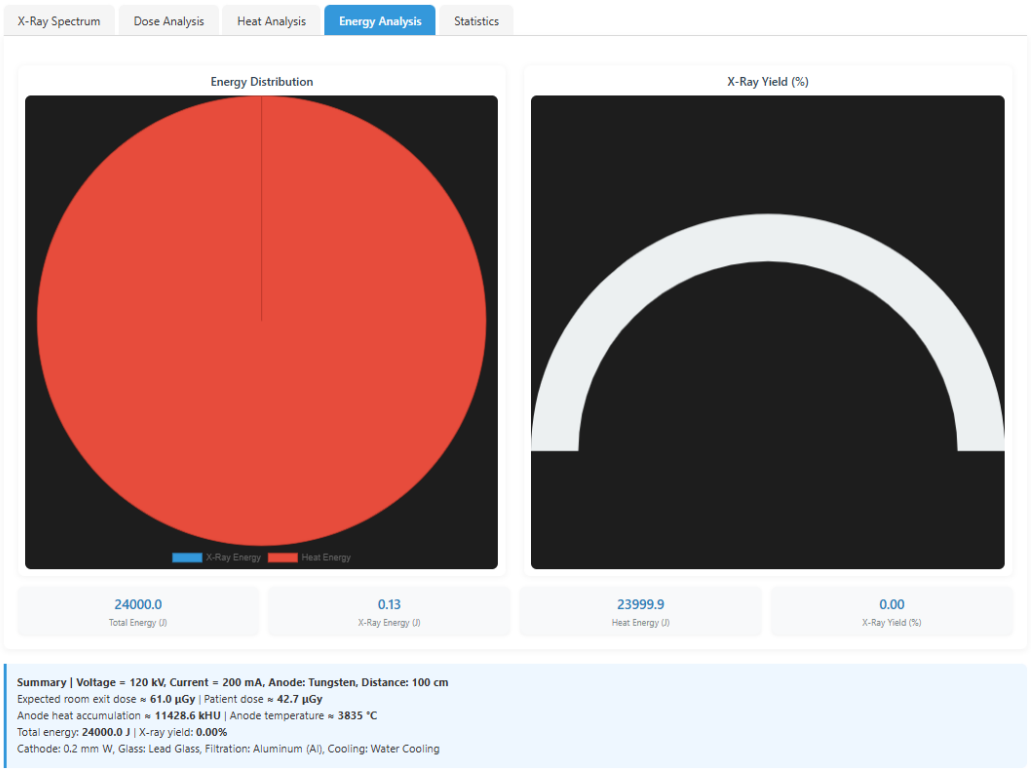


Figure 5. A screen output presenting the energy analysis results, including relevant data and visualizations

Figure 6. A screen output displaying a graph of statistical data, providing a visual representation of key metrics and trends

The statistics graph (screen output is shown in Figure 6) is a bar chart displaying the number of changes made to various parameters, including tube voltage (kV), filament current (mA), exposure time (ms), filtration, anode angle, and rotation speed (RPM). Alongside the chart, numerical indicators show the total number of parameter changes, the elapsed simulation time, and the accumulated total dose. Each time the user adjusts a parameter, the corresponding bar's value increases, providing a visual record of

*Sorumlu Yazar/Corresponding Author: aakyuz@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 94 – 106, 2025
DOI: 10.62301/usmtd.1707321

interaction frequency. Meanwhile, the simulation time and total dose values update in real time, offering continuous feedback on the experiment's duration and radiation exposure.

4. Conclusion

Simulator offers a robust and interactive educational tool specifically designed for students in biomedical engineering and biomedical device technology. The simulator bridges theoretical knowledge and practical understanding by combining realistic animations of electron beams, X-ray generation, and anode rotation with detailed, real-time graphical analyses of X-ray spectra, radiation dose, heat accumulation, and energy efficiency. The ability to manipulate critical parameters such as tube voltage, filament current, filtration, and anode properties allows students to explore the complex interdependencies within X-ray tube operation safely and cost-effectively, without the risks associated with physical equipment. From an educational perspective, the simulator enhances learning by providing immediate visual and quantitative feedback, fostering active experimentation and critical thinking. Statistical tracking promotes reflective learning and safety awareness by assisting teachers and students in monitoring participation and parameter exploration. Students' conceptual understanding of X-ray physics is strengthened by this tool, which also gets them ready for practical uses in medical imaging and device maintenance. Overall, the simulator is a great addition to biomedical device technology curricula because it encourages practical learning, enhances technical proficiency, and reinforces safety principles crucial for aspiring industry professionals. The simulations' source codes are available on GitHub (https://aliakyuz12.github.io/Interactive-X-Ray-Tube-Simulation/) and can be used for educational purposes.

5. References

[1] S.F. Keevil, Physics and medicine: a historical perspective, Lancet 379 (9825) (2012) 1517–1524.

[2] F. Nüsslin, Wilhelm Conrad Röntgen: The scientist and his discovery, Phys. Medica 79 (2020) 65–68.

[3] Y.M. Al-Worafi, Radiology and Medical Imaging Research in Developing Countries, in: Handbook of Medical and Health Sciences in Developing Countries: Education, Practice, and Research, Springer Int. Publ., Cham, 2024, pp. 1–29.

[4] J.G.J. Lftta, A.N.A.A. Zahra, A.H.J. Ashour, A.H.K. Nasser, A.N. Shanawa, X-Rays and Their Uses on The Human Body, Curr. Clin. Med. Educ. 2 (8) (2024) 91–115.

[5] M.S. Linet, et al., Cancer risks associated with external radiation from diagnostic imaging procedures, CA Cancer J. Clin. 62 (2) (2012) 75–100.

[6] M. Kurudirek, Effective atomic numbers and electron densities of some human tissues and dosimetric materials for mean energies of various radiation sources relevant to radiotherapy and medical applications, Radiat. Phys. Chem. 102 (2014) 139–146.

[7] S. Prabhu, D.K. Naveen, S. Bangera, B.S. Bhat, Production of x-rays using x-ray tube, J. Phys. Conf. Ser. 1712 (1) (2020) 012036.

[8] M.J. Yaffe, J.A. Rowlands, X-ray detectors for digital radiography, Phys. Med. Biol. 42 (1) (1997).

[9] C. Cao, et al., Emerging X-ray imaging technologies for energy materials, Mater. Today 34 (2020) 132–147.

[10] A.H. Zamzam, et al., A systematic review of medical equipment reliability assessment in improving the quality of healthcare services, Front. Public Health 9 (2021) 753951.

[11] A. Haleem, M. Javaid, R.P. Singh, R. Suman, Medical 4.0 technologies for healthcare: Features, capabilities, and applications, Internet Things Cyber-Phys. Syst. 2 (2022) 12–30.

[12] F. Albert, Principles and applications of x-ray light sources driven by laser wakefield acceleration, Phys. Plasmas 30 (5) (2023) 050902.

[13] M. Berger, Q. Yang, A. Maier, X-ray Imaging, in: Medical Imaging Systems: An Introductory Guide, 2018, pp. 119–145.

*Sorumlu Yazar/Corresponding Author: aakyuz@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 94 – 106, 2025
DOI: 10.62301/usmtd.1707321

[14] H. Jung, Basic physical principles and clinical applications of computed tomography, Prog. Med. Phys. 32 (1) (2021) 1–17.

[15] M.K.M. Alharbi, A. Zhou, M. Naunton, R. Davidson, C. Makanjee, Innovations in X-ray tube design and instrumentation for conventional radiological applications: a scoping review, Imaging Sci. J. (2025) 1–16.

[16] J.A. Seibert, X-ray imaging physics for nuclear medicine technologists. Part 1: Basic principles of x-ray production, J. Nucl. Med. Technol. 32 (3) (2004) 139–147.

[17] T. Bashore, Fundamentals of x-ray imaging and radiation safety, Catheter. Cardiovasc. Interv. 54 (1) (2001) 126–135.

[18] National Center for Biotechnology Information, PubChem Element Information, https://pubchem.ncbi.nlm.nih.gov/element (accessed 26.05.25).

*Sorumlu Yazar/Corresponding Author: aakyuz@mehmetakif.edu.tr

ISSN: 2618-6055 / 9, (1), 107 – 117, 2025
DOI:10.62301/usmtd.1716436

Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi
International Journal of Sustainable Engineering and Technology
ISSN: 2618-6055 / Cilt:9, Sayı:1, (2025), 107 – 117
Araştırma Makalesi – Research Article

SARGILI BETON BASINÇ DAYANIMININ YAPAY ZEKÂ VE OPTİMİZASYON TABANLI YAKLAŞIMLARLA MODELLENMESİ

*Abdullah GÜNDOĞAY[1]

[1]Süleyman Demirel Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, İnşaat Mühendisliği Bölümü, Isparta

(Geliş/Received: 10.06.2025, Kabul/Accepted: 23.06.2025, Yayınlanma/Published: 30.06.2025)

ÖZ

Betonarme elemanların kesit özellikleri, sargılı beton davranışını belirleyen temel faktörler arasında yer almakta olup, bu davranışın analitik hesaplara doğru biçimde yansıtılması büyük önem arz etmektedir. Bu kesit özelliklerine bağlı çok sayıda değişkenin belirlenmesi ve analiz edilmesi zaman alıcı olabilmektedir. Bu yüzden günümüzde birçok yapay zekâ tabanlı modelleme yöntemleri kullanılmaya başlanılmıştır. Yapılan çalışmada Türkiye Bina Deprem Yönetmeliği'ne uygun olarak elde edilen betonarme kare kolon kesitlerinin sargılı beton basınç dayanım değerlerinin belirlenmesinde yapay zekâ ve optimizasyon tabanlı analitik yaklaşımlar birlikte ele alınarak incelenmiştir. Karar ağacı algoritması ile sargılı beton basınç dayanımında etkili olan dört temel girdi parametresi belirlenmiş; bu girdilerle farklı alt küme yapılarına sahip adaptif ağ tabanlı bulanık çıkarım sistemi (ANFIS) modelleri oluşturulmuştur. Ayrıca parçacık sürü optimizasyonu (PSO) kullanılarak dört farklı model oluşturulmuştur. Elde edilen sonuçlara göre, ANFIS modelleri daha yüksek doğruluk sunarken, PSO modelleri parametrik formülasyon ve işlem verimliliği açısından avantaj sağlamış; her iki yöntemin üstünlükleri karşılaştırmalı olarak ortaya konulmuştur. Ayrıca, PSO algoritmasıyla oluşturulan modeller arasında test, eğitim ve doğrulama verileri üzerinde belirgin farklılıklar oluşmadığı gözlemlenmiş, bu modellerin benzer doğruluk düzeylerinde tutarlı sonuçlar verdiği görülmüştür. ANFIS ve PSO'nun güçlü yönlerinin birlikte değerlendirilmesiyle, gelecekte daha genellenebilir ve uyarlanabilir hibrit modellerin geliştirilmesi mümkün görülmektedir.

Anahtar kelimeler: Sargılı beton, basınç dayanımı, yapay zekâ, optimizasyon.

MODELING OF CONFINED CONCRETE COMPRESSIVE STRENGTH WITH ARTIFICIAL INTELLIGENCE AND OPTIMIZATION BASED APPROACHES

ABSTRACT

The cross-sectional properties of reinforced concrete elements are among the basic factors determining the behavior of confined concrete, and it is of great importance that this behavior is accurately reflected in analytical calculations. Determining and analyzing a large number of variables depending on these cross-sectional properties can be time-consuming. Therefore, many artificial intelligence-based modeling methods have started to be used today. In the study, artificial intelligence and optimization-based analytical approaches were examined together in determining the confined concrete compressive strength values of reinforced concrete square column sections obtained in accordance with the Turkish Building Earthquake Code. Four basic input parameters effective in confined concrete compressive strength were determined with the decision tree algorithm; adaptive network-based fuzzy inference system (ANFIS) models with different subset structures were created with these inputs. In addition, four different models were created using particle swarm optimization (PSO). According to the results obtained, ANFIS models provided higher accuracy, while PSO models provided advantages in terms of parametric

*Sorumlu Yazar/Corresponding Author: abdullahgundogay@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 107 – 117, 2025
DOI:10.62301/usmtd.1716436

formulation and processing efficiency; the superiorities of both methods were demonstrated comparatively. In addition, it was observed that there were no significant differences between the models created with the PSO algorithm on the test, training, and check data, and it was seen that these models gave consistent results at similar accuracy levels. By evaluating the strengths of ANFIS and PSO together, it seems possible to develop more generalizable and adaptable hybrid models in the future. According to the obtained results, the superiorities of ANFIS and PSO-based models were comparatively demonstrated.

Keywords: Confined concrete, compressive strength, artificial intelligence, optimization.

1. Giriş (Introduction)

Türkiye Bina Deprem Yönetmeliği'nde (TBDY) [1] betonarme binaların doğrusal olmayan yöntemler kullanılarak şekildeğiştirmeye göre değerlendirilmesinde beton için sargısız ve sargılı beton modelleri verilmiştir. Betonarme elemanların kesit özelliklerine bağlı olarak sargılı beton basınç dayanımlarının ve davranışlarının doğru bir şekilde belirlenebilmesi kesit hesapları ile doğrusal elastik olmayan davranışlarının gerçeğe yakın olarak belirlenebilmesi için büyük öneme sahiptir. Bu durum özellikle deprem esnasında binanın ayakta kalmasını sağlayan kolon elemanlar için daha da önemlidir. Literatürde betonarme elemanların kesit özelliklerinin değişiminin sargılı beton modellerinin davranışına etkisinin incelendiği birçok çalışma bulunmaktadır [2-10].

Beton ve betonarme elemanların özelliklerinin belirlenmesi sınırlı ve uzun süre gerektirebileceğinden dolayı sargılı beton basınç dayanımının tahmininde veri madenciliği ve yapay zekâ tabanlı modelleme yöntemleri günümüzde kullanılmaya başlanmıştır [11-18]. Yapay zekâ tabanlı modelleme yöntemleri içinde özellikle adaptif ağ tabanlı bulanık çıkarım sistemi (ANFIS), öğrenme kabiliyeti ve bulanık mantıkla belirsizliği yönetme yeteneği nedeniyle mühendislik alanlarında yaygın olarak tercih edilmektedir [19-20]. Betonun ve betonarme elemanların basınç dayanımlarının tahmin edilmesinde sıkça kullanılmıştır [21-23]. Ayrıca genetik algoritma, parçacık sürü optimizasyonu (PSO), vb. doğadan esinlenen optimizasyon teknikleri, analitik modellerin doğruluk düzeyini arttırarak parametre kalibrasyonunda etkili çözümler sunduğu için kullanılmaya başlanmıştır [24-25]. Literatürde yapılan çalışmalar genel olarak incelendiğinde ANFIS ve PSO gibi yöntemlerin ayrı ayrı şekilde kullanıldığı görülmüştür.

Bu çalışmada, kesit özellikleri TBDY'ye [1] göre belirlenmiş betonarme kare kolonların sargılı beton basınç dayanımının tahmin edilmesinde yapay zekâ temelli modelleme ve optimizasyon tabanlı analitik yaklaşımların bütünleşik biçimde ele alındığı özgün bir yöntem önerilmektedir. Literatürde sıklıkla karşılaşılan ANFIS modellerinin veri güdümlü öğrenme kabiliyeti, bu çalışmada karar ağacı algoritması ile desteklenerek giriş parametrelerinin seçiminde nesnel ve etkili bir yöntem olarak uygulanmıştır. Belirlenen dört temel parametre (sargısız beton basınç dayanımı, sargı donatısı kol adedi, sargı donatısının kolları arası mesafe ve sargı donatısı oranı) sargı etkisinin temsil gücü yüksek değişkenler olup modelin performansını önemli ölçüde arttırmıştır. Buna ek olarak, PSO algoritması kullanılarak dört farklı yapılandırmada oluşturulan ampirik denklemler hem sabit katsayılı hem de serbest parametreli senaryolarda kök ortalama karesel hata (RMSE) ve mutlak hata (AE) minimizasyonuna odaklı olarak optimize edilmiştir. Geliştirilen bu modellerin doğruluk, hata payı ve istatistiksel anlamlılığı; regresyon (R^2), RMSE ve ortalama mutlak hata (MAE) ile çok yönlü olarak değerlendirilerek ANFIS ile PSO temelli modellerin üstünlükleri karşılaştırmalı olarak ortaya konulmuştur. Elde edilen bulgular, sargılı beton basınç dayanımının tahmininde modern hesaplama yaklaşımlarının mühendislik uygulamalarına entegrasyonuna yönelik önemli katkılar sunarak literatüre metodolojik açıdan yenilikçi bir perspektif kazandırmaktadır.

2. Materyal ve Metot (Material and Method)

2.1. Kolonların kesit özellikleri (Section properties of columns)

Betonarme kolonların kesit özellikleri belirlenirken ülkemizde tasarım aşamasında güncel olarak kullanılmakta olan TBDY [1] ile Betonarme Yapıların Tasarım ve Yapım Kuralları (TS500) [26] dikkate alınmıştır. Bu yönetmeliklerde verilen sınır değerlere uyulmuştur. Kolonların kesit özelliklerinin değerleri Tablo 1'de detaylı olarak sunulmuştur.

*Sorumlu Yazar/Corresponding Author: abdullahgundogay@sdu.edu.tr

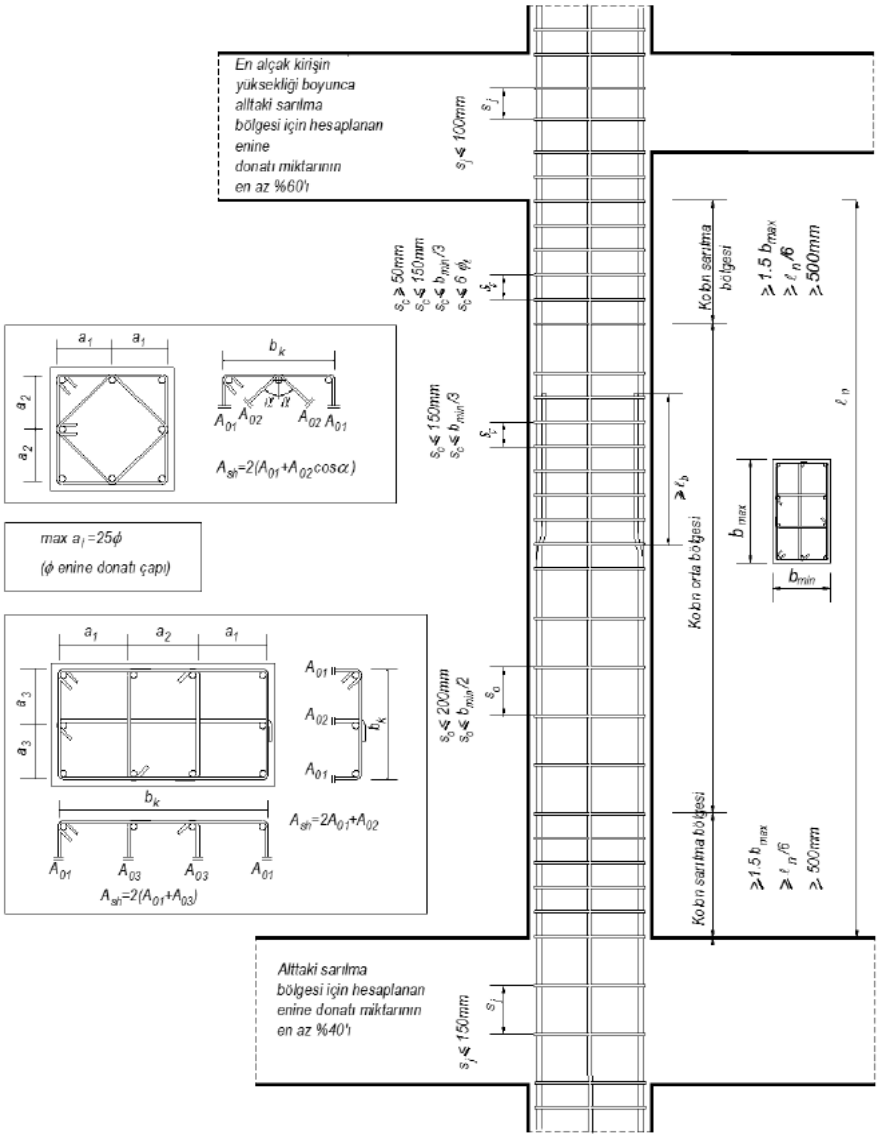
ISSN: 2618-6055 / 9, (1), 107 – 117, 2025
DOI:10.62301/usmtd.1716436

Tablo 1. Kolonların kesit özellikleri (Value ranges of section properties)

Özellik	Değer Aralığı
Beton basınç dayanımı (MPa)	25-30-35-40-45-50
Kesit boyutları (mm)	300-350-400-450-500
Boyuna donatı çapı (mm)	14-16-18-20-22-24-26-28-30
Boyuna donatı adedi	4-8-12-16-20-24-28-32
Sargı donatısı çapı (mm)	8-10-12
Sargı donatısı kol adedi	2-3-4-5-6-7-8-9
Sargı donatısı aralığı (mm)	50-75-100-125

Tablo 1'de verilen değerler her bir kesite uygulanamamıştır. Bunun nedeni TBDY [1] ve TS 500 [26] yönetmeliklerinde verilen diğer tasarım şartlarını sağlayabilmektir. Bu maddeler aşağıda sırasıyla verilmiştir.

- Minimum enine donatı oranı (Denklem 1)
- Minimum enine donatı çapının boyuna donatı çapına göre sınırlandırılması (Denklem 2)
- Minimum ve maksimum boyuna donatı oranları (%1-4)
- Boyuna donatılar arası mesafelerin sınırlandırılması
- Sarılma bölgelerinde enine donatı aralığının sınırlandırılması (Şekil 1)
- Enine donatı kolları arası yatay mesafenin sınırlandırılması (Şekil 1)



Şekil 1. Enine donatı aralığının ve kollarının sınırlandırılması (Limitation of the spacing and branches of transverse reinforcement)

*Sorumlu Yazar/Corresponding Author: abdullahgundogay@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 107 – 117, 2025
DOI:10.62301/usmtd.1716436

$$\rho_{s,min} = 0.3\frac{f_{ctd}}{f_{ywd}}\, b_w \quad (1)$$

$$\emptyset/_3 < \emptyset_s \quad (2)$$

Denklem 1 ve 2'de verilen $\rho_{s,min}$, enine donatı oranını; f_{ctd}, betonun tasarım çekme dayanımını; f_{ywd}, enine donatının tasarım akma dayanımını; b_w, kesit genişliğini; $\emptyset$, boyuna donatı çapını; $\emptyset_s$ ise enine donatı çapını ifade etmektedir.

Yukarıda maddeler halinde verilen tasarım şartlarının da sağlandığı 7556 adet betonarme kolon kesiti oluşturularak sargılı beton basınç dayanımları hesaplanmıştır.

2.2. Sargılı beton basınç dayanımı (Confined concrete compressive strength)

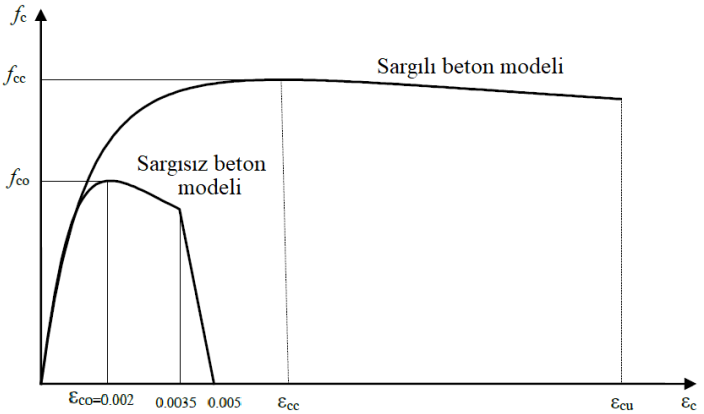
Çalışma kapsamında oluşturulan betonarme kolon kesitlerinin sargılı beton basınç dayanımları TBDY'ye [1] uygun olarak aşağıda verilen denklemler yardımıyla hesaplanmıştır. Sargılı betonun davranışına ait gerilme-birim şekildeğiştirme diyagramı Şekil 2'de sargısız beton ile kıyaslamalı olarak sunulmuştur.

$$k_e=\left(1-\frac{\sum a_i^2}{6b_o h_o}\right)\left(1-\frac{s}{2b_o}\right)\left(1-\frac{s}{2h_o}\right)\left(1-\frac{A_s}{b_o h_o}\right)^{-1} \quad (3)$$

$$f_e=(f_{ex}+f_{ey})/2 \quad , \quad f_{ex}=k_e\rho_x f_{yw} \quad , \quad f_{ey}=k_e\rho_y f_{yw} \quad (4)$$

$$f_{cc}=\left(2.254\sqrt{1+7.94\frac{f_e}{f_{co}}}-2\frac{f_e}{f_{co}}-1.25\right)f_{co} \quad (5)$$

Denklem 3 ile 6 arasında verilen k_e, sargılama etkinlik katsayısını; a_i, boyuna donatılarının eksenleri arasındaki mesafeyi; s, enine donatı aralığını; As, toplam boyuna donatı alanını; b_o ile h_o, enine donatının eksenleri arasındaki kesit boyutlarını; f_e, etkili sargılama basıncını; f_{ex} ile f_{ey}, ilgili yönlerdeki sargılama basıncını; ρ_x ile ρ_y, ilgili yönlerdeki enine donatıların hacimsel oranını; f_{yw}, enine donatı akma dayanımını; f_{co}, sargısız beton dayanımını; f_{cc}, ise sargılı beton basınç dayanımını ifade etmektedir.



Şekil 2. TBDY'ye göre beton modellerine ait gerilme-birim şekil değiştirme diyagramları (Stress-strain diagrams of concrete models according to TBDY) [1]

2.3.Karar ağacı (Decision tree)

Karar ağacı, genel olarak mevcut verileri özelliklerine göre çeşitli dallara ayırarak bir karar vermeye veya tahmin yapmaya çalışan bir yapıdır. Yani her karar ağacı öncelikle kökten başlayarak dallara ayrılır ve her iç düğümünde veri, belirlenen bir özelliğe göre bölünür. Yaprak düğümlerinde ise karar verilmiş

*Sorumlu Yazar/Corresponding Author: abdullahgundogay@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 107 – 117, 2025
DOI:10.62301/usmtd.1716436

olur. Böylelikle karmaşık veriler bile anlamlı gruplara ayrılarak her veri noktası için uygun karar kolayca bulunabilir.

2.4. Adaptif ağ tabanlı bulanık çıkarım sistemi (Adaptive neuro-fuzzy inference system)

ANFIS, özellikle doğrusal olmayan sistemlerin modellemesinde tercih edilmektedir. Bulanık mantığın yorumlama yeteneği ile yapay sinir ağlarının öğrenme yeteneğini birleştiren melez bir yapay zekâ modelidir [27]. Sugeno tipi bulanık çıkarım sistemini kullandığı için sistemin parametrelerini veriye dayalı olarak optimize etmektedir. Model mimarisi, yapay bir sinir ağı biçiminde olup genellikle girdi, üyelik fonksiyonu, kural, normalizasyon ve çıkış katmanlarından oluşan bir yapıya sahiptir [28]. Girdi katmanında, model tarafından alınan girdiler tanımlanır. Üyelik fonksiyonu katmanında her bir girdi değişkeni bulanık kümeler haline getirilir. Kural Katmanında, bulanık kurallar uygulanarak her bir düğüm bir kural tarafından gösterilir. Normalizasyon katmanında, her bir kuralın ateşleme kuvveti hesaplanarak normalize edilir. Çıkış katmanında ise normalize edilmiş ateşleme kuvvetlerine göre sonuç değerleri hesaplanır.

ANFIS'te oluşturulan modelin güvenilir sonuçlar verebilmesi için eğitim, test ve doğrulama veri kümelerinin doğru bir şekilde belirlenmesi gerekmektedir. Eğitim verileri, oluşturulan modelin öğrenme aşamasında kullandığı ana veri kümesi olduğu için verilerin %60'ını; doğrulama verileri, oluşturulan modelin aşırı öğrenme yapıp yapmadığını kontrol etmek amacıyla kullanıldığı için verilerin %20'sini; test verileri ise modelin genel performansını ölçmek için kullanıldığı için verilerin %20'sini kapsamalıdır. Bunun için çalışmada eğitim amacıyla 4533 adet veri, doğrulama amacıyla 1512 adet veri, test için ise 1511 adet veri kullanılmıştır. Eğitim, test ve doğrulama veri kümelerine ait tanımlayıcı istatiksel bilgiler girdi parametreleri (f_{co}, kesit boyutları (b, h), Ø, boyuna donatı adedi (n_b), sargı donatısı çapı ($Ø_s$), sargı donatısı adedi (n_s), s, A_s, a_i, enine donatı oranı (ρ_s) ve boyuna donatı oranı (ρ)) ile çıktı için Tablo 2 ile Tablo 4 arasında sunulmuştur.

Tablo 2. Eğitim veri kümesi için istatiksel bilgiler (Statistical information for the training dataset)

	f_{co} (MPa)	b, h (mm)	Ø (mm)	n_b	$Ø_s$ (mm)	n_s	s (mm)	A_s (mm^2)	a_i (mm)	ρ_s (%)	ρ (%)	f_{cc} (MPa)
Ortalama	35.96	486.89	19.73	18.84	10.48	5.17	64.17	5734.43	110.83	1.48	2.37	62.72
Standart Hata	0.13	1.35	0.06	0.10	0.02	0.03	0.26	40.52	0.71	0.01	0.01	0.22
Standart Sapma	8.46	91.10	4.13	7.00	1.55	1.80	17.83	2727.78	47.71	0.57	0.80	14.63
Basıklık	-1.21	-0.90	-0.40	-0.76	-1.22	-0.79	0.45	-0.13	1.05	1.54	-1.05	-0.30
Çarpıklık	0.23	-0.43	0.50	0.06	-0.45	0.33	1.05	0.68	1.26	1.25	0.16	0.45
En Küçük	25	300	14	4	8	2	50	1017.90	54.00	0.60	1.01	36.05
En Büyük	50	600	30	32	12	9	125	14137.20	262.00	3.70	3.98	114.28
Güvenirlik Düzeyi	0.25	2.65	0.12	0.20	0.05	0.05	0.52	79.43	1.39	0.02	0.02	0.43

Tablo 3. Test veri kümesi için istatiksel bilgiler (Statistical information for the test dataset)

	f_{co} (MPa)	b, h (mm)	Ø (mm)	n_b	$Ø_s$ (mm)	n_s	s (mm)	A_s (mm^2)	a_i (mm)	ρ_s (%)	ρ (%)	f_{cc} (MPa)
Ortalama	35.95	486.14	19.70	18.86	10.48	5.16	63.15	5719.46	110.49	1.51	2.37	63.12
Standart Hata	0.22	2.35	0.11	0.18	0.04	0.05	0.46	70.10	1.20	0.02	0.02	0.38
Standart Sapma	8.46	91.50	4.12	7.01	1.55	1.78	17.70	2725.05	46.69	0.59	0.80	14.73
Basıklık	-1.20	-0.90	-0.39	-0.76	-1.21	-0.76	0.72	-0.01	0.85	1.66	-1.02	-0.06
Çarpıklık	0.24	-0.43	0.51	0.06	-0.45	0.36	1.17	0.70	1.19	1.29	0.17	0.49
En Küçük	25	300	14	4	8	2	50	1017.90	54.00	0.61	1.01	36.81
En Büyük	50	600	30	32	12	9	125	14137.20	262.00	3.70	3.98	113.69
Güvenirlik Düzeyi	0.43	4.62	0.21	0.35	0.08	0.09	0.89	137.51	2.36	0.03	0.04	0.74

*Sorumlu Yazar/Corresponding Author: abdullahgundogay@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 107 – 117, 2025
DOI:10.62301/usmtd.1716436

Tablo 4. Doğrulama veri kümesi için istatiksel bilgiler (Statistical information for the check dataset)

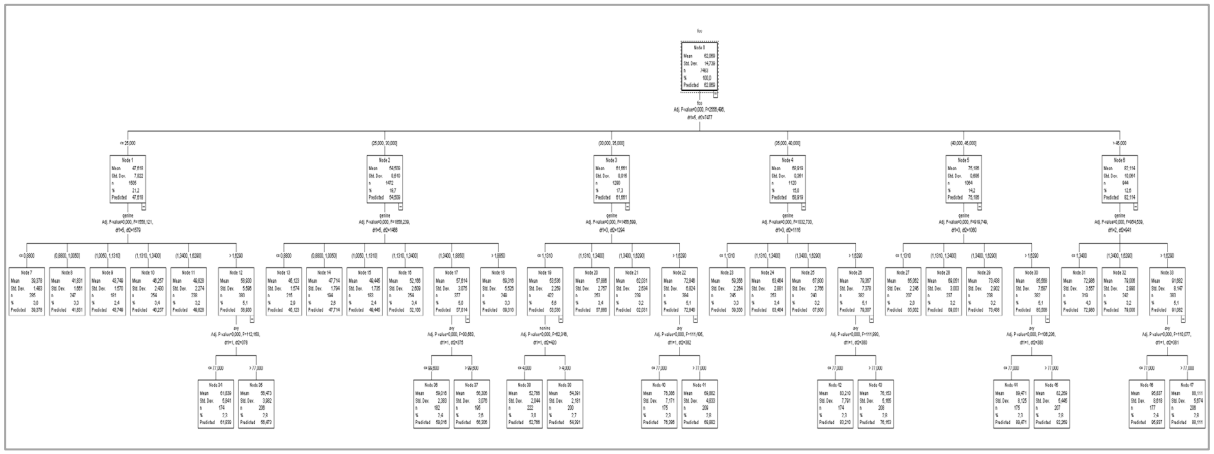
	f_{co} (MPa)	b, h (mm)	Ø (mm)	n_b	$Ø_s$ (mm)	n_s	s (mm)	A_s (mm^2)	a_i (mm)	ρ_s (%)	ρ (%)	f_{cc} (MPa)
Ortalama	35.96	487.07	19.74	18.85	10.48	5.23	63.86	5740.18	109.40	1.50	2.37	63.18
Standart Hata	0.22	2.34	0.10	0.18	0.04	0.05	0.46	70.12	1.20	0.02	0.02	0.39
Standart Sapma	8.47	90.88	4.05	7.04	1.55	1.83	17.98	2726.43	46.80	0.60	0.80	15.12
Basıklık	-1.20	-0.89	-0.34	-0.77	-1.22	-0.81	0.58	-0.22	1.11	1.36	-1.05	-0.15
Çarpıklık	0.24	-0.44	0.52	0.07	-0.45	0.33	1.11	0.65	1.27	1.21	0.14	0.50
En Küçük	25	300	14	4	8	2	50	1017.90	54.00	0.60	1.01	35.92
En Büyük	50	600	30	32	12	9	125	14137.20	262.00	3.70	3.98	114.58
Güvenirlik Düzeyi	0.43	4.58	0.20	0.36	0.08	0.09	0.91	137.54	2.36	0.03	0.04	0.76

2.5. Parçacık sürü optimizasyonu (Particle swarm optimization)

PSO, popülasyon tabanlı optimizasyon algoritmasıdır ve doğada yaşayan balık toplulukları, kuş sürüleri gibi organizmaların grup hareketlerinden esinlenerek geliştirilmiştir. Algoritmada, arama uzayındaki her bir çözüm adayına parçacık adı verilmektedir. Her parçacık hız ve konum vektörlerine sahip olmaktadır [29]. Sürüdeki tüm parçacıklar, zamanla hem kendi hem de sürüdeki en iyi konumlarına göre kendilerini güncelleyerek çözüm uzayındaki en iyi noktaya doğru hareket ederler. Algoritmanın durdurma kriterine kadar yinelemeler yapılarak sürü içindeki en iyi parçacık bulunmaya çalışılır.

3. Araştırma Bulguları (Research Findings)

Yapılan çalışma kapsamında öncelikle her bir betonarme kolon kesitinin sargılı beton basınç dayanımı 11 adet girdiye bağlı olarak elde edilmiştir. Girdi sayısı oldukça fazla olduğundan dolayı ANFIS modellerinde zamanlama açısından çok uzun süreler alacağı için öncelikle karar ağacı kullanılarak en fazla etkili olan girdi parametreleri belirlenmiştir. 7556 adet betonarme kolon kesitine ait verilerin karar ağacından değerlendirilmesi sonucunda f_{co} ve ρ_s en önemli iki parametre olarak belirlenmiştir. İkinci etapta ise sırasıyla a_s ve n_s bunlara en yakın girdi parametreleri olarak bulunmuştur (Şekil 3). Bu parametrelere göre ANFIS modelleri üç ve dört girdili olarak kurulmuştur. Girdilerin alt küme sayıları belirlenirken karar ağacındaki dağılımları dikkate alınmıştır. Üç girdili model için 40 adet, dört girdili model için ise 100 adet olmak üzere toplam 140 adet ANFIS modelinin analizi gerçekleştirilmiştir. Analizler sonucunda veri kümeleri için elde edilen R^2, RMSE ve MAE değerleri Tablo 5'te kıyaslamalı olarak sunulmuştur. ANFIS model sayısının fazlalığından dolayı üç ve dört girdili modellerin her biri için girdi parametrelerine bağlı olarak veri kümeleri için en iyi sonuçları veren modellere ait değerler Tablo 5'te verilmiştir.



Şekil 3. Karar ağacı modeli (Decision tree model)

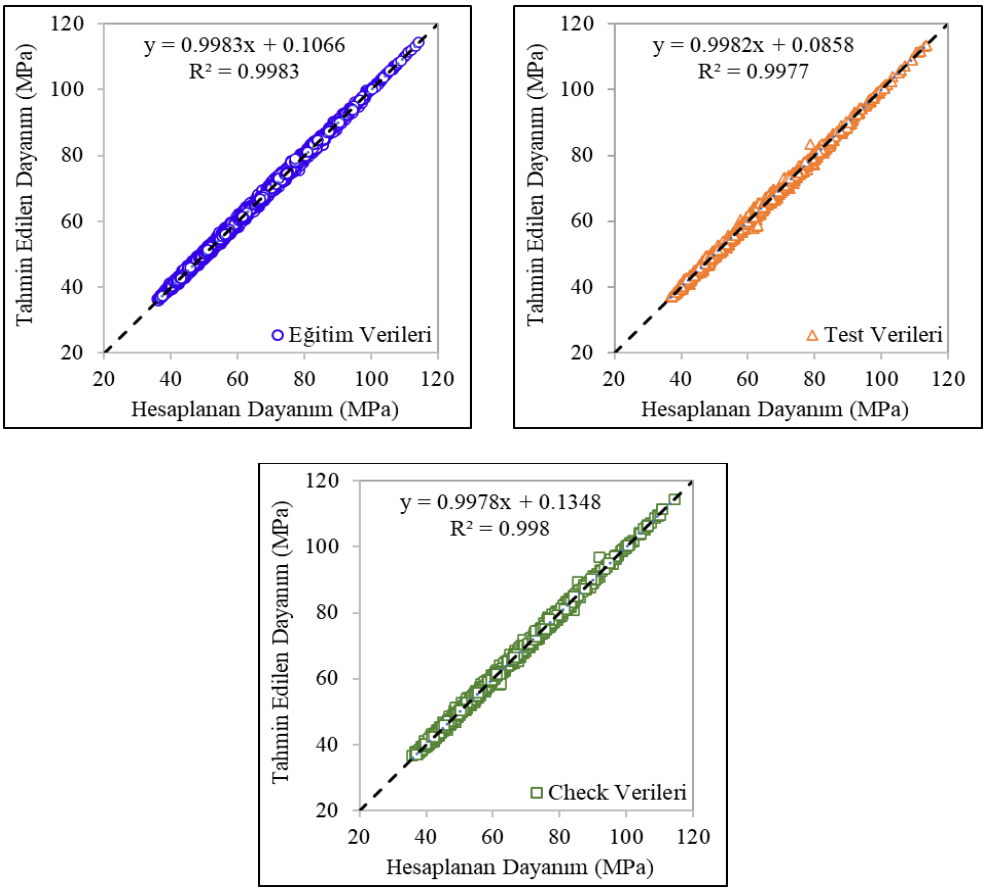
*Sorumlu Yazar/Corresponding Author: abdullahgundogay@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 107 – 117, 2025
DOI:10.62301/usmtd.1716436

Tablo 5. ANFIS modellerine ait analiz sonuçları (Analysis results of ANFIS models)

Modeller	Girdi Parametreleri	Eğitim			Test			Doğrulama		
		R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE
6-6-5	f_{co}- a_s- ρ_s	0,994	1,158	0,015	0,994	1,182	0,015	0,994	1,190	0,015
6-6-6		0,993	1,183	0,015	0,993	1,211	0,016	0,994	1,205	0,016
6-4-5	f_{co}- n_s- ρ_s	0,997	0,749	0,010	0,997	0,770	0,010	0,997	0,780	0,010
6-5-4		0,997	0,746	0,010	0,997	0,765	0,010	0,997	0,780	0,010
6-5-5		0,997	0,742	0,010	0,997	0,764	0,010	0,997	0,781	0,010
6-6-5		0,997	0,742	0,010	0,997	0,769	0,010	0,997	0,787	0,010
6-6-6		0,997	0,743	0,010	0,997	0,766	0,010	0,997	0,782	0,010
6-2-4-6	f_{co}- n_s - a_s- ρ_s	0,998	0,711	0,009	0,998	0,730	0,009	0,998	0,726	0,009
6-2-6-5		0,998	0,690	0,009	0,998	0,724	0,009	0,998	0,717	0,009
6-2-6-6		0,998	0,698	0,009	0,998	0,730	0,009	0,998	0,715	0,009
6-3-6-5		0,998	0,666	0,008	0,998	0,712	0,009	0,998	0,695	0,009
6-3-6-6		0,998	0,667	0,008	0,998	0,715	0,009	0,998	0,696	0,009
6-4-4-6		0,998	0,661	0,008	0,998	0,708	0,009	0,998	0,697	0,009
6-4-6-6		0,998	0,641	0,008	0,998	0,717	0,009	0,998	0,693	0,008
6-5-5-6		0,998	0,599	0,007	0,998	0,701	0,008	0,998	0,670	0,008
6-6-3-3		0,998	0,699	0,009	0,998	0,734	0,009	0,998	0,736	0,009
6-6-5-6		0,998	0,646	0,008	0,997	0,812	0,009	0,997	0,774	0,009
6-6-6-6		0,998	0,635	0,008	0,997	0,825	0,009	0,997	0,771	0,009

Tablo 5'te, modeller sütunundaki sayıların her biri sırasıyla girdi parametrelerinin alt küme sayısını göstermektedir. Analiz sonuçları incelendiğinde, 6-5-5-6 modeli eğitim, test ve kontrol veri kümelerinde en yüksek R^2, en düşük RMSE ve MAE değerleri elde edilmiştir. Bu modele ait saçılım diyagramları da Şekil 4'te sunulmuştur.



Şekil 4. 6-5-5-6 modeli için saçılım diyagramları (Scatter diagrams for the 6-5-5-6 model)

*Sorumlu Yazar/Corresponding Author: abdullahgundogay@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 107 – 117, 2025
DOI:10.62301/usmtd.1716436

PSO algoritması kullanılarak hem sabit katsayılı hem de serbest parametreli senaryolar dikkate alınarak RMSE ve AE minimizasyonuna odaklı dört farklı model oluşturulmuştur. Sabit katsayılı RMSE (Model 1), RMSE (Model 2), sabit katsayılı AE (Model 3) ve AE (Model 4) modellerine ait denklemler sırasıyla aşağıda sunulmuştur. Ayrıca bu modellere ait R^2, RMSE ve MAE için elde edilen değerler kıyaslamalı olarak Tablo 6'da sunulmuştur.

$$f_{cc} = 1{,}1804f_{co} + 0{,}5491n_s - 0{,}0130a_s + 14{,}2597\rho_s - 2{,}2418 \quad (1)$$

$$f_{cc} = 1{,}1659f_{co} + 0{,}3772n_s - 0{,}0194a_s + 14{,}1747\rho_s \quad (2)$$

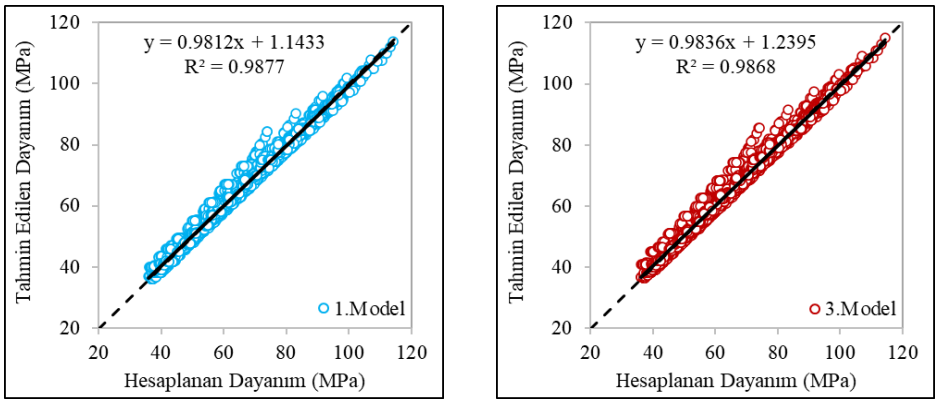
$$f_{cc} = 1{,}1694f_{co} + 0{,}6450n_s - 0{,}0015a_s + 14{,}8703\rho_s - 4{,}2638 \quad (3)$$

$$f_{cc} = 1{,}1461f_{co} + 0{,}2946n_s - 0{,}0138a_s + 14{,}6640\rho_s \quad (4)$$

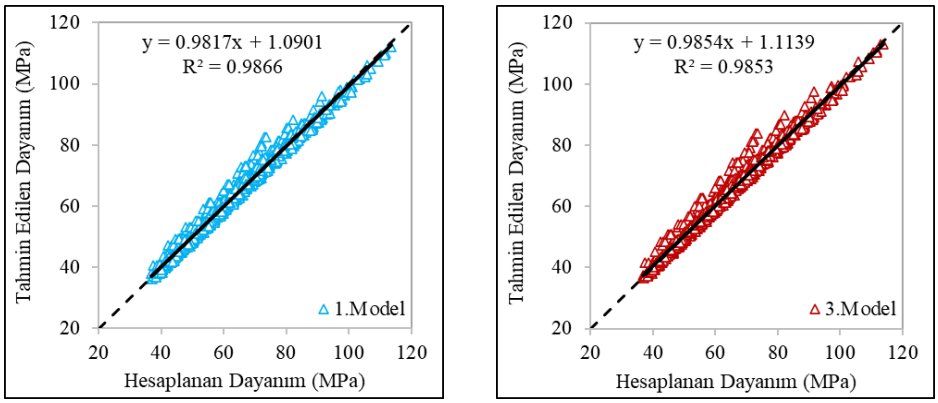
Tablo 6. PSO modellerine ait analiz sonuçları (Analysis results of 4 inputs ANFIS models)

Modeller	Eğitim			Test			Doğrulama		
	R^2	RMSE	MAE	R^2	RMSE	MAE	R^2	RMSE	MAE
Model 1	0,988	1,626	0,020	0,987	1,708	0,020	0,988	1,659	0,020
Model 2	0,987	1,654	0,021	0,986	1,742	0,021	0,988	1,692	0,021
Model 3	0,987	1,696	0,019	0,985	1,795	0,019	0,987	1,728	0,019
Model 4	0,987	1,720	0,020	0,985	1,827	0,020	0,987	1,761	0,020

Tablo 6 incelendiğince eğitim, test ve doğrulama veri kümeleri için R^2 ve RMSE için en iyi sonuçların Model-1'den, MAE için ise Model-3'ten elde edildiği görülmektedir. Bu modellerin veri kümeleri için saçılım diyagramları Şekil 5 ile Şekil 7 arasında karşılaştırmalı olarak verilmiştir.



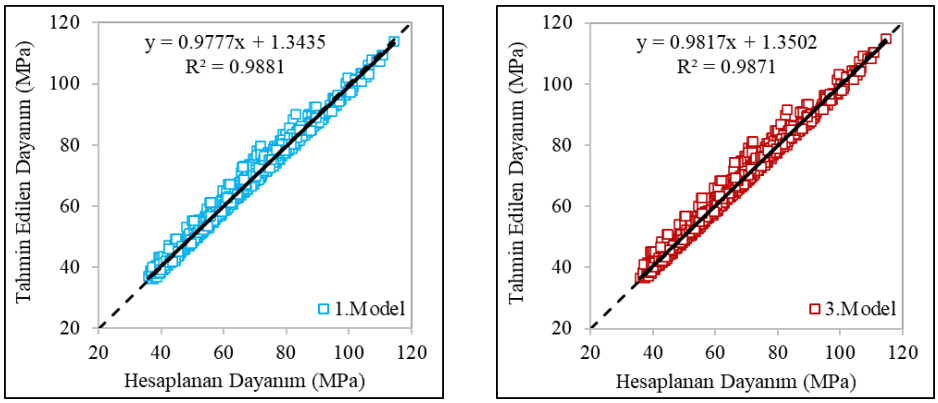
Şekil 5. Model 1 ve Model 3'teki eğitim verilerinin saçılım diyagramları (Scatter plots for the training data in Models 1 and 3)



Şekil 6. Model 1 ve Model 3'teki test verilerinin saçılım diyagramları (Scatter plots for the test data in Models 1 and 3)

*Sorumlu Yazar/Corresponding Author: abdullahgundogay@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 107 – 117, 2025
DOI:10.62301/usmtd.1716436



Şekil 7. Model 1 ve Model 3'teki doğrulama verilerinin saçılım diyagramları (Scatter plots for the check data in Models 1 and 3)

4. Tartışma ve Sonuç (Results and Discussion)

Bu çalışmada, TBDY'ye [1] uygun olarak modellenen 7556 adet betonarme kare kolon kesiti kullanılarak sargılı beton basınç dayanımının tahmini için yapay zekâ ve optimizasyon tabanlı iki farklı yöntem olan ANFIS ve PSO sistematik biçimde karşılaştırılmıştır. Bunun için öncelikle betonarme kolonlara ait 11 farklı girdi parametresi tanımlanmış olup, bu parametrelerin model üzerindeki etkilerini belirlemek amacıyla karar ağacı algoritması uygulanmıştır. Karar ağacı analizi sonucunda, sargılı beton basınç dayanımını en fazla etkileyen dört temel girdi parametresi olarak f_{co}, ρ_s, a_s ve n_s belirlenmiştir. Bu dört temel girdi parametresi kullanılarak hem ANFIS hem de PSO tabanlı modeller oluşturulmuş ve sargılı beton basınç dayanımının tahmini açısından performansları karşılaştırılmıştır.

Modelleme sonuçlarına göre 6-5-5-6 yapısına sahip dört girdili ANFIS modeli eğitim, test ve doğrulama veri kümelerinde oldukça başarılı sonuçlar vermiştir. Bu bulgular, karar ağacı ile yapılan değişken seçiminin model performansını arttırmada etkili olduğunu ve ANFIS'in az sayıda ama anlamlı girdilerle yüksek başarı sağlayabildiğini göstermiştir.

PSO algoritması ile oluşturulan RMSE ve MAE değerlerini minimum yapmak üzerine kurulmuş dört farklı model sonuçlarına göre eğitim, test ve doğrulama veri kümeleri üzerinde çok belirgin farklar oluşmadığı görülmüştür. Bu durum PSO'nun sargılı beton basınç dayanımını temsil eden karmaşık ve çok değişkenli ilişkileri etkin bir biçimde modelleyebildiğini ve özellikle analitik denklem tabanlı tahminlerde etkili bir araç olduğunu göstermektedir.

ANFIS modelleri, özellikle yüksek doğruluk düzeyleri ve düşük hata oranları ile öne çıkarken; PSO modelleri ise parametrik yalınlık, analitik formülasyon üretimi ve işlem verimliliği açısından avantaj sağlamıştır. Her iki yöntemin elde ettiği sonuçlar, yapay zekâ ve optimizasyon tabanlı yaklaşımların bu alanda etkili tahmin araçları sunduğunu ortaya koymakta; ANFIS'in öğrenme yeteneği ile PSO'nun optimizasyon gücünün birlikte değerlendirilmesinin ileride daha güçlü modeller geliştirilmesine olanak sağlayacağını göstermektedir.

Gelecekte yapılabilecek çalışmalar için çeşitli beton modelleri, yapay zekâ ve optimizasyon modelleri kullanılarak araştırmanın genişletilmesi, ulusal ve uluslararası literatürün gelişmesine katkıda bulunacaktır. Literatürde yaygın tercih edilen sargılı beton modellerinin basınç dayanımları arasındaki uyum incelenerek modelleme çalışmaları yapılabilir. Özellikle yüksek basınç dayanımına sahip betonarme kesitler için deneysel çalışmalar yapılarak elde edilen sonuçların analitik hesaplar ile uyumu araştırılabilir.

5. Kaynaklar (References)

[1] Türkiye Bina Deprem Yönetmeliği, Afet ve Acil Durum Yönetimi Başkanlığı, Ankara, Türkiye, 2018.

[2] A. İlki, N. Kumbasar, Sargılı beton için mevcut modellerin deneysel verilerle karşılaştırılması, Teknik Dergi 12 (3) (2001) 2419–2433.

*Sorumlu Yazar/Corresponding Author: abdullahgundogay@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 107 – 117, 2025
DOI:10.62301/usmtd.1716436

[3] M.Y. Kaltakcı, A. Köken, Ü.S. Yımaz, Eksenel yük altındaki çelik lifli ve lifsiz etriyeli betonarme kolonların davranışının deneysel ve analitik olarak incelenmesi, Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi 8 (1) (2006) 65–85.

[4] H.B. Özmen, M. İnel, H. Bilgin, Sargılı beton davranışının betonarme eleman ve sistem davranışına etkisi, Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi 22 (2) (2007) 375–383.

[5] E. Işık, M. Özdemir, İ.B. Karaşin, A. Karaşin, Betonarme yapılarda kullanılan malzeme modellerinin karşılaştırılması, Bitlis Eren Üniversitesi Fen Bilimleri Dergisi 8 (3) (2019) 968–984.

[6] S.B. Yüksel, S. Foroughi, Betonarme Kolonların Sargısız ve Sargılı Beton Dayanımının Analitik Olarak Araştırılması, Konya Journal of Engineering Sciences 7 (3) (2019) 612–631.

[7] S. Foroughi, R. Jamal, B. Yüksel, TBDY 2018 ve Mander modeline göre sargılı betonun gerilme-şekil değiştirmesinin araştırılması, El-Cezeri Sci. Eng. J. 8 (1) (2021) 363–375.

[8] S. Foroughi, B. Yüksel, Investigation of nonlinear behavior of the reinforced concrete columns for different confined concrete models, Politeknik Dergisi 25 (4) (2021) 1447–1462.

[9] E. Kılıç, M.F. Güllü, Görece narin betonarme perde duvarlarda mevcut beton malzeme modellerinin etkilerinin irdelenmesi Bilecik Şeyh Edebali Üniversitesi Fen Bilimleri Dergisi 9 (2) (2022) 799–818.

[10] A. Gündogay, Comparison of confined concrete models, 1st International Conference on Innovative Academic Studies, Konya, Turkey, September 10-13, 2022, pp. 1376–1382.

[11] U. Sahin, İ. Bedirhanoglu, A fuzzy model approach to stress–strain relationship of concrete in compression, Arabian Journal for Science and Engineering 39 (2014) 4515–4527.

[12] C. Özel, A. Topsakal, Veri madenciliği kullanarak beton basınç dayanımının belirlenmesi, Cumhuriyet Üniversitesi Fen Edebiyat Fakültesi Fen Bilimleri Dergisi 35 (1) (2014) 1–11.

[13] S. Murtazaoğlu, K. Yetilmezsoy, B. Doran, CFRP ile güçlendirilmiş betonarme kolonların basınç dayanımının çoklu regresyon modelleriyle tahmini, Erciyes Üniversitesi Fen Bilimleri Enstitüsü Fen Bilimleri Dergisi 31 (3) (2015) 172–178.

[14] H.T. Öztürk, Beton basınç dayanımının JAYA ve öğretme-öğrenme tabanlı optimizasyon (TLBO) algoritmalarıyla modellenmesi, Journal of Investigations on Engineering and Technology 1 (2) (2018) 24–29.

[15] S. Yörübulut, O. Dogan, F. Erdugan, S. Yörübulut, Tahribatsız yöntem verileri kullanılarak yapay sinir ağı ve regresyon yöntemi ile beton basınç dayanımının tahmin edilmesi, International Journal of Engineering Research and Development 12 (2) (2020) 769–776.

[16] A. Gündoğay, Lifli polimer ile sargılanan betonarme kolonların gerilme-şekil değiştirme ilişkisinin anfis yöntemi ile elde edilmesi, Journal of Innovations in Civil Engineering and Technology, 6 (2) (2024) 111–130.

[17] M.A. Bülbül, Beton basınç dayanımı tahmini için özellik mühendisliği ve makine öğrenimi tabanlı hibrit bir yaklaşım, Erciyes Üniversitesi Fen Bilimleri Enstitüsü Fen Bilimleri Dergisi 40 (1) (2024) 10–19.

[18] M.G. Altun, A.H. Altun, Yüksek performanslı betonun basınç dayanımının farklı makine öğrenimi algoritmaları ile tahmin edilmesi, Journal of Innovative Engineering and Natural Science, 5 (1) (2025) 347–361.

[19] M. Saltan, S. Saltan, A. Şahiner, Fuzzy logic modeling of deflection behavior against dynamic loading in flexible pavements, Construction and Building Materials 21 (7) (2007) 1406–1414.

[20] B. Gültekin, G. Dogan, İnşaat mühendisliğinde yapay zekâ çalışmaları, İleri Mühendislik Çalışmaları ve Teknolojileri Dergisi 2 (2) (2021) 117–138.

[21] S. Subaşı, İ. Şahin, B. Çomak, Tahribatsız test sonuçları kullanılarak uçucu kül ikameli betonlarda basınç dayanımının anfıs ile tahmini, Uluslararası Teknolojik Bilimler Dergisi 2 (3) (2010) 9–16.

*Sorumlu Yazar/Corresponding Author: abdullahgundogay@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 107 – 117, 2025
DOI:10.62301/usmtd.1716436

[22] C. Özel, O. Soykan, Betonun basınç mukavemetinin taze beton özelliklerinden tahmini için anfıs modeli, Uluslararası Teknolojik Bilimler Dergisi 4 (1) (2012) 30–45.

[23] F. Kaya, N. Keskin, M.A. Çakıroğlu, Değişik geometrili çelik lif ilaveli betonarme kirişlerin tahribatsız deney yöntemleri ile elde edilen basınç dayanımlarının anfıs metoduyla tahmin edilmesi, Uluslararası Teknolojik Bilimler Dergisi 6 (1) (2014) 44–56.

[24] R. Temür, G. Bekdaş, Betonarme konsol istinat duvarlarının optimum tasarımı, Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi 24 (6) (2018) 1043–1050.

[25] Güler, M. Ulaş, M. Açıkgenç Ulaş, Efektif kendiliğinden yerleşen hafif beton dayanımı tahmini için farklı makine öğrenmesi algoritmalarının karşılıklı değerlendirilmesi, Fırat Üniversitesi Mühendislik Bilimleri Dergisi 37 (1) (2025) 251–261.

[26] Betonarme Yapıların Tasarım ve Yapım Kuralları, Türk Standardları Enstitüsü, Ankara, Türkiye, 2000.

[27] K. Saplıoğlu, R. Acar, K-means kümeleme algoritması kullanılarak oluşturulan yapay zeka modelleri ile sediment taşınımının tespiti, Bitlis Eren Üniversitesi Fen Bilimleri Dergisi, 9 (1) (2020) 306–322.

[28] T.S. Küçükerdem, M. Kilit, K. Saplioglu, Bulanık çıkarım sistemlerinde kullanılan küme sayılarının K-ortalamalar ile belirlenmesi ve baraj hacmi modellenmesi: Kestel barajı örneği, Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi, 25 (8) (2019) 962–967.

[29] K. Saplioglu, T.S.K. Ozturk, R. Acar, Optimization of open channels using particle swarm optimization algorithm, Journal of Intelligent & Fuzzy Systems, 39 (1) (2020) 399–405.

***Sorumlu Yazar/Corresponding Author:** abdullahgundogay@sdu.edu.tr

ISSN: 2618-6055 / 9, (1), 118 – 132, 2025
DOI: 10.62301/usmtd.1712846

Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi
International Journal of Sustainable Engineering and Technology
ISSN: 2618-6055 / Cilt:9, Sayı:1, (2025), 118 – 132
Araştırma Makalesi – Research Article

CARBON FOOTPRINT THROUGH THE LENS OF INFORMATION AND COMMUNICATION TECHNOLOGY

*Fatih GENÇTÜRK[1], Serdar BİROĞUL[2,3]

[1]Isparta University of Applied Sciences, Faculty of Technology, Department of Computer Engineering, Isparta
[2]Düzce University, Faculty of Engineering, Department of Computer Engineering, Düzce
[3]Nakhchivan State University, Faculty of Architecture and Engineering, Department of Electronics and Information Technologies, Nakhchivan

(Geliş/Received: 03.06.2025, Kabul/Accepted: 26.06.2025, Yayınlanma/Published: 30.06.2025)

ABSTRACT

The carbon footprint is one of the key indicators used to evaluate the environmental impacts of human activities and to guide sustainability policies. It refers to the total amount of greenhouse gases emitted into the atmosphere by individuals, institutions, or products. The carbon footprint encompasses not only direct emissions such as those resulting from energy production but also indirect sources, including manufacturing processes, logistics operations, and waste management. Today, these measurements are carried out based on various standards and methodologies. In particular, advancements in sensor technologies, the widespread adoption of cloud-based systems, and increased computational capacity have made it possible to monitor carbon footprints dynamically and in real time. These technological developments enable more precise tracking of environmental impacts over time and contribute to raising public awareness by providing personalized recommendations based on individual consumption habits. This study provides a comprehensive examination of the definition and significance of the carbon footprint, relevant standards, calculation methods, pricing mechanisms, and innovative technological approaches.

Keywords: Carbon footprint, Information and communication technology, Carbon footprint calculation, Emission reduction

BİLGİ VE İLETİŞİM TEKNOLOJİLERİ PERSPEKTİFİNDEN KARBON AYAK İZİ

ÖZ

Karbon ayak izi, insan faaliyetlerinin çevresel etkilerini değerlendirmek ve sürdürülebilirlik politikalarını yönlendirmek amacıyla kullanılan temel göstergelerden biridir. Bu gösterge, bireyler, kurumlar veya ürünler tarafından atmosfere salınan toplam sera gazı miktarını ifade eder. Karbon ayak izi yalnızca enerji üretimi gibi doğrudan emisyonları değil, aynı zamanda üretim süreçleri, lojistik faaliyetler ve atık yönetimi gibi dolaylı emisyon kaynaklarını da kapsamaktadır. Günümüzde bu ölçümler çeşitli standartlar ve metodolojiler temelinde gerçekleştirilmektedir. Özellikle sensör teknolojilerindeki ilerlemeler, bulut tabanlı sistemlerin yaygınlaşması ve artan hesaplama kapasitesi sayesinde karbon ayak izinin dinamik ve gerçek zamanlı olarak izlenmesi mümkün hale gelmiştir. Bu teknolojik gelişmeler, çevresel etkilerin zaman içinde daha hassas bir şekilde takip edilmesini sağlamakta; bireysel tüketim alışkanlıklarına yönelik öneriler sunarak toplumsal farkındalığın artmasına katkıda bulunmaktadır. Bu çalışmada karbon ayak izinin tanımı, önemi, ilgili standartlar, hesaplama yöntemleri, fiyatlandırma mekanizmaları ve yenilikçi teknolojik yaklaşımlar kapsamlı bir biçimde ele alınmaktadır.
Anahtar kelimeler: Karbon ayak izi, Bilgi ve iletişim teknolojileri, Karbon ayak izi hesaplama, Emisyon azaltma

*Sorumlu Yazar/Corresponding Author: fatihgencturk@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 118 – 132, 2025
DOI: 10.62301/usmtd.1712846

1. Introduction

The atmosphere absorbs a portion of the solar radiation emitted by the Earth's surface and cools by reflecting some of it back into space. However, a significant part of this radiation is absorbed by greenhouse gases in the atmosphere, causing both the atmosphere and the Earth's surface to warm. If the atmosphere lacked this property, the Earth's average temperature of 15°C would drop to approximately -18°C [1]. Increased atmospheric concentrations of carbon dioxide, chlorofluorocarbons, methane, and nitrogen oxides further amplify the greenhouse effect, which is considered the primary driver of global climate change. According to the Intergovernmental Panel on Climate Change (IPCC), the average global temperature on land and in the oceans increased by 0.85°C between 1880 and 2012 [2]. As a consequence of this temperature rise, glaciers are melting rapidly, sea levels are rising, and shifts in evaporation and precipitation patterns are leading to more frequent droughts and floods. Among these cascading effects, the depletion of water resources stands out as one of the most critical threats to sustainable living and environmental balance [3,4].

The primary gases contributing to the greenhouse effect are carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O), chlorofluorocarbons (CFCs), and fluorinated gases. These gases originate from both natural processes and human activities, and their atmospheric concentrations have increased significantly since 1750 [2]. As of 2011, the concentrations of CO_2, CH_4, and N_2O had risen by 40%, 150%, and 20%, respectively, compared to pre-industrial levels, reaching 391 ppm, 1,803 ppb, and 324 ppb [2]. In response to these rising levels, numerous international discussions and initiatives have emerged to address greenhouse gas emissions. Efforts began with the First World Climate Conference in 1979 and continued with the establishment of the IPCC in 1988. In 1992, the United Nations Conference on Environment and Development was held, during which the United Nations Framework Convention on Climate Change (UNFCCC) was signed and adopted. The Kyoto Protocol, signed by 160 countries in 1997, emphasized the measurement and reporting of emissions, along with the implementation of mitigation strategies. It came into force in 2005. Later, in 2015, the 21st Conference of the Parties (COP21) convened in Paris, resulting in the adoption of the Paris Agreement, which entered into force in 2016 [4,5].

In addition to the studies and discussions conducted, technical committees within the International Organization for Standardization (ISO) have published three key standards: ISO 14064-1, ISO 14064-2 and ISO 14064-3 to provide standardized methods for tracking and managing greenhouse gas emissions. ISO 14064-1 outlines the principles and requirements for quantifying and reporting greenhouse gas emissions and removals at the organizational level [6]. ISO 14064-2 focuses on project-level activities, providing guidance for calculating, monitoring, and reporting efforts aimed at reducing or eliminating emissions [7]. ISO 14064-3 defines the criteria and processes for validating or verifying greenhouse gas declarations and managing assurance activities [8].

Although systematic efforts to monitor greenhouse gas emissions in Turkey began relatively recently, an official inventory has been maintained since 1990. In 1990, total greenhouse gas emissions were 228.4 million tons of CO_2e, and this figure rose to 558.3 million tons by 2022 [9]. While per capita CO_2e emissions were calculated as 4.1 tons/person in 1990, this value increased to 6.5 tons/person by 2023 [10]. The total greenhouse gas emissions for the Republic of Turkey are projected to reach 1,244.13 million tonnes of CO_2e by 2030, based on estimates derived from an artificial neural network model. This result significantly exceeds Turkey's commitment of 929 Mt CO_2e emissions for the year 2030, as declared at the Paris Climate Summit [11].

Carbon dioxide (CO_2) constitutes 76% of anthropogenic greenhouse gas emissions, followed by methane (CH_4) at 16%, nitrogen oxides (NO_x) at 6%, and fluorinated gases at 2%. The total amount of greenhouse gases emitted directly or indirectly by individuals, organizations, activities, or products is referred to as the carbon footprint [12]. Figure 1 illustrates the direct and indirect sources contributing to the carbon footprint, along with associated mitigation strategies.

In this context, the carbon footprint has emerged as a widely recognized indicator of greenhouse gas emissions. It has played a significant role in raising public awareness of climate change and the environmental impacts of global warming, and is expected to continue doing so in the future [13].

*Sorumlu Yazar/Corresponding Author: fatihgencturk@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 118 – 132, 2025
DOI: 10.62301/usmtd.1712846

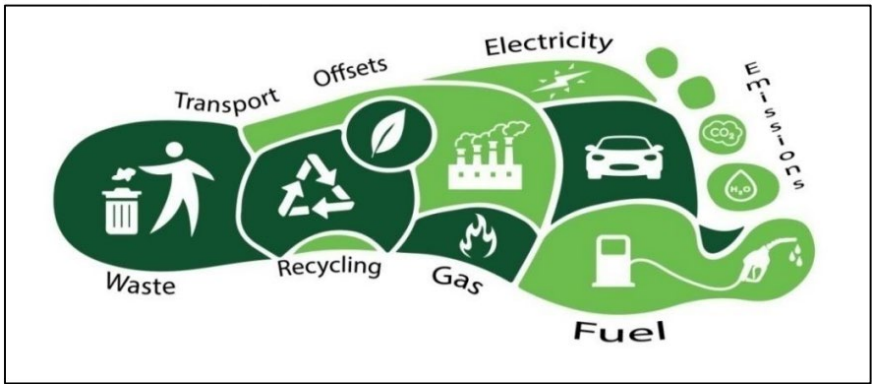


Figure 1. Overview of carbon footprint drivers and mitigation strategies

This study aims not only to examine the concept of the carbon footprint from technical and methodological perspectives, but also to investigate the role of information and communication technologies (ICT) from an interdisciplinary perspective. It offers a comparative evaluation of existing calculation methods, technological advancements, and emission mitigation strategies discussed in the literature, while also highlighting the innovative opportunities enabled by ICT-based systems. Accordingly, the primary objective of the study is to demonstrate how digital technologies can be more effectively leveraged to monitor and manage carbon footprints in alignment with sustainability goals.

2. Carbon Footprint Calculation Methodology

Determining carbon footprints is critically important for guiding sustainability initiatives, evaluating and enhancing the effectiveness of emission reduction strategies, mitigating global climate change, and understanding environmental impacts. To support these objectives, a variety of approaches, methodologies, and tools have been developed from simple online calculators for individuals and businesses to advanced, life cycle-based scientific models. These methods allow carbon footprint assessments to be conducted across multiple levels, including countries, cities, organizations, companies, households, and individuals [13]. However, despite the availability of sophisticated tools, different carbon calculators may yield inconsistent results when using similar inputs [14]. In some cases, the estimated values can differ by several metric tons per activity [15].

According to the EN ISO 14064-1 standard, greenhouse gas emissions that constitute a carbon footprint are categorized into three main scopes. Scope 1 (direct emissions) refers to emissions from sources owned or directly controlled by the organization, such as on-site fuel combustion. Scope 2 (indirect energy-related emissions) includes emissions generated during the production of electricity, heat, or steam purchased from external providers. Scope 3 (other indirect emissions) covers emissions indirectly resulting from the organization's activities but occurring from sources not owned or controlled by the organization. This includes upstream and downstream processes such as supply chain operations, product use, and waste management [4,6].

In this context, a review of existing studies shows that carbon footprint calculations typically begin by identifying the relevant activities, products, and processes. Next, the scopes and their contents are defined. Once the scopes are established, data related to the selected activities are collected. These data are then converted into carbon dioxide equivalents (CO_2e) using appropriate unit conversions. Calculations are performed by applying emission factors corresponding to the type of energy or resource consumed. Finally, the total carbon footprint is obtained by aggregating the calculated emission values.

Although general steps for carbon footprint calculation are commonly applied in the literature, there is still no universally accepted method, standard, or consensus on how to calculate personal carbon footprints [16,17]. In most cases, carbon footprint is calculated by multiplying the activity data that cause emissions with the relevant emission factor (see Equation 1).

$$E_S = AD_S \; x \; EF_S \qquad (1)$$

*Sorumlu Yazar/Corresponding Author: fatihgencturk@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 118 – 132, 2025
DOI: 10.62301/usmtd.1712846

Greenhouse gas emissions from a specific source (Es) are calculated by multiplying the activity data (ADs) associated with that source by the relevant emission factor (EFs). Activity data represent a quantifiable measure of resource use, such as liters of gasoline or kilowatt-hours of electricity, while emission factors are coefficients used to convert activity data into greenhouse gas emissions.

Once the total greenhouse gas emissions (CO_2, CH_4, NO_x) from all sources are calculated, they are aggregated and expressed as carbon dioxide equivalents (CO_2e) (see Equation 2). CO_2e is a standardized unit commonly used to represent greenhouse gas emissions. It quantifies the global warming potential of different gases by converting them into the equivalent amount of CO_2 [18].

$$E_s = \sum_{fuels} \left(E_{s,fuels}\right) \quad (2)$$

The emission factor and oxidation factor values used in carbon footprint calculations may vary depending on the specific methodology, standard, or guideline employed. In most studies, commonly referenced sources include IPCC Tier methods [12,19–21], the GHG Protocol developed by WRI/WBCSD [22], Annexes of relevant international agreements [23], and the UK Department for Environment, Food and Rural Affairs (DEFRA) [12]. Table 1 presents selected DEFRA conversion factors for 2024 for illustrative purposes [24].

Table 1. DEFRA 2024-based greenhouse gas conversion factors by fuel type for scope 1 and scope 2 emissions

	Fuel Type	Unit	kg CO_2e	kg CO_2	kg CH_4	kg N_2O
Scope 1	Natural Gas	m^3	2.04542	2.04140	0.00307	0.00095
	Coal	tonnes	2904.95234	2632	240.352	32.60034
	Gasoline	litres	2.0844	2.07047	0.00806	0.00587
	Diesel	litres	2.51279	2.47960	0.00029	0.0329
Scope 2	Electricity	kWh	0.20705	0.20493	0.0009	0.00122

The DEFRA conversion factors, which are widely used in organizational carbon footprint assessments in the UK, calculate emissions by multiplying activity data with the corresponding emission factors. Similarly, the IPCC 2006 methodology, developed as part of international greenhouse gas reporting standards, uses an emission factor-based approach to quantify emissions from energy use and industrial processes. The IPCC 2006 guidelines define three calculation tiers, namely Tier 1, Tier 2, and Tier 3, which represent different levels of methodological complexity and data specificity for estimating greenhouse gas emissions [25].

The Tier-1 method is easier to apply than the others. In this method, the emission value is calculated by multiplying the amount of fuel consumed by the emission factor and the oxidation factor (see Equation 3). The emission factor is a coefficient that varies by fuel type and is typically expressed in kilograms of gas per terajoule (kg/TJ). It represents the amount of greenhouse gases released into the atmosphere as a result of fuel combustion. In Tier 1 calculations, the oxidation factor is taken as 1 [19].

$$Emission = Fuel\ Amount\ x\ Emission\ Factor\ x\ Oxidation\ Factor \quad (3)$$

In the Tier 2 method, country-specific or region-specific emission factors and fuel properties are used to improve accuracy. The emission value is calculated as shown in Equation 4. In this equation, the calorific value refers to the lower heating value of the fuel and is expressed in kilocalories (kcal). MW represents the molar mass of carbon and carbon dioxide, expressed in grams per mole (g/mol). The symbol C denotes the carbon content percentage of the fuel by mass.

$$Emission = Fuel\ Amount\ x\ Calorific\ Value\ x\ Carbon\ Content\ x\ Oxidation\ Factor\ x\ \frac{MWco_2}{MWc}\ x\ C \quad (4)$$

*Sorumlu Yazar/Corresponding Author: fatihgencturk@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 118 – 132, 2025
DOI: 10.62301/usmtd.1712846

The Tier 3 method is a facility-level approach that incorporates country-specific methodologies and typically requires more detailed input data. This approach accounts for various parameters such as the type of fuel used, combustion technology, emission control methods, and maintenance practices. It also considers fuel characterization and operating conditions, all of which must be specifically determined for each facility [19].

The scope and depth of calculations vary significantly across carbon footprint assessment studies. Developed carbon calculators are designed to estimate not only individual carbon footprints but also those of organizations and products. Mulrow et al. [26] evaluated 31 publicly accessible calculators capable of estimating personal carbon footprints. Their analysis showed that simpler tools generally calculate emissions based solely on energy-related activities. In contrast, more comprehensive calculators also incorporate lifestyle and consumption factors such as diet and travel behavior. Some applications further provide tailored recommendations to help users reduce their carbon emissions. The authors emphasized that input parameters vary widely among personal carbon footprint calculators, and that no standardized framework has yet been established.

Carbon footprint calculators are also used at the organizational level. In a study conducted by Harangozo and Szigeti [27], the carbon footprint of a hypothetical company was evaluated under three scenarios involving varying energy consumption levels and supplier activity characteristics. To assess the consistency of the results, the researchers examined whether the calculators produced similar outputs when provided with identical input data. The findings revealed that the reliability of the calculators was insufficient. The authors attributed this inconsistency primarily to variations in calculation methods and emission factors [13].

Product-based carbon footprint assessments differ from individual or organizational assessments in both scope and methodology. These calculations are more complex, as they account for greenhouse gas emissions generated throughout the entire life cycle of a product. Due to the dependence on supply chain structures and life cycle stages, it is not feasible to develop a universal product-based calculator that accommodates the wide variety of product types and categories [13]. Kim and Neff [28] analyzed eight carbon calculators designed to estimate and communicate indirect emissions from food consumption in the United States of America (USA). Their study revealed that most calculators tend to overlook diet-related emissions. Moreover, they highlighted that dietary choices significantly contribute to indirect greenhouse gas emissions, underscoring the need for more robust and comprehensive calculation methodologies.

3. Carbon Emission Pricing Methods

There are two primary methods for pricing carbon emissions: carbon taxation and carbon trading. Carbon trading is a market-based approach to reducing emissions, and the Emissions Trading System (ETS) is its most widely implemented regulatory form. The carbon tax imposes a fixed price on greenhouse gas emissions for all emitters, providing a direct economic incentive to reduce emissions. In contrast, carbon trading involves setting an overall emissions cap and allowing participants who exceed their allocated limits to purchase additional emission allowances. This system requires the establishment of a functioning market for the trading of emission permits [29,30].

Although both carbon taxation and carbon trading aim to reduce emissions, they differ significantly in terms of policy design, implementation mechanisms, and potential economic impacts. These differences are summarized in Table 2 [30,31].

*Sorumlu Yazar/Corresponding Author: fatihgencturk@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 118 – 132, 2025
DOI: 10.62301/usmtd.1712846

Table 2. Comparative overview of carbon tax and carbon trading: key features and implementation characteristics

Feature	Carbon Tax	Carbon Trading
Emission Reduction Potential	Emission reduction is not guaranteed due to the "polluter pays" approach.	More likely to reduce emissions as it involves emission caps.
Regulatory Integration	Easier to implement as it can be integrated into the existing tax system.	Requires detailed design as it introduces a new trading mechanism.
Applicability	Can be applied to all sectors and companies emitting carbon.	Requires differentiation among sectors.
Pricing Mechanism	Prices are clear and transparent.	May create a market prone to manipulation.
Cost Impact	Creates additional costs for firms.	Offers trading opportunities that may offset costs.

Carbon prices vary significantly across countries. As of April 2024, carbon taxes and ETSs have been adopted in 75 countries worldwide. Together, these instruments cover approximately 24% of global greenhouse gas emissions. Revenues from carbon pricing mechanisms surpassed $100 billion in 2023, marking a notable increase compared to 2022. Carbon pricing is implemented through a range of policy frameworks, either at the national level or within subnational jurisdictions. As of April 1, 2024, direct carbon prices under existing tax and trading schemes range from as low as $1 to over $160 per tonne of CO_2e. In countries such as Ukraine, Argentina, and Indonesia, carbon prices remain at the lower end of the scale (approximately $1 to 5/t$CO_2e$). In contrast, Switzerland, Norway, Finland, and the Netherlands report prices ranging from $90 to $120/t$CO_2e$. However, most existing prices remain well below the levels required to align with the Paris Agreement targets, which are estimated to be at least $226 to 385/t$CO_2e$ for 2024. According to the World Bank's 2024 report, countries with operational carbon pricing schemes are illustrated in Figure 2 [32].

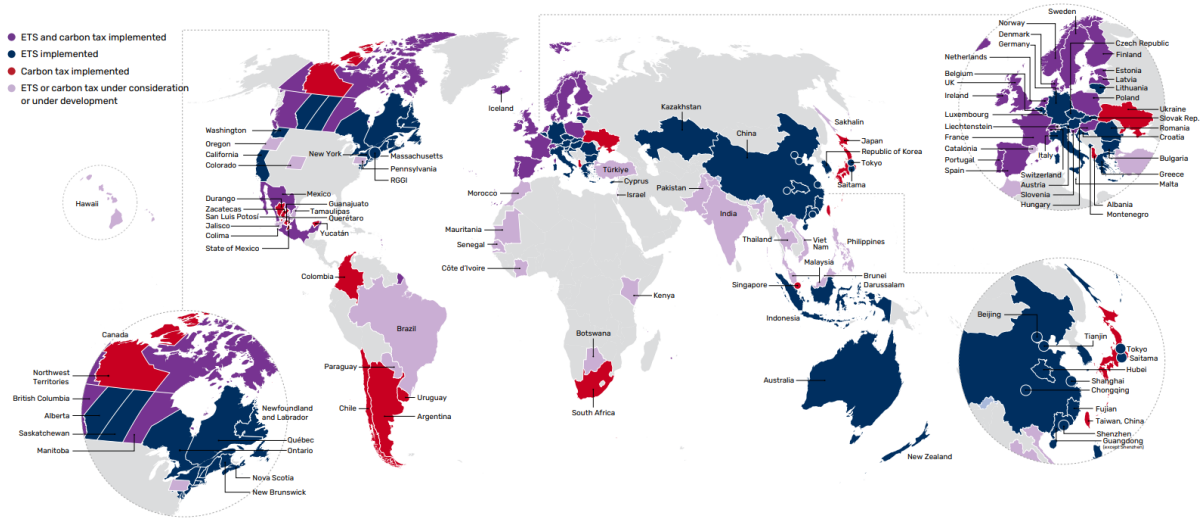


Figure 2. Worldwide Implementation of Carbon Pricing Mechanisms in 2024: ETS and Carbon Taxes

The unit cost of carbon pricing varies significantly across countries worldwide. A carbon tax is a policy instrument that imposes a fee on entities based on the amount of carbon dioxide they emit, aiming to reduce overall greenhouse gas emissions [33]. It is a widely recognized pricing mechanism and may be implemented independently or alongside emissions trading schemes. This approach offers emission-reducing entities several mitigation strategies, such as switching to low-carbon fuels, improving energy efficiency, or altering production processes and input sources. The concept of a carbon tax was first

*Sorumlu Yazar/Corresponding Author: fatihgencturk@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 118 – 132, 2025
DOI: 10.62301/usmtd.1712846

introduced in academic and policy discussions during the 1970s and was initially implemented by Northern European countries in the 1990s. Finland became the first country to apply a national carbon tax in 1990, targeting fossil fuel consumption in sectors such as transportation, electricity generation, and heating [30].

In addition to carbon taxation, carbon trading is a market-based mechanism designed to lower the overall cost of reducing greenhouse gas emissions from human activities [5]. It contributes to both air quality improvement and climate change mitigation, which are essential goals in addressing global warming. The carbon trading system requires companies to limit their greenhouse gas emissions to predefined thresholds and to verify their emissions through certifiable units that can be traded in the market. Within this system, emission allowances can be bought and sold among regulated entities. These certificates represent a quantifiable amount of emissions; for example, one allowance may correspond to one metric ton of carbon dioxide equivalent (CO_2e) [5,34].

Historically, the first emissions trading scheme originated from the Acid Rain Program, which was established under the Clean Air Act Amendments of 1990 to reduce sulfur dioxide (SO_2) emissions, especially within the electricity generation sector [35]. Subsequently, the European Union Emissions Trading System (EU ETS), recognized as the world's largest carbon market, was launched in 2005 [30,36].

Expanding beyond internal market mechanisms, the European Green Deal proposes the implementation of carbon border adjustment mechanisms to extend carbon taxation to products imported into the European Union (EU). It outlines targets across key sectors, including clean energy transition, biodiversity, sustainable industrial and food systems, construction and renovation, and zero pollution. Among the instruments developed to meet these objectives, the carbon border adjustment mechanism is designed to impose a carbon price on imports from non-EU countries in order to prevent carbon leakage and reduce global emissions [37]. This mechanism introduces a carbon levy on specific imported product groups, such as cement, fertilizers, aluminum, electricity, and iron and steel products. Over time, the list has expanded to include raw materials like agglomerated iron ore, ferrochrome, ferromanganese, and kaolin. In addition, manufactured goods such as hydrogen, screws, and bolts, which fall under the iron and steel category, have also been added to the scope. The implementation of this mechanism is expected to negatively impact the export competitiveness of carbon-intensive products and sectors, particularly in countries that trade extensively with the EU. However, during the transitional phase of the mechanism, which will run from 2023 to 2026, no financial payments are expected [38].

4. Literature Review

According to IPCC's Sixth Assessment Report published in 2023, as of 2019, 34% of global greenhouse gas emissions originated from the energy sector, 24% from industry, 22% from agriculture, forestry, and land use, 15% from transportation, and 6% from buildings [39]. These findings underscore the critical importance of transitioning to low-carbon technologies, particularly within the energy sector, which is the largest source of emissions. For instance, while coal-fired power plants are estimated to emit between 675 and 1689 gCO_2e/kWh over their life cycle [40], renewable sources such as offshore wind energy can reduce this figure to as low as 5.3-13 gCO_2e/kWh [41]. Furthermore, life cycle assessments of solar and wind systems demonstrate that improvements in the design and deployment phases of these technologies can play a decisive role in reducing emissions [42]. Comparative analyses covering all stages of energy production reveal that fossil fuels have the highest life cycle emissions, whereas nuclear, hydroelectric, and wind power exhibit the lowest emission levels [43].

As digitalization accelerates, the share of the ICT sector in global greenhouse gas emissions is increasing significantly, and its long-term impact is expected to become even more substantial. Indeed, some projections indicate that the ICT sector's contribution to global greenhouse gas emissions could reach as high as 23% by 2030 [44]. Freitag et al. [45] report that the ICT sector currently accounts for between 1.8% and 2.8% of global emissions. However, when the full upstream supply chain and life-cycle emissions are considered, this estimate may increase to between 2.1% and 3.9%. At this stage, selecting regions and time periods with lower carbon intensity becomes essential for minimizing emissions. Dodge et al. [46] propose a framework for evaluating the carbon intensity of software on cloud platforms and emphasize that the environmental impact of computing operations varies depending on when and

*Sorumlu Yazar/Corresponding Author: fatihgencturk@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 118 – 132, 2025
DOI: 10.62301/usmtd.1712846

where they are executed. Accordingly, users are encouraged to utilize cloud resources during times and in locations with lower carbon intensity to minimize their environmental impact.

The environmental footprint of artificial intelligence (AI) systems has emerged as an increasingly critical issue. Kirkpatrick [47] highlights that both the training and deployment of large-scale AI models entail substantial energy consumption, consequently resulting in significant CO_2 emissions. This impact is further exacerbated when data centers are powered by fossil fuel-based energy sources. Similarly, Yu et al. [48] estimate that the cumulative annual carbon footprint of 79 AI systems released between 2020 and 2024 may reach 102.6 $MtCO_2e$, highlighting the necessity of regulatory interventions in the sector. The environmental implications of AI-assisted software development have also become a key area of investigation. Cheung et al. [49] show that software development using large language models (LLMs) produces, on average, 32.72% more carbon emissions than conventional manual development approaches.

Blockchain technologies and bioinformatics have become increasingly scrutinized due to their high energy demands. Bitcoin's annual electricity consumption is reported to be approximately 45.8 TWh, with corresponding carbon emissions ranging from 22.0 to 22.9 $MtCO_2e$ [50]. These levels are comparable to the annual greenhouse gas emissions of certain nations. In bioinformatics, computationally intensive tasks such as genetic analyses and biological simulations have been shown to consume substantial amounts of energy, which in turn generates a considerable carbon footprint. To address this issue, Grealey et al. [51] suggest various mitigation strategies, including software optimization, careful hardware selection, and enhancements in data center efficiency.

While ICT is an emissions-intensive sector, it also holds substantial potential to serve as a tool for emissions mitigation. This duality is evident across sectors and applications, typically supported by empirical research and field implementations. In particular, AI and machine learning (ML) have been applied in industries such as manufacturing, agriculture, and logistics to reduce carbon emissions.

In a large-scale empirical study, Lu and Liau [52] found that a 1% increase in AI adoption across industrial enterprises resulted in a 0.0395% reduction in sectoral carbon emission intensity. Likewise, Rajendiran et al. [53] showed that ML-based forecasting and logistics optimization algorithms, such as Random Forest and Gradient Boosted Machines, can effectively reduce transport-related emissions in the e-commerce sector.

Ji et al. [54] highlighted the potential of digital agriculture for both direct and indirect emission mitigation. By integrating IoT, big data, blockchain, and modular greenhouse systems, the study demonstrated how agricultural productivity can be enhanced while enabling the quantification and trading of carbon sinks. Their proposed architecture combines renewable energy systems with green production and finance modules. Nonetheless, infrastructural disparities and limited digital access in rural areas remain key barriers to implementation.

In a field-based case study, Andrae et al. [55] proposed a practical energy-saving model for the ICT supply chain. Accelerated life cycle assessments were conducted on four ICT products (one modem and three access modules), identifying high-impact components. On-site energy audits were then carried out in accordance with the International Performance Measurement and Verification Protocol (IPMVP), followed by the implementation of energy conservation measures. These efforts led to annual savings of 27,000 MWh of energy and 25,700 tons of CO_2e, amounting to a 1% reduction in the operator's Scope 3 emissions.

In addition to these applications, it has been demonstrated that artificial intelligence can, in certain domains, have a lower carbon footprint than human labor. Tomlinson et al. [56] report that AI systems emit 130 to 2,900 times less CO_2 than humans when performing tasks such as writing and image generation. However, these findings do not account for broader dimensions such as social implications and transformations in the labor market.

Personal carbon footprint analysis is regarded as an effective tool for shaping individual decision-making and enhancing environmental awareness. Nevertheless, estimates based on static methodologies have shown limited effectiveness in influencing behavioral change [26]. In this context, Rahman et al.

*Sorumlu Yazar/Corresponding Author: fatihgencturk@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 118 – 132, 2025
DOI: 10.62301/usmtd.1712846

[15] proposed a specialized framework called the Open Carbon Footprint Framework, which could serve as a foundation for the development of carbon footprint calculation applications. This framework provides developers with access to a cloud-based information infrastructure that supports the integration of real-time sensor data, offering a dynamic and responsive resource for emissions estimation. Based on this framework, an application named the Ubiquitous Carbon Footprint Calculator was developed, enabling users to calculate their carbon footprints. This application allows users to monitor their emissions regardless of location and supports conscious decision-making related to environmental impact. Furthermore, the design and features of a mobile Global Positioning System (GPS) application, developed specifically for the iPhone platform, are detailed. This GPS-based application is capable of suggesting the most environmentally friendly route to the user, thereby enhancing fuel efficiency and reducing emissions. The functionality and applicability of the study are demonstrated through two example scenarios. In the first scenario, a graduate student working in a laboratory equipped with smart sensors is shown to reduce their carbon footprint by optimizing indoor heating. Temperature and humidity data obtained from the student's smartphone are combined with environmental data from external sensors to calculate real-time emissions. The application provides instant feedback and recommends actionable strategies, such as adjusting the ambient temperature, to lower energy-related carbon output. The second scenario focuses on the behavior of a sales representative who is frequently on the move. The system tracks the representative's travel routes via GPS, generates speed profiles to evaluate driving efficiency, and proposes alternative, lower-emission routes. In addition, the application encourages emission reduction by enabling users to carpool with others who share similar travel paths, further contributing to sustainable mobility.

Andersson [16] developed a mobile application in Sweden that estimates users' greenhouse gas emissions through a hybrid approach combining user-provided data, official government records, and financial transaction data. The system utilized these three data sources to generate individualized carbon footprint estimates. In particular, users could link their private bank accounts and credit cards to the app, allowing financial transactions to be analyzed and matched with Sweden-specific multi-regional environmental input-output data. By categorizing spending across different consumption domains, the system enabled the estimation of emissions associated with various activities. All data transmissions were encrypted and stored on secure servers in compliance with Swedish and European Union regulations. Users' expenditures were classified based on financial transaction data, allowing for detailed insights into their carbon profiles. Depending on the policies of individual banks, the system could access transaction histories ranging from three months to five years. This historical data allowed users to monitor the evolution of their carbon footprints over time. Emission estimates were calculated using GWP100 (Global Warming Potential over 100 years) conversion factors integrated into the application.

One of the key components of a carbon footprint is emissions generated by transportation activities. Therefore, accurately tracking and monitoring transportation-related emissions through carbon footprint calculators is essential for reducing global greenhouse gas outputs. In this context, Ajufo and Bekaroo [57] developed a personal, transportation-based carbon footprint calculator called TCTracker, which eliminates the need for manual data entry. The mobile application leverages GPS functionality and built-in artificial intelligence to monitor user behavior and estimate emissions. In a related effort, Wang et al. [58], addressed carbon reduction in the petroleum distribution network by formulating and solving a low-carbon inventory routing problem. Their model aims to minimize total costs while accounting for carbon emissions. To solve the model, they proposed a hybrid approach that combines an adaptive genetic algorithm with a greedy algorithm, and they validated its performance using a practical case study. The study was structured in two parts: the first part examined model effectiveness under varying carbon tax scenarios, and the second part applied the model to optimize routing for petroleum tankers with different loading capacities. Similarly, Jabali et al. [59], investigated the trade-offs among carbon emissions, travel time, and fuel consumption in time-dependent vehicle routing problems using a tabu search algorithm. Their study also discussed the implications of setting carbon emission thresholds in transportation planning. In the aviation sector, Tsai et al. [60], introduced a mixed activity-based costing decision model for green fleet planning. They evaluated how carbon emissions affect the operating costs of airline logistics under the regulatory constraints of the European Union Emissions Trading Scheme.

In addition to academic research, many companies also engage in efforts to reduce their carbon footprints by providing consumers with sustainability ratings or certifications for the services and

*Sorumlu Yazar/Corresponding Author: fatihgencturk@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 118 – 132, 2025
DOI: 10.62301/usmtd.1712846

products they offer. Such initiatives not only promote consumer awareness but also help influence user preferences by offering environmental information about the purchased product or service. For example, Booking.com [61] has classified the sustainability ratings of many hotels listed in its database into four levels, based on 29 criteria related to waste management, water consumption, energy and greenhouse gas emissions, location and community impact, and nature conservation. This categorization enables consumers to make informed decisions regarding the environmental performance of their accommodation choices. In the aviation sector, Travalyst [62], in collaboration with Google, has developed the "Travel Impact Model" (TIM), a framework that calculates the life cycle emissions of flights. Similarly, platforms such as Skyscanner [63] and Google Flights [64] provide users with estimated carbon emission data for selected flights. These calculations are based on several criteria, including travel distance, cabin class, baggage allowance, and aircraft type. These examples illustrate how environmental transparency in consumer services contributes to the foundation of individual carbon emission tracking and management. While the number of such implementations continues to grow, their common objective is to empower individuals to make environmentally responsible decisions in daily life.

5. Results and Discussion

This section provides a multidimensional analysis of the methodological diversity in carbon footprint assessment approaches and the current trends shaping this field. In addition to the dominant traditional, static calculation methods in the literature, newer models supported by real-time data and emerging technologies are comparatively examined. The potential of user-centered applications designed to influence daily lifestyle choices, the contribution of corporate transparency in product and service-level emission data to decision-making processes, and the dual role of ICT in both increasing and reducing carbon emissions are also discussed. Throughout the section, ICT-based solutions developed across different sectors are evaluated to explore their contribution to environmental sustainability from a broad and integrated perspective.

The concept of carbon footprinting remains fundamental to assessing the environmental impact of individuals, organizations, and products. Conventional carbon footprint calculation methods typically aim to quantify this impact by measuring greenhouse gas emissions arising from energy consumption, transportation, production, and related activities. In addition to these traditional approaches, recent studies have introduced dynamic models and analytical techniques that consider temporal variability in environmental impacts and emissions to obtain more precise results [16,57]. These contemporary approaches emphasize the use of real-time data enabled by emerging technologies, including sensor systems and smart devices. As a result, research in carbon footprinting is advancing significantly, offering deeper insights into environmental effects, supporting the achievement of sustainability goals, and contributing to the development of more effective strategies to mitigate global climate change. This dual function of ICT highlights the necessity for balanced strategies that can simultaneously leverage its monitoring and optimization capabilities while minimizing its environmental costs.

Traditional studies on carbon footprint assessment lack the ability to dynamically update data in response to user actions due to their static structures. Although there are sector-based or scenario-specific dynamic applications such as route optimization for vehicles [55,58] and electricity consumption optimization[52,54] that consider the carbon footprint, comprehensive systems capable of guiding daily life remain insufficient. In this context, the development of a smart carbon optimization application that incorporates carbon footprint considerations and facilitates daily living, potentially even offering actionable recommendations, would be highly beneficial. By processing data collected through personal smartphones, smartwatches, or various sensors, a dynamic system can be developed that not only offers users novel and unconventional experiences but also contributes to reducing their carbon footprint. A multi-criteria application that provides guidance on topics such as public versus private transportation, fuel types like gasoline or diesel, route planning, accommodation and facilities, food choices, lighting, and heating, based on cost, carbon footprint, and time, could significantly enhance both awareness and the adoption of more environmentally friendly lifestyles at the individual level. Furthermore, environmentally conscious behaviors such as support for recycling should also be taken into consideration in reducing carbon footprints. For example, recyclable waste deposited in smart recycling machines should be recognized as a positive contributor to an individual's overall carbon reduction.

*Sorumlu Yazar/Corresponding Author: fatihgencturk@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 118 – 132, 2025
DOI: 10.62301/usmtd.1712846

In addition to behavior-responsive applications, infrastructure-level strategies play a critical role in reducing the environmental impact of digital systems. One promising solution is the broader adoption of edge computing architectures in place of centralized cloud systems. By enabling local data processing, edge computing minimizes the energy demands of large-scale data centers while also improving latency. This approach is particularly relevant in applications involving IoT, artificial intelligence, and mobile services, where both responsiveness and efficiency are key concerns. Reducing ICT-related emissions also requires a shift toward low-carbon energy sources. Data centers and digital platforms should prioritize the use of electricity generated from renewable energy. Moreover, geographical and temporal optimization—executing high-energy tasks in regions and at times with lower grid carbon intensity—can play a significant role in emission reduction. Encouraging cloud service providers to implement carbon-aware scheduling and infrastructure planning can help align digital growth with climate goals.

The implementation of environmentally friendly policies by businesses can pave the way for more comprehensive efforts and contribute to the development of emission reduction mechanisms. By transparently presenting sustainability assessments of their products and services or disclosing emission data obtained through product life cycle analyses, businesses can play an active role in decision-making processes. The data provided by companies may serve as a valuable resource for the learning mechanisms of intelligent systems. Furthermore, such data can help lay the groundwork for autonomous software that functions as an assistant by evaluating not only time and cost but also emission-related expenses. Moreover, integrating energy-efficient coding practices and carbon-aware software design into corporate digital services could enhance environmental performance across the software lifecycle. Carbon labeling of software products may also help users and institutions make more informed choices.

Another noteworthy point is that ICT has both increasing and decreasing effects on the carbon footprint across various sectors. For instance, the use of artificial intelligence systems can be considered a significant source of emissions [48,49]. However, these systems can also contribute to emission reduction by enhancing efficiency [58,59]. Similarly, while energy efficiency improvements in the manufacturing industry can help control carbon emissions [52], the high energy consumption of data centers may increase the environmental burden [47]. In the retail and e-commerce sectors, AI-supported supply chain planning and demand forecasting algorithms help prevent overstocking, thereby reducing emissions related to production and transportation [53]. In the transportation and logistics sector, GPS-based route optimization and fleet management systems reduce fuel consumption and thus contribute to lower emissions [59,60]. Nevertheless, the carbon footprint of ICT-based solutions themselves, particularly stemming from the hardware and infrastructure used in data processing, storage, and transmission, should not be overlooked [49,50]. This dual impact reveals that ICT can serve both as a means of promoting sustainability and as a potential source of environmental concern. Beyond infrastructure and corporate design, digital tools at the individual level should evolve from merely providing information to actively guiding low-carbon behavior. For instance, mobility apps can suggest routes based on minimal emissions rather than just cost or speed, while e-commerce platforms could display the carbon impact of products to support sustainable consumption.

6. Conclusion

This study examined the evolution of carbon footprint assessment approaches, focusing particularly on the integration of ICT into sustainability efforts. The findings emphasize a shift from conventional, static models to dynamic, user-responsive, and data-driven systems supported by real-time analytics. ICT-based tools have demonstrated significant potential in enhancing carbon monitoring, guiding behavioral changes, and optimizing resource usage across sectors.

Nevertheless, ICT itself is a contributor to global emissions due to the high energy demands of data centers, cloud infrastructures, and AI-based applications. This dual role of ICT as both an enabler and a driver of emissions calls for balanced strategies that simultaneously leverage its capabilities and minimize its environmental footprint. Key recommendations include the adoption of edge computing architectures, carbon-aware infrastructure planning, renewable energy integration, and energy-efficient software development practices. Furthermore, encouraging corporate transparency, carbon labeling of digital services, and behavior-oriented user interfaces are critical in aligning technological innovation with sustainability goals.

*Sorumlu Yazar/Corresponding Author: fatihgencturk@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 118 – 132, 2025
DOI: 10.62301/usmtd.1712846

Future research should evaluate the real-world feasibility, scalability, and cross-sector applicability of these approaches. It is essential to examine how ICT-driven systems can be effectively embedded into institutional, infrastructural, and behavioral frameworks. Policy incentives, regulatory support, and interdisciplinary collaboration will play a vital role in ensuring that these tools are not only technologically sound but also socially acceptable and ethically robust. Exploring how individuals respond to carbon-aware digital systems and how these influence long-term behavioral change also presents a valuable direction for further investigation.

In conclusion, the sustainable digital transformation of carbon footprint assessment requires more than technological advancement. It demands systemic thinking, collective responsibility, and alignment between digital innovation and environmental ethics. ICT should be seen not only as a source of impact but as an integral part of the solution in the transition toward a low-carbon future.

7. References

[1] Y. Serengil, Küresel Isınma ve Olası Ekolojik Sonuçları, İstanbul Üniversitesi Orman Fakültesi Dergisi 45 (1-2) (1995) 135-152.

[2] M. Türkeş, IPCC İklim Değişikliği 2013: Fiziksel Bilim Temeli Politikacılar için Özet Raporundaki Yeni Bulgu ve Sonuçların Bilimsel Bir Değerlendirmesi, İklim Değişikliğinde Son Gelişmeler: IPCC 2013 Raporu Paneli Bildiriler Kitapçığı, İstanbul Politikalar Merkezi, Sabancı Üniversitesi, İstanbul (2013) 8-18

[3] B. Davarcıoğlu, Küresel İklim Değişikliği ve Uyum Çalışmaları: Türkiye Açısından Değerlendirilmesi, Mesleki Bilimler Dergisi 7 (2) (2018).

[4] M.M. Yatarkalkmaz, M.B. Özdemir, The calculation of greenhouse gas emissions of a family and projections for emission reduction, Journal of Energy Systems 3 (2019) 96–110. https://doi.org/10.30521/jes.566516.

[5] E. Bilgiç, İklim Değişikliği İle Mücadelede Emisyon Ticareti ve Türkiye Uygulaması, Uzmanlık Tezi, T.C. Çevre ve Şehircilik Bakanlığı Strateji Geliştirme Başkanlığı, Ankara, (2017).

[6] TS EN ISO 14064-1, Greenhouse Gases-Part 1: Specification with guidance at the organization level for quantification and reporting of greenhouse gas emissions and removals, Turkish Standards Institution, Ankara, Turkey, (2007).

[7] TS EN ISO 14064-2, Greenhouse Gases-Part 2: Specification with guidance at the project level for quantification, monitoring and reporting of greenhouse gas emission reductions or removal enhancements, Turkish Standards Institution, Ankara, Turkey, (2007).

[8] TS EN ISO 14064-3, Greenhouse Gases-Part 3: Specification with guidance for the validation and verification of greenhouse gas assertions, Turkish Standards Institution, Ankara, Turkey, (2007).

[9] TÜİK Kurumsal, Sera Gazı Emisyon İstatistikleri 1990–2022, https://data.tuik.gov.tr/Bulten/Index?p=Sera-Gazi-Emisyon-Istatistikleri-1990-2022-53701 (accessed 10.06.25).

[10] TÜİK Kurumsal, Sera Gazı Emisyon İstatistikleri 1990–2023, https://data.tuik.gov.tr/Bulten/Index?p=Sera-Gazi-Emisyon-Istatistikleri-1990-2023-53974 (accessed 20.05.25).

[11] H. Pabuçcu, T. Bayramoğlu, Yapay Sinir Ağları ile CO_2 Emisyonu Tahmini: Türkiye Örneği, Gazi Üni. İktisadi ve İdari Bilimler Fakültesi Dergisi 18 (3) (2016) 762–778

[12] A.K. Seyhan, M. Çerçi, IPCC Tier 1 ve DEFRA Metotları ile Karbon Ayak İzinin Belirlenmesi: Erzincan Binali Yıldırım Üniversitesi'nin Yakıt ve Elektrik Tüketimi Örneği, J. Nat. Appl. Sci. 26 (2022) 386–397. https://doi.org/10.19113/sdufenbed.1061021.

[13] F. Scrucca, G. Barberio, V. Fantin, P.L. Porta, M. Barbanera, Carbon Footprint: Concept, Methodology and Calculation, in: S.S. Muthu (Ed.), Carbon Footprint Case Studies: Municipal Solid Waste Management, Sustainable Road Transport and Carbon Sequestration, Springer, Singapore, 2021, pp. 1–31. https://doi.org/10.1007/978-981-15-9577-6_1.

*Sorumlu Yazar/Corresponding Author: fatihgencturk@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 118 – 132, 2025
DOI: 10.62301/usmtd.1712846

[14] J.P. Padgett, A.C. Steinemann, J.H. Clarke, M.P. Vandenbergh, A comparison of carbon calculators, Environmental Impact Assessment Review 28 (2008) 106–115. https://doi.org/10.1016/j.eiar.2007.08.001.

[15] F. Rahman, C. O'Brien, S.I. Ahamed, H. Zhang, L. Liu, Design and implementation of an open framework for ubiquitous carbon footprint calculator applications, Sustainable Computing: Informatics and Systems 1 (2011) 257–274. https://doi.org/10.1016/j.suscom.2011.06.001.

[16] D. Andersson, A novel approach to calculate individuals' carbon footprints using financial transaction data – App development and design, Journal of Cleaner Production 256 (2020) 120396. https://doi.org/10.1016/j.jclepro.2020.120396.

[17] T. Wiedmann, J. Minx, A definition of 'carbon footprint,' 1st ed., Nova Publishers, 2008.

[18] K. Valls-Val, M.D. Bovea, Carbon footprint in Higher Education Institutions: a literature review and prospects for future research, Clean Techn Environ Policy 23 (2021) 2523–2542. https://doi.org/10.1007/s10098-021-02180-2.

[19] A.S. Toröz, Determination of the carbon footprint of a port reception facility for ship generated waste, Master's Thesis, Istanbul Technical University, 2015.

[20] T. Atabey, The calculation of the carbon footprint: The city of Diyarbakir, Master's Thesis, Firat University, 2013.

[21] G. Binboğa, A. Ünal, Sürdürülebilirlik Ekseninde Manisa Celal Bayar Üniversitesi'nin Karbon Ayak Izinin Hesaplanmasina Yönelik Bir Araştirma, UİİİD (2018) 187–202. https://doi.org/10.18092/ulikidince.323532.

[22] L. Ozawa-Meida, P. Brockway, K. Letten, J. Davies, P. Fleming, Measuring carbon performance in a UK University through a consumption-based carbon footprint: De Montfort University case study, Journal of Cleaner Production 56 (2013) 185–198. https://doi.org/10.1016/j.jclepro.2011.09.028.

[23] K. Kumaş, A. Akyüz, A. Güngör, Burdur Mehmet Akif Ersoy Üniversitesi Bucak Yerleşkesi Yükseköğretim Birimlerinin Karbon Ayak İzi Tespiti, Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi (2019). https://doi.org/10.28948/ngumuh.598212.

[24]Greenhouse gas reporting: conversion factors 2024, GOV.UK (2024). https://www.gov.uk/government/publications/greenhouse-gas-reporting-conversion-factors-2024 (accessed 22.04.25).

[25] C. Peter, A. Fiore, C. Nendel, C. Xiloyannis, Improving the accounting of land-based emissions in Carbon Footprint of agricultural products: comparison between IPCC Tier 1, Tier 2 and Tier 3 approaches, LCA Food 2014 (2014). https://doi.org/10.1029/2001GB001812.

[26] J. Mulrow, K. Machaj, J. Deanes, S. Derrible, The state of carbon footprint calculators: An evaluation of calculator design and user interaction features, Sustainable Production and Consumption 18 (2019) 33–40. https://doi.org/10.1016/j.spc.2018.12.001.

[27] G. Harangozo, C. Szigeti, Corporate carbon footprint analysis in practice – With a special focus on validity and reliability issues, Journal of Cleaner Production 167 (2017) 1177–1183. https://doi.org/10.1016/j.jclepro.2017.07.237.

[28] B. Kim, R. Neff, Measurement and communication of greenhouse gas emissions from U.S. food consumption via carbon calculators, Ecological Economics 69 (2009) 186–196. https://doi.org/10.1016/j.ecolecon.2009.08.017.

[29] D. Nerudová, M. Dobranschi, Pigouvian Carbon Tax Rate: Can It Help the European Union Achieve Sustainability?, in: P. Huber, D. Nerudová, P. Rozmahel (Eds.), Competitiveness, Social Inclusion and Sustainability in a Diverse European Union: Perspectives from Old and New Member States, Springer International Publishing, Cham, 2016, pp. 145–159. https://doi.org/10.1007/978-3-319-17299-6_8.

[30] E. Toma, Potential impacts of a carbon tax on the day-ahead market prices and the electricity generation mix, Master's Thesis, Istanbul Technical University, 2020.

*Sorumlu Yazar/Corresponding Author: fatihgencturk@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 118 – 132, 2025
DOI: 10.62301/usmtd.1712846

[31] G. Bavbek, Carbon Taxation Policy Case Studies, EDAM Energy & Climate Change Climate Action Paper Series 2016/4 (2016)

[32] World Bank, State and Trends of Carbon Pricing 2024, World Bank, Washington, DC, (2024).

[33] S. Yıldız, Sürdürülebilir Kalkınma İçin Karbon Vergisi, Muhasebe ve Vergi Uygulamaları Dergisi / Journal of Accounting & Taxation Studies 10 (2017) 367–384. https://doi.org/10.29067/muvu.333079.

[34] H. Aliusta, B. Yılmaz, H. Kırlıoğlu, Küresel Isınmayı Önleme Sürecinde Uygulanan Piyasa Temelli İktisadi Araçlar: Karbon Ticareti ve Karbon Vergisi, İJMEB 12 (2016) 382–401.

[35] S.M. Schennach, The Economics of Pollution Permit Banking in the Context of Title IV of the 1990 Clean Air Act Amendments, Journal of Environmental Economics and Management 40 (2000) 189–210. https://doi.org/10.1006/jeem.1999.1122.

[36] K. Pamukçu, Küresel Emisyon Ticareti Sistemi İçin Bir Model: Avrupa Birliği Emisyon Ticareti Programı, İstanbul Üniversitesi Siyasal Bilgiler Fakültesi Dergisi 11 (2007) 17–42.

[37] B.E. Yüksel, M. Özcan, E. Ocaklı, Türkiye Gönüllü Karbon Piyasaları'nın Değerlendirilmesi, DÜBİTED 10 (2022) 10–25. https://doi.org/10.29130/dubited.1101215.

[38] S. Ünsal, The evaluation of the carbon tax implementation within the framework of Green Reconciliation in terms of the EU and Türkiye, Master's Thesis, Anadolu University, 2023.

[39] IPCC Core Writing Team [Calvin et al.], Climate Change 2023: Synthesis Report. Contribution of WGs I–III to the Sixth Assessment Report, H. Lee & J. Romero (eds.), IPCC, Geneva, Switzerland (2023). https://doi.org/10.59327/IPCC/AR6-9789291691647.

[40] M. Whitaker, G.A. Heath, P. O'Donoughue, M. Vorum, Life Cycle Greenhouse Gas Emissions of Coal-Fired Electricity Generation, Journal of Industrial Ecology 16 (2012) S53–S72. https://doi.org/10.1111/j.1530-9290.2012.00465.x.

[41] N.Y. Amponsah, M. Troldborg, B. Kington, I. Aalders, R.L. Hough, Greenhouse gas emissions from renewable energy sources: A review of lifecycle considerations, Renewable and Sustainable Energy Reviews 39 (2014) 461–475. https://doi.org/10.1016/j.rser.2014.07.087.

[42] D. Nugent, B.K. Sovacool, Assessing the lifecycle greenhouse gas emissions from solar PV and wind energy: A critical meta-survey, Energy Policy 65 (2014) 229–244. https://doi.org/10.1016/j.enpol.2013.10.048.

[43] D. Weisser, A guide to life-cycle greenhouse gas (GHG) emissions from electric supply technologies, Energy 32 (2007) 1543–1559. https://doi.org/10.1016/j.energy.2007.01.008.

[44] A.S.G. Andrae, T. Edler, On Global Electricity Usage of Communication Technology: Trends to 2030, Challenges 6 (2015) 117–157. https://doi.org/10.3390/challe6010117.

[45] C. Freitag, M. Berners-Lee, K. Widdicks, B. Knowles, G. Blair, A. Friday, The climate impact of ICT: A review of estimates, trends and regulations, (2021). https://doi.org/10.48550/arXiv.2102.02622.

[46] J. Dodge, T. Prewitt, R. Tachet Des Combes, E. Odmark, R. Schwartz, E. Strubell, A.S. Luccioni, N.A. Smith, N. DeCario, W. Buchanan, Measuring the Carbon Intensity of AI in Cloud Instances, in: 2022 ACM Conference on Fairness, Accountability, and Transparency, ACM, Seoul Republic of Korea, 2022, pp. 1877–1894. https://doi.org/10.1145/3531146.3533234.

[47] K. Kirkpatrick, The Carbon Footprint of Artificial Intelligence, Commun. ACM 66 (2023) 17–19. https://doi.org/10.1145/3603746.

[48] Y. Yu, J. Wang, Y. Liu, P. Yu, D. Wang, P. Zheng, M. Zhang, Revisit the environmental impact of artificial intelligence: the overlooked carbon emission source?, Front. Environ. Sci. Eng. 18 (2024) 158. https://doi.org/10.1007/s11783-024-1918-y.

[49] K.S. Cheung, M. Kaul, G. Jahangirova, M.R. Mousavi, E. Zie, Comparative Analysis of Carbon Footprint in Manual vs. LLM-Assisted Code Development, arXiv.Org (2025). https://doi.org/10.1145/3711919.3728678.

*Sorumlu Yazar/Corresponding Author: fatihgencturk@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 118 – 132, 2025
DOI: 10.62301/usmtd.1712846

[50] C. Stoll, L. Klaaßen, U. Gallersdörfer, The Carbon Footprint of Bitcoin, Joule 3 (2019) 1647–1661. https://doi.org/10.1016/j.joule.2019.05.012.

[51] J. Grealey, L. Lannelongue, W.-Y. Saw, J. Marten, G. Méric, S. Ruiz-Carmona, M. Inouye, The Carbon Footprint of Bioinformatics, Molecular Biology and Evolution 39 (2022) msac034. https://doi.org/10.1093/molbev/msac034.

[52] Y. Lu, Z. Liao, The influence of AI application on carbon emission intensity of industrial enterprises in China, Sci Rep 15 (2025) 12585. https://doi.org/10.1038/s41598-025-97110-3.

[53] G.R. Rajendiran, R. Ayyadurai, Reducing E-Commerce Carbon Footprint Through AI-Driven Warehouse and Supply Chain Optimization, Reduct. Footprint AI Appl. 9 (2023).

[54] H. Ji, W. Bi, J. Yan, S. Wu, H. Han, ICTs-driven Agriculture Contributes to the Mission of Carbon Reduction, in: Interdisciplinary Research in Technology and Management, 1st ed., CRC Press, London, 2024, pp. 453–462. https://doi.org/10.1201/9781003430469-53.

[55] A. Andrae, L. Hu, L. Liu, J. Spear, K. Rubel, Delivering Tangible Carbon Emission and Cost Reduction through the ICT Supply Chain, International Journal of Green Technology 3 (2017) 1–10. https://doi.org/10.30634/2414-2077.2017.03.1.

[56] B. Tomlinson, R.W. Black, D.J. Patterson, A.W. Torrance, The carbon emissions of writing and illustrating are lower for AI than for humans, Sci Rep 14 (2024) 3732. https://doi.org/10.1038/s41598-024-54271-x.

[57] C.A.M. Ajufo, G. Bekaroo, An Automated Personal Carbon Footprint Calculator for Estimating Carbon Emissions from Transportation Use, in: Proceedings of the International Conference on Artificial Intelligence and Its Applications, Association for Computing Machinery, New York, NY, USA, 2021, pp. 1–7. https://doi.org/10.1145/3487923.3487935.

[58] S. Wang, F. Tao, Y. Shi, Optimization of Inventory Routing Problem in Refined Oil Logistics with the Perspective of Carbon Tax, Energies 11 (2018) 1437. https://doi.org/10.3390/en11061437.

[59] O. Jabali, T.V. VanWoensel, A.G.D. Kok, Analysis of Travel Times and CO2 Emissions in Time-Dependent Vehicle Routing, Prod. Oper. Manag. 21 (2012) 1060-1074. https://journals.sagepub.com/doi/abs/10.1111/j.1937-5956.2012.01338.x (accessed 20.03.25).

[60] W.-H. Tsai, K.-C. Lee, J.-Y. Liu, H.-L. Lin, Y.-W. Chou, S.-J. Lin, A mixed activity-based costing decision model for green airline fleet planning under the constraints of the European Union Emissions Trading Scheme, Energy 39 (2012) 218–226. https://doi.org/10.1016/j.energy.2012.01.027.

[61] Booking, <https://www.booking.com/> (accessed 14.05.25).

[62] Travalyst, <https://travalyst.org/work/aviation-industry/> (accessed 12.05.25).

[63] Skyscanner, <https://www.skyscanner.com.tr/> (accessed 10.04.25).

[64] Google Flights, <https://www.google.com/travel/flights> (accessed 22.04.25).

*Sorumlu Yazar/Corresponding Author: fatihgencturk@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 133 – 149, 2025
DOI: 10.62301/usmtd.1720980

Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi
International Journal of Sustainable Engineering and Technology
ISSN: 2618-6055 / Cilt:9, Sayı:1, (2025), 133 – 149
Araştırma Makalesi – Research Article

BİLGİSAYAR BÖLÜMÜ ÖĞRENCİLERİ İÇİN MAKİNE ÖĞRENME ALGORİTMALARI İLE KARİYER ÖNERİSİ VE AÇIKLANABİLİR YAPAY ZEKA İLE DEĞERLENDİRİLMESİ

*Emine ALTUNBAŞ, Cevriye ALTINTAŞ[2]

[1] Isparta Uygulamalı Bilimler Üniversitesi, Teknoloji Fakültesi, Bilgisayar Mühendisliği Bölümü, Isparta

(Geliş/Received: 16.06.2025, Kabul/Accepted: 28.06.2025, Yayınlanma/Published: 30.06.2025)

ÖZ

Bu çalışmada, bilgisayar mühendisliği öğrencilerinin teknik becerilerine ve akademik performanslarına göre hangi mesleki alanda çalışabileceklerinin makine öğrenme algoritmaları ile tahmin edilmesi ve bu tahminlerin açıklanabilir yapay zekâ (AYZ) teknikleriyle değerlendirilmesi amaçlanmıştır. 174 öğrenciden oluşan özgün veri kümesinde, genel not ortalaması, programlama dili yeterlilikleri, proje bilgileri ve staj alanları yer almaktadır. AdaBoost, Decision Tree, Gradient Boosting, KNN, Logistic Regression, Naive Bayes, Random Forest, SVC algoritmaları LazyClassifier ile kıyaslanmış ve en başarılı Gradient Boosting ve Decision Tree modelleri hiperparametre optimizasyonu ile eğitilmiştir. En yüksek doğruluk Gradient Boosting algoritması ile %97 olarak elde edilmiştir. Modelin kararları SHAP ve LIME algoritmalarıyla yorumlanarak modelin açıklanabilirliği sağlanmıştır. Sonuçlar, önerilen yöntemin kariyer yönlendirme ve eğitim planlamasında kullanılabilir olduğunu göstermektedir.

Anahtar kelimeler: Sınıflandırma, LIME, SHAP, Makine Öğrenmesi, Gradyan Artırma

PREDICTION WITH MACHINE LEARNING ALGORITHMS FOR COMPUTER ENGINEERING DEPARTMENT AND EVALUATION WITH EXPLAINABLE ARTIFICIAL INTELLIGENCE

ABSTRACT

In this study, it is aimed to predict which professional field computer engineering students can work in according to their technical skills and academic performance using machine learning algorithms and to evaluate these predictions using explainable artificial intelligence (AI) techniques. The original dataset consisting of 174 students includes variables such as general grade point average, programming language proficiency, project information and internship areas. Eight different classification algorithms were compared with LazyClassifier and the most successful models were trained with hyperparameter optimization. The highest accuracy was obtained as 97% with the Gradient Boosting algorithm. The model's decisions were interpreted with SHAP and LIME algorithms and the explainability of the model was ensured. The results show that the proposed method can be used in career guidance and educational planning.

Keywords: Classification, LIME, SHAP, Machine learning, Gradient Boosting.

***Sorumlu Yazar/Corresponding Author:** yl2230696009@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 133 – 149, 2025
DOI: 10.62301/usmtd.1720980

1. Giriş (Introduction)

Bilgisayar mühendisliği öğrencilerinin kariyer planlaması, yapay zekâ ve veri bilimi alanlarında önemli bir araştırma konusu olarak karşımıza çıkmaktadır. Hem ulusal hem de uluslararası literatürde, kariyer tahmini ve yönlendirme amacıyla çeşitli yapay zekâ ve makine öğrenmesi algoritmalarının kullanıldığı çalışmalar mevcuttur. Ancak bu çalışmaların birçoğu genelleştirilmiş öğrenci gruplarına odaklanmakta olup bilgisayar mühendisliği öğrencilerine özel olarak uygulanmamış ya da AYZ algoritmaları ile yorumlanmasını kapsamamıştır. Bu bölümde yer alan kaynak özetleri önerilen çalışmanın literatürdeki önemini ve doldurulması gereken boşlukları analiz etmektedir.

Prasanna ve Haritha (2019), bilgisayar bilimi ve bilgi teknolojileri alanında öğrencilerin kariyer seçimlerine yardımcı olmak için makine öğrenmesi tabanlı bir akıllı kariyer rehberliği ve öneri sistemi geliştirmişlerdir. Öğrencilerin notlarına ve ilgi alanlarına dayalı olarak uygun alan seçimi için öngörüler sunmayı amaçlamışlardır. Çalışmada, öğrencilerin ilgi alanları ve yetenekleri hakkında bilgi toplamak için anket ve beceri testleri kullanılarak kapsamlı bir veri seti oluşturulmuştur. 1000 öğrenciden toplanan veriler üzerinde karar ağaçları, destek vektör makineleri, rastgele orman, Naive Bayes, lojistik regresyon ve doğrusal diskriminant analizi gibi çeşitli makine öğrenmesi algoritmalarını uygulanmıştır. Analiz sonuçlarına göre, Lojistik Regresyon ve Doğrusal Diskriminant Analizi yöntemleri sırasıyla %94 ve %82 doğruluk oranlarıyla en iyi performansı göstermiştir. Çalışma, geliştirilen sistemin kurum-öğrenci ilişkilerini güçlendirmede ve kurumun itibarını artırmada yardımcı olacağını ve gelecekte daha fazla veri ile derin sinir ağları ve zaman serisi analizi gibi ileri tekniklerin kullanılabileceğini önermektedir [5].

Akgün (2019), Hacettepe Üniversitesi Endüstri Mühendisliği Bölümü mezunlarının verilerini kullanarak, yeni mezun olacak öğrencilere iş bulma konusunda sektör önerisi yapabilecek bir model geliştirmiştir. Araştırmacı, 1969-2018 yılları arasındaki 235 mezunun verilerini kullanarak çeşitli makine öğrenmesi algoritmaları (Lojistik Regresyon, En Yakın Komşuluk, SVM, Gaussian Naive Bayes, Karar Ağacı ve Rassal Orman) ile bir tavsiye sistemi oluşturmuştur. Farklı tekrar sayıları (100, 200, 500 ve 1000) ile testler yapılmış ve sonuçlarda Rassal Orman algoritmasının %67.46 doğruluk oranıyla en iyi performansı gösterdiği, Precision değerinde SVM algoritmasının (%0.770), Recall ve F1 Score değerlerinde ise En Yakın Komşuluk algoritmasının (sırasıyla %0.660 ve %0.646) en yüksek sonuçları verdiği tespit edilmiştir. Ayrıca, mezun bilgilerinin artmasıyla tavsiye sisteminin doğruluk oranının yükselebileceğini belirtmiş ve gelecek çalışmalar için lisansüstü eğitim bilgileri, işe giriş şekli, yandal/çift anadal bilgileri gibi ek özelliklerin de modele dahil edilmesini, farklı program ve üniversitelerde de benzer sistemlerin uygulanmasını önermiştir [1].

Al-Dossari vd. (2020), Suudi Arabistan'daki BT mezunlarının kariyer yolu seçimlerine yardımcı olmak amacıyla bir makine öğrenimi tabanlı öneri sistemi olan CareerRec'i geliştirmişlerdir. Çalışma, BT sektöründeki 2255 çalışanın verileriyle eğitilen beş farklı makine öğrenimi algoritmasını karşılaştırmıştır. Araştırmada kullanılan algoritmalar arasında XGBoost, %70,47 doğruluk oranı ile en iyi performansı göstermiştir. CareerRec sistemi, BT mezunlarının becerilerine en uygun kariyer yollarını seçmelerine yardımcı olmayı amaçlamaktadır. Sonuçlar, XGBoost'un diğer makine öğrenimi modellerine göre daha etkili olduğunu ortaya koymuş, ancak sistemin performansının daha fazla veri ve derin öğrenme modelleriyle iyileştirilebileceği belirtilmiştir. Gelecekteki çalışmalarda, BT çalışanlarının becerilerini daha doğru bir şekilde ölçmek için alternatif yaklaşımlar kullanılması önerilmiştir [2].

Saeed vd. (2021), mühendislik öğrencilerine kariyer önerileri sağlamak için evrişimsel sinir ağı (ESA) tabanlı bir kariyer öneri sistemi geliştirmişlerdir. Çalışmada, çeşitli üniversitelerden toplanan bir kariyer veri seti kullanılarak, mühendislik öğrencilerinin nitelikleri, ilgi alanları ve becerilerine göre en uygun işlerin tahmin edilmesi amaçlanmıştır. Kariyer önerileri, derin doğal dil işleme tabanlı ESA modeli ile gerçekleştirilmiş ve performansı artırmak için 512 gizli katman kullanılmıştır. Çalışma, öneri sisteminin içerik tabanlı filtreleme teknikleri ile karşılaştırıldığında %84 doğrulukla daha iyi sonuçlar verdiğini göstermiştir. Bu öneri sistemi, mühendislik disiplinleri arasında iş arama sürecinde maliyet ve zamandan tasarruf sağlarken, öğrencilere en uygun iş fırsatlarını sunmuştur. Çalışmada gelecekteki araştırmalar için daha geniş veri kümeleri ve diğer eğitim alanlarına genişletilmesi gerektiğini vurgulanmakta, bu modelin genelleştirilebilirliği için farklı veri kümeleri üzerinde testler önerilmektedir [7].

*Sorumlu Yazar/Corresponding Author: yl2230696009@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 133 – 149, 2025
DOI: 10.62301/usmtd.1720980

Çelik (2022), makine öğrenmesi yöntemlerinin eğitim alanında kullanımına odaklanmıştır. Lise öğrencilerinin dört yıl boyunca derslerde aldıkları notlar ve diğer bilgiler ile seçtikleri üniversite bölümleri arasındaki ilişkiyi incelemiştir. Mezun olmuş 447 öğrencinin verileri ve bu öğrencilerin tercih ettiği 43 farklı bölüm belirlenmiştir. Veri setindeki gözlem birimlerinin bağımlı değişkendeki sınıf sayısına oranının düşük olması nedeniyle, benzer bölümler birleştirilerek 43, 23, 9 ve 6 sınıf içeren dört farklı etiketleme yaklaşımı oluşturulmuştur. Her bir etiketleme yaklaşımı için 10 farklı makine öğrenmesi algoritması test edilmiş ve sonuçlara göre XGBoost algoritması ilk yaklaşımda %43, ikinci yaklaşımda %53, dördüncü yaklaşımda %75 başarı oranıyla en iyi performansı gösterirken, üçüncü etiketleme yaklaşımında KNN algoritması %65 başarı oranıyla öne çıkmıştır. Araştırmacı, veri setindeki gözlem birimi sayısının önemli ölçüde artırılması durumunda, bölümlerin gruplandırılmasına gerek kalmadan daha başarılı tahminler yapılabileceğini vurgulamıştır [3].

Guleria ve Sood (2023), makine öğrenmesi ve AYZ tekniklerini kullanarak kariyer danışmanlığı için bir çerçeve geliştirmeye odaklanmıştır. Araştırmada, Beyaz Kutu ve Siyah Kutu modelleri, öğrencilerin akademik başarıları ve istihdam edilebilirlik niteliklerini analiz etmek için kullanılmıştır. Yazarlar, Naive Bayes, Lojistik Regresyon, Karar Ağacı, SVM, KNN ve Topluluk modelleri gibi çeşitli makine öğrenme algoritmalarını eğiterek, Naive Bayes modelinin Geri Çağırma (%91,2) ve F-Ölçümü (%90,7) açısından en iyi sonuçları verdiğini tespit etmişlerdir. Çalışma, kariyer danışmanlığında kullanılabilecek etkili bir model önerirken, kullanılan veri kümesinin sınırlı nitelik kapsamı ve örneklem büyüklüğü gibi sınırlamaları vurgulamıştır. Sonuç olarak, Guleria ve Sood, sosyo-ekonomik faktörler ve genişletilmiş veri kümeleri ile daha kapsamlı çalışmalar yapmayı önermiş, bu kapsamda derin öğrenme ve gözetimsiz öğrenme teknikleriyle kariyer danışmanlığı modellerini geliştirmeyi planlamışlardır [4].

Nizar vd. (2024), bilgisayar bilimleri öğrencileri arasında yazılım beceri setlerini kullanarak tahmin edilmesi ve değerlendirilmesi için evrişimsel sinir ağları ve açıklanabilir yapay zekâ yöntemlerini kullanmıştır. Hindistan'daki çeşitli kurumlardan bilgisayar bilimi öğrencilerden yaşam ve teknik becerilerine ilişkin 2004 öğrencinin anket verisi toplanmıştır. Önerilen modelin performansı, accuracy, precision, recall, F1 score ve AUC değeri gibi ölçütleri kullanılarak değerlendirilmiştir. ANN, LightGBM, AdaBost, XGBBoost ve önerilen CNN modelleri için performans değerlendirme ölçütleri kullanılmıştır. Modelin tahminlerinin yorumlanması içinde SHAP, LIME ve LRP açıklanabilir yapay zekâ teknikleri kullanılmıştır. Simülasyon sonuçları, önerilen CNN-XAI modelinin tahmin doğruluğu açısından diğer temel modellerden daha iyi performans gösterdiğini göstermektedir [6].

2. Materyal ve Metot (Material and Method)

2.1.Makine Öğrenmesi (Machine Learning)

Makine öğrenmesi günümüzde birçok disiplinin kesişim noktasında yer alan ve veri analizi ile modelleme süreçlerini otomatikleştiren bir alan olarak öne çıkmaktadır. Bu teknoloji özellikle büyük veri setlerinin işlenmesi ve analiz edilmesi gereken durumlarda önemli bir rol oynamaktadır. Makine öğrenmesi istatistiksel yöntemlerin ve algoritmaların bir kombinasyonu olarak, verilerden öğrenme ve bu öğrenmeyi yeni verilere uygulama yeteneği sunar. Bu nedenle, makine öğrenmesinin çeşitli uygulama alanları ve yöntemleri, disiplinler arası bir yaklaşım gerektirmektedir [8].

Makine öğrenmesi algoritmaları birçok sektörde olduğu gibi eğitim sektöründe de kullanımı giderek artmaktadır. Bu çalışmada bilgisayar mühendisliği öğrencilerinin kariyer tahmini için rastgele orman sınıflandırıcısı, destek vektör makinesi, K en yakın komşu, karar ağaçları, gradient boosting, adaboost ve gaussian navie bayes algoritmaları olmak üzere sekiz farklı algoritma seçilmiştir. Bu bölümde seçilen sekiz algoritma kısaca incelenmiştir.

2.2.Veri Kümesinin Hazırlanması (Preparation of Dataset)

Bilgisayar mühendisliği öğrencilerinin akademik öğrenme süreçlerinden elde edilen bilgiler ile gelecekte hangi kariyer alanında çalışabileceklerinin makine öğrenmesi ile tahmin edilmesi ve sonuçların AYZ algoritmaları ile yorumlaması için bilgisayar mühendisliği ve yazılım mühendisliği bölümünden mezun olmuş ve özel sektörde çalışan kişiler üzerinden yapılan görüşme, form doldurma ve linkedin profilleri üzerinden kazıma ile çalışmaya özgün bir veri kümesi

*Sorumlu Yazar/Corresponding Author: yl2230696009@isparta.edu.tr

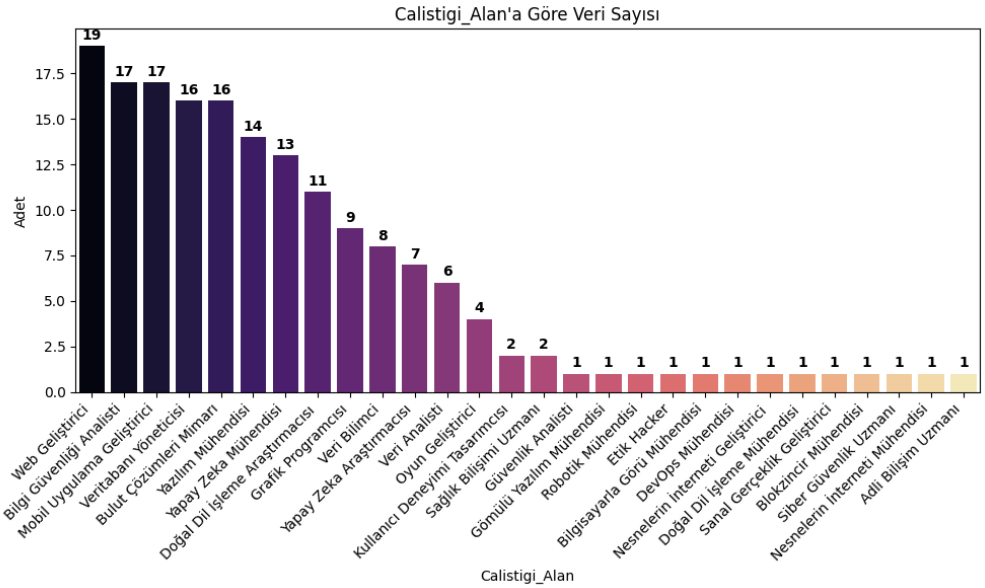
ISSN: 2618-6055 / 9, (1), 133 – 149, 2025
DOI: 10.62301/usmtd.1720980

oluşturulmuştur. Veri kümesi özellikleri; genel not ortalaması (GNO), Staj alanı (ilgili_alan), geliştirdiği en önemli proje (proje), şu anda çalıştığı alan (calistiği_alan), veri tabanı bilgisi, Python, Java, Csharp, C++ programlama dillerinden oluşmaktadır. Veri kümesi içerisinde bildiği programlama dilleri ve veri tabanı bilgisi çok iyi, ortalama ve zayıf olarak ölçeklendirilmiştir (Şekil 1). Aynı şekilde genel not ortalaması da 2 ile 2.5 arasında olanlar ortalama, 2.5 ile 3 arasında olanlar iyi, 3 ile 3.5 arasında olanlar çok iyi ve 3.5 ile 4 arasında olanlar mükemmel olarak ölçeklendirilmiştir. Veri kümesi toplamda 174 kayıttan oluşmaktadır.

GNO	Ilgili_Alan	Proje	Calistigi_Alan	Veritabani	Python	Java	Csharp	C++
Ortalama	Yapay Zeka	Chatbot Geliştirme	Yapay Zeka Araştırmacısı	Çok İyi	Çok İyi	Zayıf	Zayıf	Zayıf
İyi	Veri Bilimi	Veri Analitiği	Veri Bilimci	Çok İyi	Ortalama	Zayıf	Zayıf	Zayıf
Ortalama	Yazılım Geliştirme	E-ticaret Web Sitesi	Yazılım Mühendisi	Çok İyi	Zayıf	Ortalama	Çok İyi	Zayıf
Ortalama	Web Geliştirme	Full-Stack Web Uygulaması	Web Geliştirici	Çok İyi	Zayıf	Çok İyi	Zayıf	Zayıf
Ortalama	Siber Güvenlik	Siber Güvenlik	Bilgi Güvenliği Analisti	Zayıf	Ortalama	Çok İyi	Zayıf	Zayıf
İyi	Yapay Zeka	Görüntü Tanıma	Yapay Zeka Mühendisi	Ortalama	Çok İyi	Zayıf	Zayıf	Zayıf
Çok İyi	Veritabanı Yönetimi	SQL Sorgu Optimizasyonu	Veritabanı Yöneticisi	Çok İyi	Ortalama	Zayıf	Zayıf	Zayıf
İyi	Bulut Bilişim	AWS Dağıtımı	Bulut Çözümleri Mimarı	Çok İyi	Zayıf	Ortalama	Zayıf	Zayıf
Ortalama	Mobil Uygulama Geliştirme	Android Uygulaması	Mobil Uygulama Geliştirici	Zayıf	Ortalama	Çok İyi	Zayıf	Zayıf
Ortalama	3D Modelleme	3D İşleme	Grafik Programcısı	Zayıf	Zayıf	Çok İyi	Zayıf	Zayıf
İyi	Yapay Zeka	Doğal Dil İşleme	Doğal Dil İşleme Araştırmacısı	Ortalama	Çok İyi	Zayıf	Zayıf	Zayıf
Ortalama	Web Geliştirme	Full-Stack Web Uygulaması	Web Geliştirici	Çok İyi	Zayıf	Ortalama	Zayıf	Zayıf

Şekil 1. Veri kümesinin ham hali (Raw dataset)

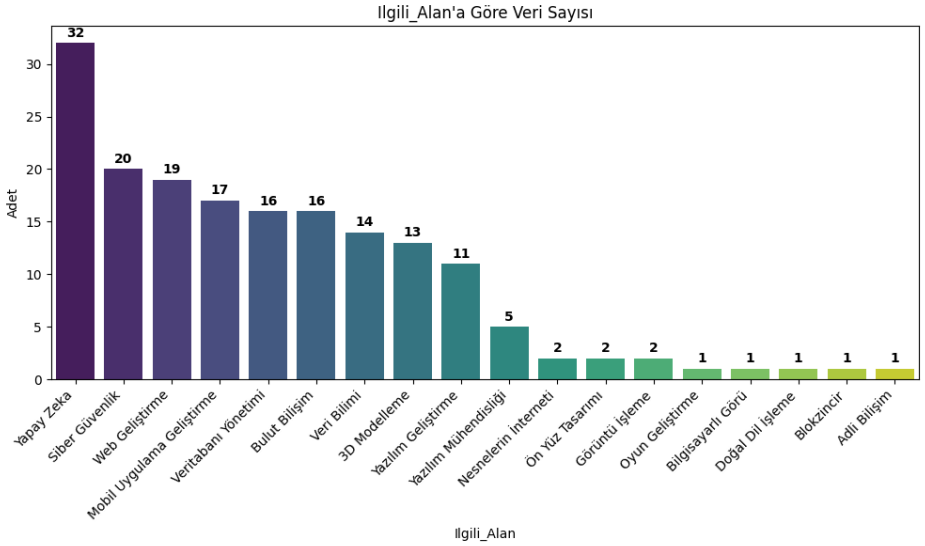
Veri kümesi içerinde veri tabanı bilgisi, Python, Java, Csharp, C++ ana programlama dilleri bilgisine göre sektörde karşılıklarını görebilmek adına Şekil 2, Şekil 3 ve Şekil 4'te çalıştığı alana, ilgilendiği alan ve genel not ortalamasına göre gruplandırılarak sayıları verilmiştir. Elde edilen gruplandırma gösteriyor ki bilgisayar mühendisliğinde güncel ve popüler her alandan veri yeterli sayıda yer almaktadır.



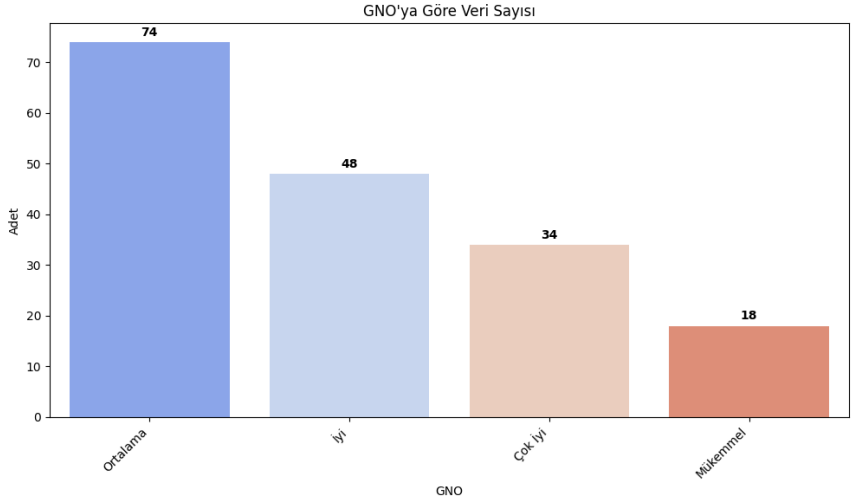
Şekil 2. Veri kümesinde yer alan verilerin çalıştığı alana göre sayı dağılımı (Number distribution of data in the dataset according to the field it operates in)

*Sorumlu Yazar/Corresponding Author: yl2230696009@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 133 – 149, 2025
DOI: 10.62301/usmtd.1720980



Şekil 3. Veri kümesinde yer alan staj ve işyeri eğitimi alanına göre sayı dağılımı (Distribution of the number of internships and workplace training in the dataset according to field)



Şekil 4. Veri kümesinde yer alan verilerin genel not ortalamasına göre sayı dağılımı (Number distribution of data in the dataset according to general grade point average)

2.3.Ön İşleme ve Kodlama (Preprocessing and Coding)

Veri kümesinde yer alan verilerin makine öğrenmesi algoritmaları tarafından kullanılması için kategorik verilerin sayısallaştırılması gerekmektedir. Veri kümesinde yer alan "ilgili alan", "proje" ve "çalıştığı alan" gibi nominal kategorik değişkenler makine öğrenmesi algoritmaları tarafından doğrudan işlenemez. Bu nedenle bu tür değişkenler, sınıf sayısının az olması durumunda etiketleme daha fazla kategoriye sahip durumlarda ise tekil sütunlara ayrılarak one-hot encoding yöntemi ile sayısal forma dönüştürülmüştür. Ön işlemelerde pandas kütüphanesinde yer alan ve kategorik değişkenleri sayısal verilere dönüştürmek için kullanılan bir ön işleme yöntemi olan get_dummies fonksiyonu kullanılmıştır.

Programlama dillerine (Python, Java, C#, C++) ve veri tabanı bilgisine ait bilgi düzeyleri "Zayıf", "Ortalama" ve "Çok İyi" şeklinde sıralı kategorilerden oluşmaktadır. Bu nedenle sıralı değişkenler olarak kabul edilerek ve sayısal karşılıkları sırasıyla 0, 1 ve 2 olacak şekilde dönüştürülmüştür. Benzer biçimde öğrencilerin genel not ortalamaları da önceden belirlenmiş

*Sorumlu Yazar/Corresponding Author: yl2230696009@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 133 – 149, 2025
DOI: 10.62301/usmtd.1720980

aralıklar kullanılarak "Ortalama" (2.0–2.5), "İyi" (2.5–3.0), "Çok İyi" (3.0–3.5) ve "Mükemmel" (3.5–4.0) şeklinde kategorize edilmiştir. Bu kategoriler de ordinal ölçeklendirme yöntemiyle 0, 1, 2 ve 3 şeklinde sayısallaştırılmıştır. Bilgisayar mühendisliği alanlar veri kümesi içerisinde karışmasın diye ilgili alan özelliğine "InDo" eki, çalıştığı alan özelliğine ise "PR" eki eklenmiştir.

2.4.Tahmin modelinin geliştirilmesi (Development of the prediction model)

Bu çalışmada mühendislik öğrencilerinin sahip oldukları akademik ve teknik yetkinliklere bağlı olarak hangi alanda çalışabileceklerinin öngörülmesi amacıyla bir makine öğrenmesi temelli tahmin modeli geliştirilmiştir. Modelleme sürecinde ilk olarak ön işleme adımlarından geçirilmiş olan veri kümesi eğitim (%80) ve test (%20) olmak üzere iki alt kümeye ayrılmıştır. Eğitim kümesi, algoritmaların öğrenme süreçlerinde kullanılmıştır. Test kümesi ise geliştirilen modelin genelleme performansını değerlendirmek amacıyla ayrılmıştır.
Bu bölümde sınıflandırma problemine en uygun makine öğrenmesi algoritmasının belirlenebilmesi amacıyla Lazy Predict kütüphanesi içerisinde yer alan LazyClassifier yapısı kullanılmıştır. LazyClassifier, farklı sınıflandırma algoritmalarını aynı veri kümesi üzerinde otomatik olarak eğitip test ederek karşılaştırmalı performans analizini gerçekleştiren bir araçtır. Bu yapı model geliştirme sürecinde deneme yanılma yöntemlerini en aza indirerek, Tablo 1' de verilen birçok popüler algoritmanın varsayılan hiperparametrelerle elde ettiği doğruluk, F1 skoru, işlem süresi gibi metrikleri tek bir tablo üzerinde sunmaktadır. Böylece belirli bir veri kümesine en iyi uyum sağlayan algoritmalar hızlı ve objektif biçimde tespit edilebilmektedir. Bu sebeplerden dolayı ilgili veri kümesine özgü sınıflandırma başarısını en üst düzeye çıkarabilecek modelin seçimi LazyClassifier çıktıları dikkate alınarak yapılmıştır. LazyClassifier sonuçlarından en yüksek doğruluğu veren modeller AdaBoost, Decision Tree, Gradient Boosting, KNN, Logistic Regression, Naive Bayes, Random Forest, SVC olarak belirlenmiştir.

Tablo 1. LazyClassifier Kütüphanesindeki Modeller (Models in the LazyClassifier Library)

Modeller		
BaggingClassifier	LogisticRegression	SVC
LinearDiscriminantAnalysis	DecisionTreeClassifier	LGBMClassifier
GaussianNB	ExtraTreesClassifier	SGDClassifier
AdaBoostClassifier	BernoulliNB	CalibratedClassifierCV
Perceptron	RidgeClassifier	LabelPropagation
KNNClassifier	Random Forest	DummyClassifier
Gradient Boosting	Naive Bayes	LinearSVC

Elde edilen sonuçlara göre en yüksek sınıflandırma performansını gösteren modellerin hiperparametreleri GridSearchCV yöntemiyle optimize edilerek nihai tahmin modeli oluşturulmuştur.

Tablo 2'de her bir model için GridSearchCV yöntemiyle gerçekleştirilen hiperparametre arama sürecinde kullanılan parametre aralıkları ile elde edilen en iyi sonuçlar gösterilmiştir. Bu süreç, modelin doğruluk performansını en üst düzeye çıkaracak konfigürasyonun belirlenmesini sağlamış ve nihai tahmin modelleri bu değerler üzerinden yapılandırılmıştır.

*Sorumlu Yazar/Corresponding Author: yl2230696009@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 133 – 149, 2025
DOI: 10.62301/usmtd.1720980

Tablo 2. GridSearchCV yöntemi ile modellerin hiperparametrelerinin aralıkları ve optimum değerleri (Ranges and optimum values of hyperparameters of models with GridSearchCV method)

Model Adı	Hiperparametre	Aralıklar	Değer
Gradient Boosting	n_estimators	[100, 150]	150
	learning_rate	[0.05, 0.1]	0.1
	max_depth	[3, 5]	5
AdaBoost	n_estimators	[50, 100]	100
	learning_rate	[0.5, 1.0]	1.0
Decision Tree	criterion	['gini', 'entropy']	entropy
	max_depth	[None, 5, 10]	10
	min_samples_split	[2, 5]	2
Logistic Regression	penalty	['l2']	l2
	C	[0.1, 1.0, 10]	1.0
	solver	['lbfgs']	lbfgs
Naive Bayes	var_smoothing	[1e-9, 1e-8, 1e-7]	1e-9
SVC	C	[0.1, 1, 10]	1
	kernel	['linear', 'rbf']	rbf
	gamma	['scale', 'auto']	scale
KNN	n_neighbors	[3, 5, 7]	5
	weights	['uniform', 'distance']	distance
	metric	['euclidean', 'manhattan']	euclidean
Random Forest	n_estimators	[100, 200]	200
	max_depth	[None, 10, 20]	20
	min_samples_split	[2, 5]	2
	max_features	['auto', 'sqrt']	sqrt

3. Araştırma Bulguları (Research Findings)

3.1. Sınıflandırma Modellerin Sonuçlarının Değerlendirilmesi (Evaluation of the Results of Classification Models)

Makine öğrenmesi modellerinin değerlendirilmesinde doğruluk, kesinlik, duyarlılık ve F1 skoru tercih edilmiştir ve her metrik değerlendirmede farklı bakış açılarıyla model performansını yansıtmaktadır. Ancak bu metrikler içerisinde özellikle F1 skoru modelin genel başarısını daha hassas ve dengeli bir şekilde yansıtan bir ölçüt olarak öne çıkmaktadır. F1 skoru, kesinlik ve duyarlılık değerlerinin harmonik ortalamasıdır. Bundan dolayı hem yanlış pozitif hem de yanlış negatif tahminlerin etkisini dengeleyecek biçimde hesaplanır. Özellikle veri setinde sınıf dengesizliği mevcutsa yalnızca doğruluk oranı yüksek olsa bile modelin belirli sınıflarda başarısız olması mümkündür. Bu gibi durumlarda yüksek doğruluk değeri yanıltıcı olabilecektir. Örneğin tüm sınıfların eşit dağılmadığı bir veri kümesinde model çoğunluk sınıfı tahmin ederek yüksek doğruluk sağlayabilir ancak azınlık sınıfı neredeyse hiç tanınmayabilir. Bu nedenle F1 skoru, modelin hem pozitif sınıfı doğru tanıma yeteneğini (duyarlılık) hem de tahminlerinin doğruluğunu (kesinlik) birlikte değerlendirmesi açısından daha kapsamlı ve hassas bir gösterge olarak kabul edilir. Bu sebeplerden dolayı bu çalışmada geliştirilen sınıflandırma modellerinin değerlendirilmesinde F1 skoru temel referans metrik olarak ele alınmış ve model karşılaştırmaları bu metrik üzerinden ağırlıklı olarak gerçekleştirilmiştir.

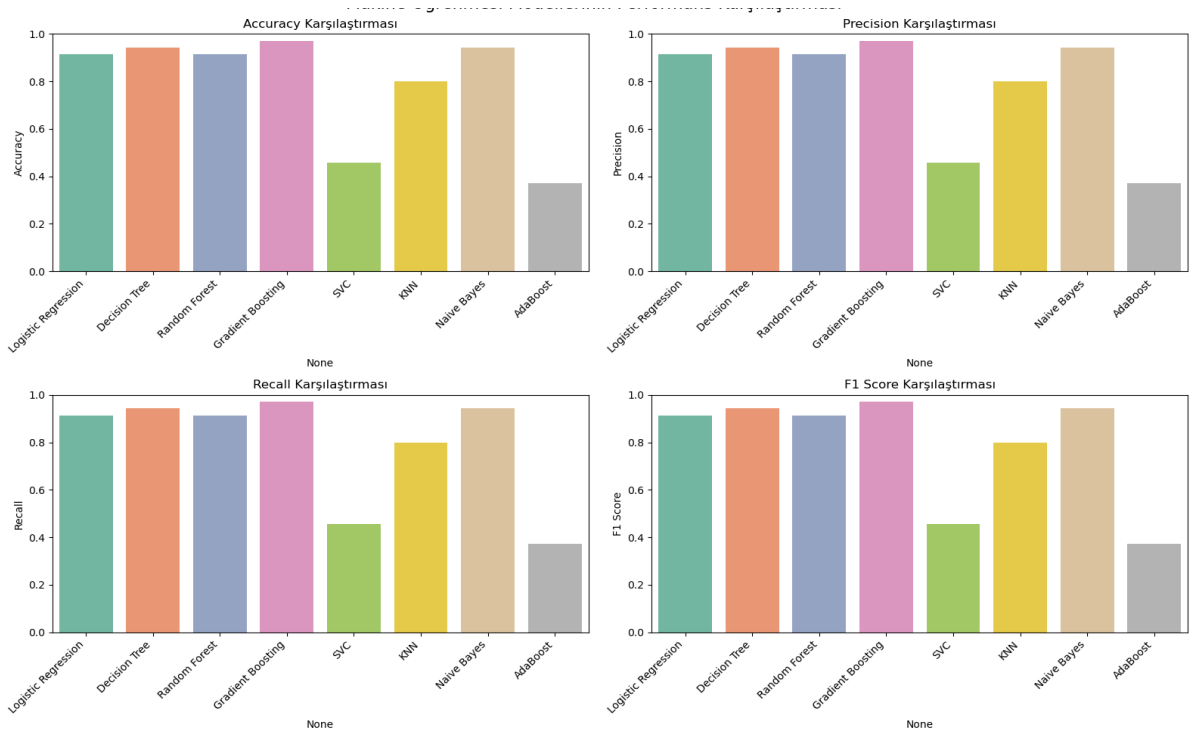
Bilgisayar mühendisliği öğrencilerine ait veri kümesinden elde edilen sınıflandırma modellerinin doğruluklarını Tablo 3'te verilmiştir. Elde edilen sonuçlara göre tüm performans metrikleri dikkate alındığında en başarılı modelin Gradient Boosting olduğu görülmektedir. Gradient Boosting ana modelimiz %97.5 doğruluk, kesinlik, duyarlılık ve F1 skoru değerleri ile diğer tüm modellerden daha iyi doğruluk göstermiştir. F1 skoru değerleri üzerinden en başarılı model ikinci ve üçüncü ise %95.4 ve %94.2 oranlarıyla Decision Tree ve Naive Bayes algoritmaları olmuştur. Logistic Regression %91.4 başarım oranlarıyla bu iki modeli takip etmiştir. Random Forest ve K-Nearest Neighbors (KNN) modelleri sırasıyla %88.5 ve %80 doğruluk oranı ile orta düzeyde başarı sağlamıştır. Buna karşın AdaBoost (%48.6) ve Support Vector Classifier (SVC) (%45.7) modelleri tüm metriklerde düşük performans göstermiştir. Bu sonuçlar doğrultusunda hiperparametre optimizasyonuna rağmen bazı algoritmaların veri kümesinin yapısına uygunluk göstermediğini ortaya koymaktadır.

*Sorumlu Yazar/Corresponding Author: yl2230696009@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 133 – 149, 2025
DOI: 10.62301/usmtd.1720980

Tablo 3. Modellerin sonuçlarının karşılaştırılması (Comparison of results of models)

Model	Doğruluk	Kesinlik	Duyarlılık	F1 Skor
Gradient Boosting	0.9714	0.9704	0.9714	0.9754
Decision Tree	0.9521	0.9528	0.9523	0.9541
Naive Bayes	0.9428	0.9428	0.9428	0.9428
Logistic Regression	0.9142	0.9142	0.9142	0.9142
Random Forest	0.8857	0.8857	0.8857	0.8857
KNN	0.8000	0.8001	0.8020	0.8030
AdaBoost	0.4857	0.4857	0.4858	0.4857
SVC	0.4571	0.4560	0.4571	0.4571

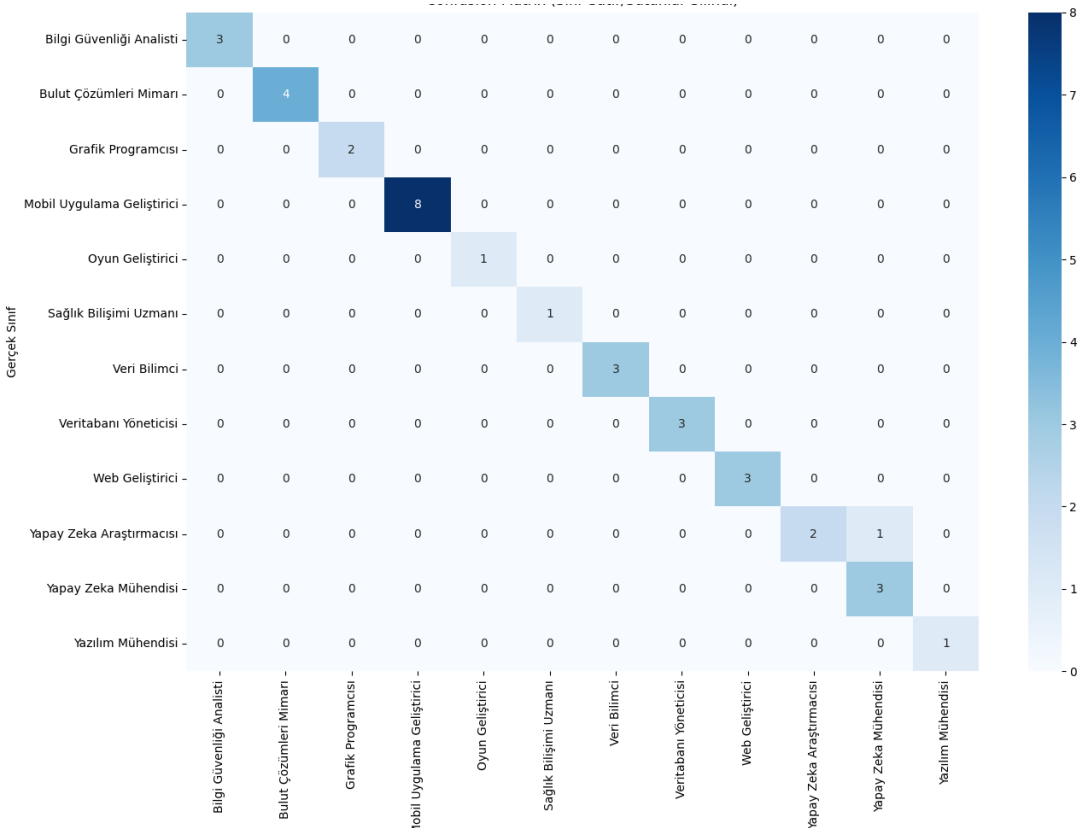


Şekil 5. Modellerin sonuçlarının karşılaştırılması (Comparison of results of models)

Sekiz farklı model içerisinde en yüksek doğruluğu elde ettiğimiz Gradient Boosting modeli üzerinden sınıflandırma sonuçlarının özellikleri açısından değerlendirilmesi yapılmıştır. Bu çalışmada sınıflandırma sürecinde en yüksek başarıyı gösteren model olan Gradient Boosting algoritmasına ait karışıklık matrisi Şekil 6'da gösterilmektedir. Karışıklık matrisi modelin her bir sınıfa ait örnekleri doğru ve yanlış sınıflandırma durumlarını ayrıntılı olarak göstermektedir.

*Sorumlu Yazar/Corresponding Author: yl2230696009@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 133 – 149, 2025
DOI: 10.62301/usmtd.1720980



Şekil 5.Sınıflandırma modelinin karışıklık matrisi (Confusion matrix of the classification model)

Çalışmada kullanılan Destek Vektör Makinesi (SVC) modeli, %45.7 doğruluk oranı ile en düşük performans gösteren algoritma olmuştur. Bu bölümde, SVC'nin neden bu veri kümesinde başarısız olduğu detaylı olarak incelenmiştir.

SVC, özellikle doğrusal çekirdek (linear kernel) kullanıldığında, veri kümesindeki karmaşık ve doğrusal olmayan ilişkileri modellemekte zorlanmıştır.

Kategorik verilerin one-hot encoding ile sayısallaştırılması sonucunda oluşan yüksek boyutlu ve seyrek özellik uzayı, SVC'nin sınırlı genelleme yeteneğini daha da olumsuz etkilemiştir.

GridSearchCV ile hiperparametre optimizasyonu yapılmasına rağmen (C=1, kernel='rbf', gamma='scale'), modelin performansı diğer algoritmalara kıyasla düşük kalmıştır. Bu durum, veri dağılımının SVC'nin varsayımlarına uygun olmadığını göstermektedir.

Veri kümesindeki bazı meslek sınıflarının örnek sayısının az olması (örneğin, "Web Geliştirici" ve "Bilgi Güvenliği Analisti" sınıflarının 3'er örnekle temsil edilmesi), SVC'nin sınıf ağırlıklarını (class weights) doğru öğrenememesine neden olmuştur.

Dengesiz dağılım, modelin marjinal sınıfları göz ardı ederek çoğunluk sınıflara yönelik önyargılı tahminler yapmasına yol açmıştır.

SVC, kernel seçimi ve regularization parametresi (C) gibi hiperparametrelere son derece duyarlıdır.

3.2.Sonuçların Açıklanabilir Yapay Zekâ Algoritmaları ile Değerlendirilmesi (Evaluation of Results with Explainable Artificial Intelligence Algorithms)

Bu çalışmada ana model olan Gradient Boosting algoritmasının sınıflandırma kararlarının yorumlanabilirliğini artırmak amacıyla AYZ yöntemlerinden SHAP ve LIME algoritmalarından yararlanılmıştır. Bu algoritmalar modelin kararlarını hangi özelliklerin ne ölçüde etkilediğini sayısal ve görsel olarak ortaya koyarak, modelin karar verme mekanizmasını daha şeffaf hale getirmektedir. SHAP algoritması daha yaygın olarak tüm veri kümesi genelinde her bir özelliğin model çıktısına katkısını

*Sorumlu Yazar/Corresponding Author: yl2230696009@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 133 – 149, 2025
DOI: 10.62301/usmtd.1720980

hesaplayarak genel yorumlama sağlamaktadır. LIME algoritması ise belirli bir alan için en etkili olan özellikleri lokal düzeyde analiz ederek modelin o örneğe özel kararını açıklamaktadır. Bu sayede modelin yüksek doğruluk oranlarına ulaşmasının ötesinde, tahmin ettiği sınıflandırma sonuçlarının altında yatan nedenler de anlaşılır kılınmıştır.

3.2.1. Sonuçların Local Interpretable Model-agnostic Explanations Tarafından Yorumlanması (Interpretation of Results by LIME)

Gradient Boosting modeli tarafından yapılan sınıflandırmaların açıklanabilirliğini sağlamak amacıyla uygulanan LIME analizi her bir alana ait örnekler özelinde yorumlayıcı bilgiler Tablo 4'te verilmiştir. LIME algoritması ile veri kümesinde yer alan veri tabanı ve programlama dillerini baz alarak bunların sınıflandırmaya etkisi incelenmiştir. Her bir görseldeki yeşil çubuklar modelin tahminini destekleyen, kırmızı çubuklar ise tahmine karşıt yönde etkide bulunan özellikleri temsil etmektedir.

Tablo 4'teki LIME analizleri incelendiğinde, Veri Bilimci sınıfında özellikle orta düzeyde Python bilgisi ve iyi seviyede Java bilgisi pozitif etkiye sahipken, Mobil Uygulama Geliştirici örneğinde Python bilgisi pozitif, düşük Java ve ileri düzey veritabanı bilgisi negatif yönde katkı sağlamıştır. Bilgi Güvenliği Analisti sınıfında orta düzey Python bilgisinin tahmini desteklediği ancak ileri veritabanı bilgisinin bu sınıfa ait tahmini olumsuz etkilediği gözlenmiştir. Ayrıca Yapay Zekâ Araştırmacısı ve Yapay Zekâ Mühendisi sınıflarında Python ve Java bilgisi pozitif etkiler üretmiştir. Buna karşın veritabanı bilgisinin yüksekliği genellikle bu tahminleri zayıflatmıştır.

Bu analizler sonucunda LIME algoritmasının sunduğu lokal açıklamalar sayesinde modelin hangi bireysel özellikleri hangi sınıflandırma kararlarında nasıl kullandığı açık biçimde gözlemlenebilmiştir. Böylece hem modelin güvenilirliği artırılmış hem de tahmin sonuçlarının kullanıcıya yorumlanabilir biçimde sunulması mümkün kılınmıştır. Özellikle karar sınırında kalan örneklerde pozitif-negatif etkilerin dengesi, modelin doğruluğunun ötesinde karar gerekçesinin şeffaflaşmasına katkı sağlamıştır.

Yapılan LIME analizlerinde, Python, Java ve Veritabanı gibi bazı teknik yetkinliklerin model kararlarını güçlü biçimde etkilediği gözlemlenmiştir. Ancak aynı analizlerde C++ ve C# programlama dillerinin model kararlarına anlamlı bir etkisi olmadığı tespit edilmiştir. Bu durumun başlıca iki nedeni bulunmaktadır:

Modelin Değişken Seçimi (Feature Importance) Kararı: Gradient Boosting algoritması, tahmin sürecinde yalnızca sınıflar arasındaki farkları en iyi açıklayan değişkenleri kullanmaktadır. Eğitim süreci sonucunda model, C++ ve C# değişkenlerinin hedef değişken olan "Çalıştığı Alan" üzerinde anlamlı bir ayrım gücüne sahip olmadığını öğrenmiş ve bu özelliklere çok düşük önem derecesi (feature importance) atamıştır. LIME algoritması, yalnızca modelin karar sürecine etki eden değişkenleri açıklamaya çalıştığı için bu değişkenler çoğu örnekte yer almamıştır.

Verideki Varyans Eksikliği: C++ ve C# değişkenlerinin örneklem içinde yeterince çeşitlilik göstermemesi, bu değişkenlerin karar süreçlerine katkısını sınırlamıştır. Örneğin, bu programlama dillerine ait yetkinlik düzeylerinin çoğu öğrenci için benzer (örneğin "Ortalama") düzeyde olması, modelin bu değişkenler üzerinden anlamlı sınıf ayrımı yapamamasına neden olmuştur.

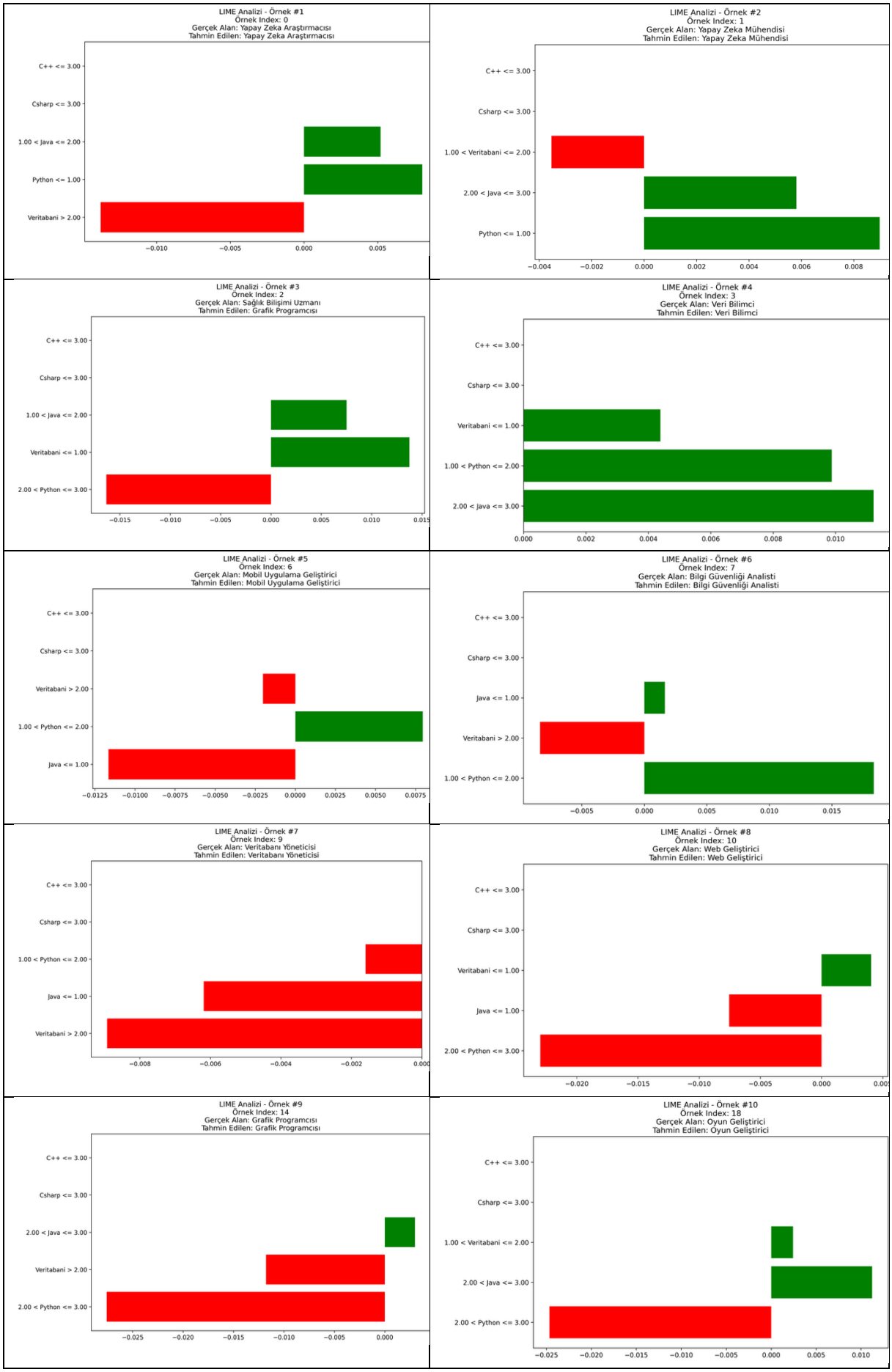
Bu nedenle, model hem istatistiksel açıdan hem de karar ağacı düzeyinde bu değişkenleri göz ardı etmiş ve LIME analizleri sonucunda bu özelliklerin etkisi çoğu grafikte görülmemiştir. Modelin daha yüksek ayrım gücüne ulaşabilmesi için bu tür değişkenlerin daha dengeli, çeşitli ve anlamlı biçimde veri kümesine yansıtılması gerekmektedir.

Sonuç olarak, LIME analizleri sadece hangi değişkenlerin kullanıldığını değil, aynı zamanda modelin karar verirken bu değişkenlere ne kadar itibar ettiğini de ortaya koyarak açıklanabilirliğe katkı sağlamaktadır.

*Sorumlu Yazar/Corresponding Author: yl2230696009@isparta.edu.tr

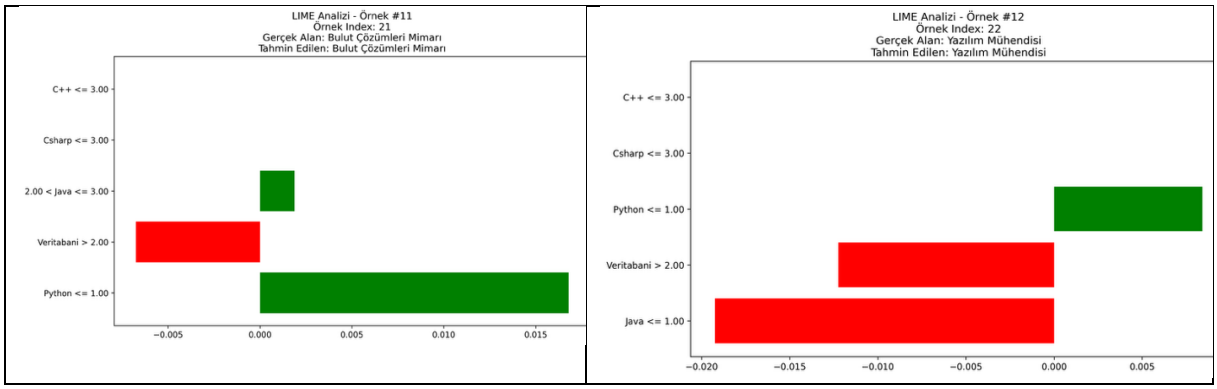
ISSN: 2618-6055 / 9, (1), 133 – 149, 2025
DOI: 10.62301/usmtd.1720980

Tablo 4. Her bir sınıf için LIME sonuçları (LIME results for each class)



*Sorumlu Yazar/Corresponding Author: yl2230696009@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 133 – 149, 2025
DOI: 10.62301/usmtd.1720980



3.2.2. Sonuçların SHapley Additive exPlanations Tarafından Yorumlanması (Interpretation of Results by SHAP)

Bu çalışmada, geliştirilen tahmin modelinin karar mekanizmasını daha iyi anlamak amacıyla uygulanan SHAP analizi ile her bir alan özelinde özniteliklerin model çıktısına etkisi detaylı biçimde Tablo 5 ‘de verilmiştir.

Veri Bilimci sınıfı için yapılan SHAP analizi, özellikle PR_Yapay Zekâ, InDo_Bulut Bilişim, InDo_Veritabanı Yönetimi ve InDo_Web Geliştirme gibi teknik alanlara ait yeterliliklerin yüksek SHAP değerlerine sahip olduğunu ve bu sınıfa tahmini güçlü biçimde etkilediğini göstermektedir.

Benzer şekilde Mobil Uygulama Geliştirici sınıfı için en belirgin öznitelikler InDo_Mobil Uygulama Geliştirme, InDo_Bulut Bilişim ve InDo_Web Geliştirme olarak öne çıkmış, bu da geliştiricilerin çalıştıkları teknoloji yığınlarının tahmine yön verdiğini göstermektedir.

Bilgi Güvenliği Analisti sınıfında InDo_Siber Güvenlik, PR_Siber Güvenlik ve InDo_Bulut Bilişim öznitelikleri en yüksek katkıyı sağlayan unsurlar olurken, Python ve Veritabanı gibi programlama bilgileri de sınırlı da olsa etki göstermiştir.

Yapay Zekâ Araştırmacısı sınıfında ise InDo_Yapay Zekâ, PR_Sinir Ağı Geliştirme, PR_Takviyeli Öğrenme gibi ileri düzey yapay zekâ tekniklerinin yoğun katkı sağladığı görülmektedir. Bu sonuç araştırma odaklı yapay zekâ pozisyonlarının model tarafından teknik yetkinliklerle güçlü şekilde ilişkilendirildiğini ortaya koymaktadır.

Yapay Zekâ Mühendisi sınıfında da InDo_Yapay Zekâ, InDo_Blokzincir, PR_Görüntü Tanıma ve InDo_Web Geliştirme gibi disiplinler SHAP değerlerine göre yüksek katkı sağlamıştır.

Veritabanı Yöneticisi sınıfında beklenildiği üzere InDo_Veritabanı Yönetimi en belirgin etkileyici unsur olurken, InDo_Bulut Bilişim, PR_Bilgisayar Adli Analisti gibi destekleyici alanlar da modelin kararına katkıda bulunmuştur.

Sağlık Bilişimi Uzmanı sınıfı için InDo_Görüntü İşleme ve InDo_Bulut Bilişim gibi teknik alanların yanı sıra PR_Bilgisayar Adli Analisti gibi alanların da etkili olması, disiplinler arası bilgi yapısının model tarafından dikkate alındığını göstermektedir.

Web Geliştirici sınıfında ise InDo_Web Geliştirme, InDo_Bulut Bilişim ve PR_SQL Sorgu Optimizasyonu gibi doğrudan uygulama geliştiriciliğiyle ilişkili teknik yeterliliklerin SHAP değerlerinde baskın olduğu gözlemlenmiştir.

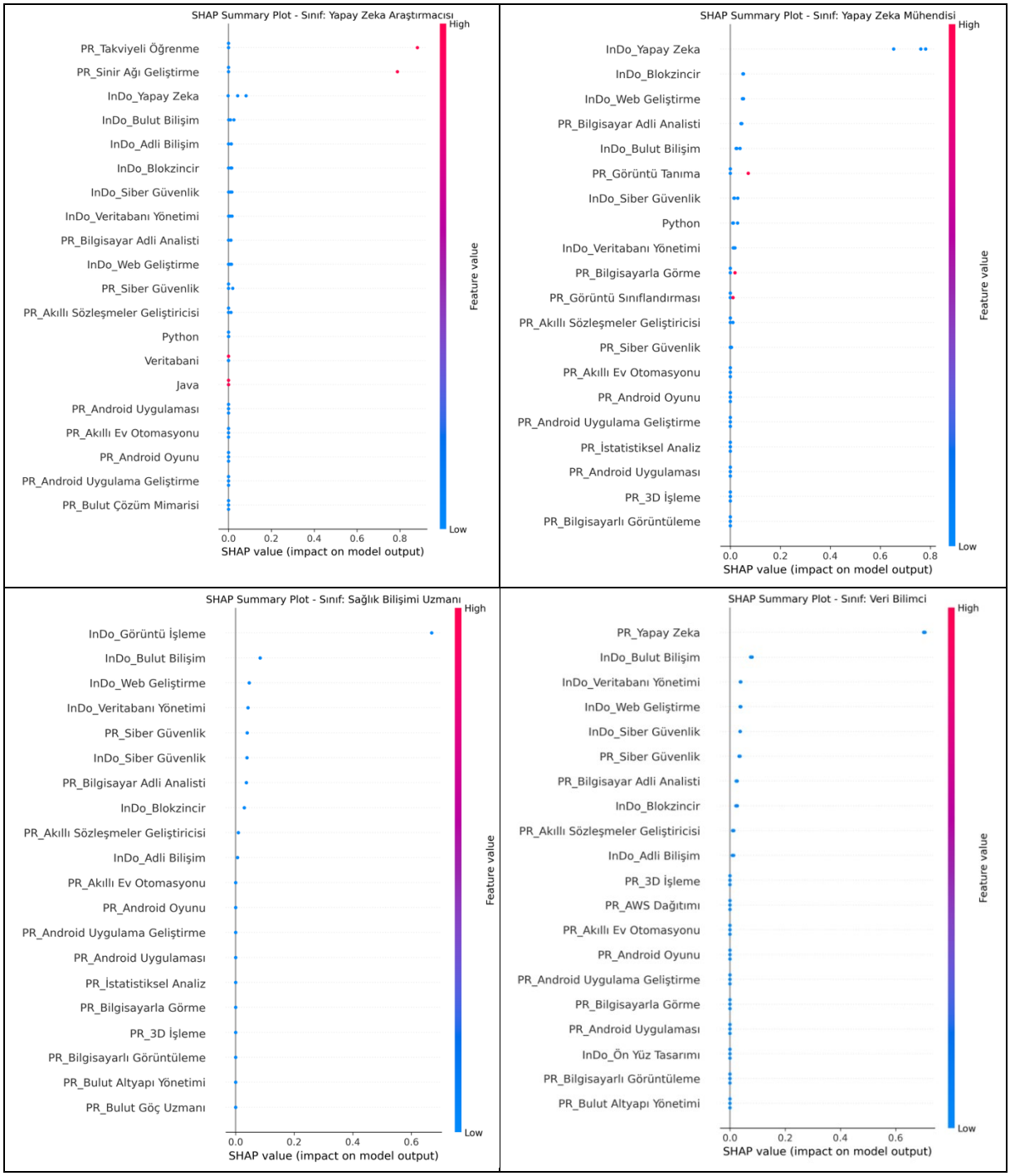
*Sorumlu Yazar/Corresponding Author: yl2230696009@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 133 – 149, 2025
DOI: 10.62301/usmtd.1720980

Son olarak Grafik Programcısı ve Oyun Geliştirici sınıfları için yapılan SHAP analizleri her iki alanda da InDo_Web Geliştirme, InDo_Siber Güvenlik ve PR_3D İşleme gibi teknik alanlara ait düşük ama tutarlı katkılar sergilemiştir. Ancak bu sınıflarda etkili olan özelliklerin SHAP katkılarının diğer meslek gruplarına kıyasla daha az ayrıştığı görülmektedir. Bu durum bu meslek gruplarına özgü belirgin teknik yeterliliklerin daha az temsil edilmesiyle açıklanabilir.

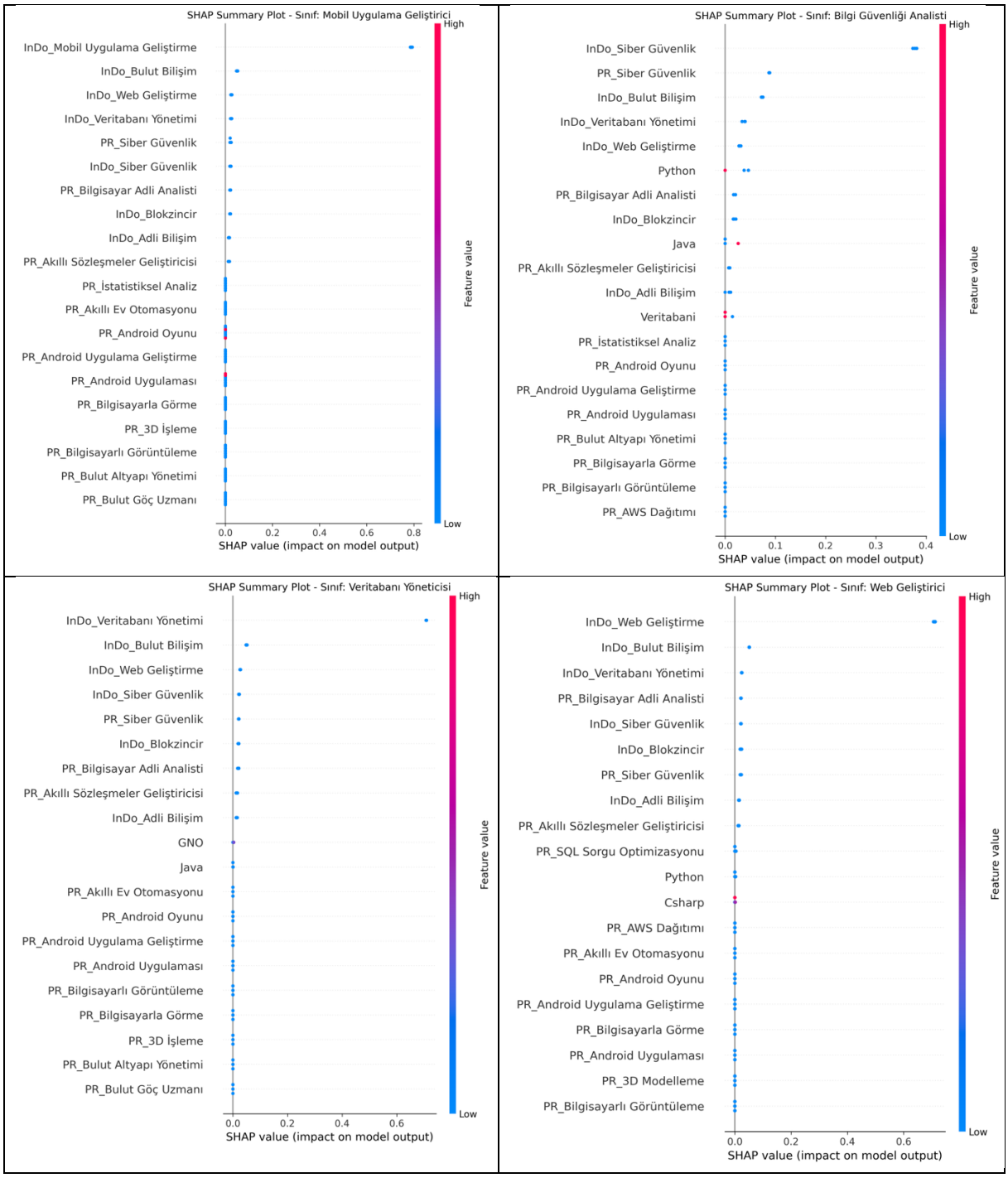
Genel olarak SHAP analizleri modelin her bir sınıfa ilişkin tahminlerinde hangi teknik yeterliliklerin öne çıktığını ortaya koymuş, modelin açıklanabilirliğini artırarak kararlarının sadece doğru değil, aynı zamanda anlaşılır ve yorumlanabilir olmasını sağlamıştır. Böylece veri bilimsel kararların güvenilirliği ve kullanıcıya sunulabilirliği üst düzeye çıkarılmıştır.

Tablo 5. Her bir sınıf için SHAP sonuçları (SHAP results for each class)



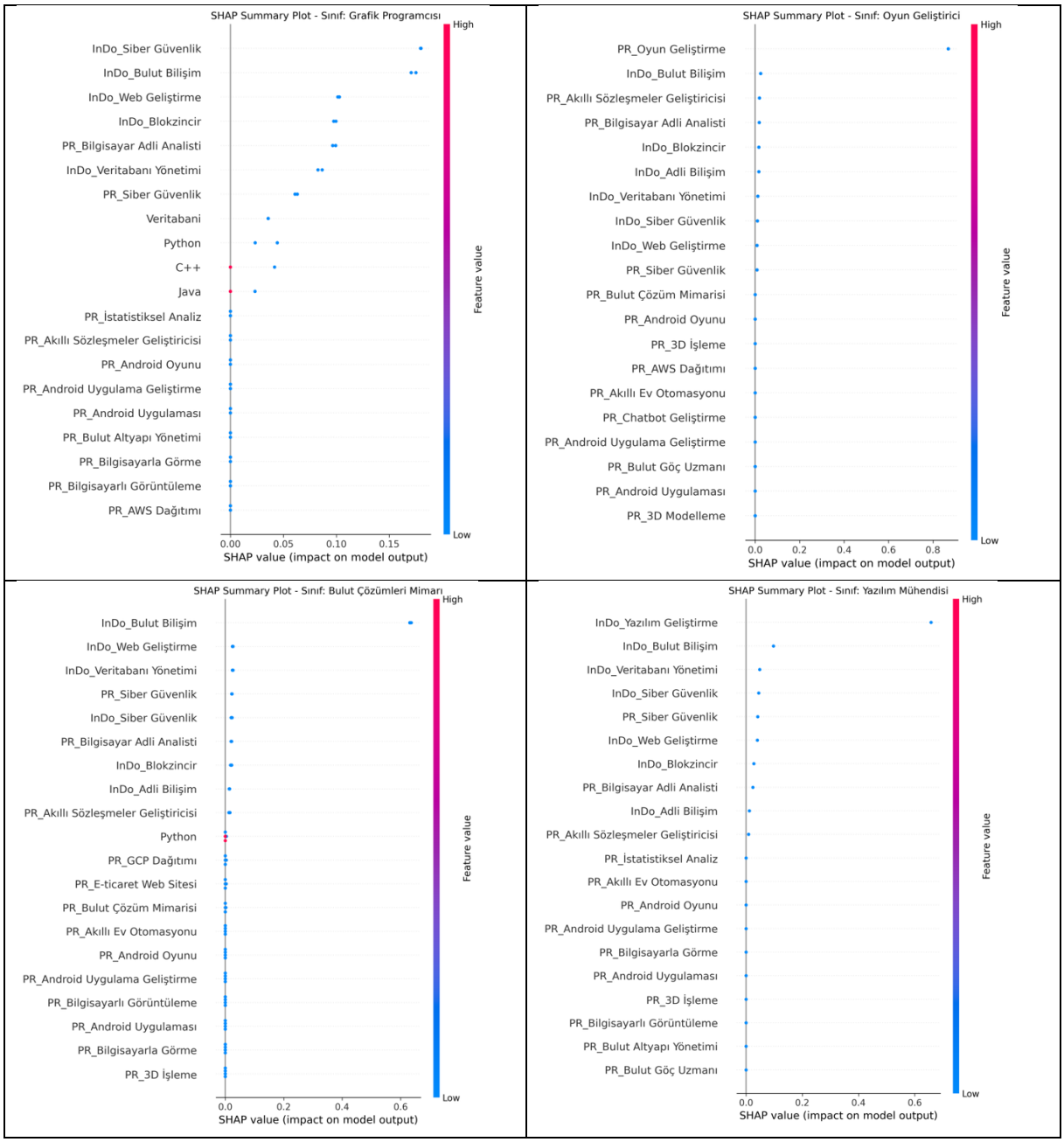
*Sorumlu Yazar/Corresponding Author: yl2230696009@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 133 – 149, 2025
DOI: 10.62301/usmtd.1720980



*Sorumlu Yazar/Corresponding Author: yl2230696009@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 133 – 149, 2025
DOI: 10.62301/usmtd.1720980



4. Tartışma ve Sonuç (Results and Discussion)

Bu çalışmada bilgisayar mühendisliği alanında öğrenim gören veya mezun olan bireylerin sahip oldukları teknik yeterlilikler, programlama bilgileri ve akademik başarı düzeylerine göre hangi meslek grubunda çalıştıklarının tahmin edilmesi amaçlanmıştır. Bu amaç doğrultusunda oluşturulan veri kümesi, genel not ortalaması (GNO), staj yapılan alan, geliştirilen proje, veri tabanı bilgisi ve Python, Java, C#, C++ gibi programlama dillerine ilişkin bilgi düzeylerinden oluşmuştur. Veri kümesi üzerinde gerçekleştirilen veri ön işleme sürecinde eksik veriler temizlenmiş kategorik değişkenler uygun biçimde kodlanmış ve ordinal değişkenler sıralı şekilde sayısallaştırılmıştır. Böylece makine öğrenmesi algoritmalarının eğitilmesine uygun, tutarlı ve dengeli bir veri yapısı elde edilmiştir.

Model seçimi aşamasında farklı sınıflandırma algoritmalarının başarımını karşılaştırmak amacıyla LazyClassifier yapısından yararlanılmış ve en yüksek doğruluk oranını sağlayan modeller tespit edilmiştir. Elde edilen karşılaştırma sonuçlarına göre Gradient Boosting algoritması yaklaşık %97 F1 skorları ile en başarılı model olarak belirlenmiştir. Bu model için GridSearchCV yöntemiyle hiperparametre optimizasyonu gerçekleştirilmiş ve optimum parametre kombinasyonları belirlenerek ana model ortaya çıkmıştır. Modelin değerlendirilmesinde yalnızca doğruluk oranı değil aynı zamanda

*Sorumlu Yazar/Corresponding Author: yl2230696009@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 133 – 149, 2025
DOI: 10.62301/usmtd.1720980

kesinlik, duyarlılık ve F1 skor gibi performans metrikleri de dikkate alınmıştır. Özellikle F1 skoru sınıflar arası dengesizliklerin etkisini azaltan hassas bir gösterge olarak ön plana çıkarılmıştır.

Sınıflandırma modelinin başarısı, karışıklık matrisi ile detaylandırılarak değerlendirilmiştir. Gradient Boosting modeli, birçok meslek sınıfını doğru şekilde sınıflandırabilmiştir. Özellikle veri bilimci, web geliştirici, mobil uygulama geliştirici gibi sınıflarda tüm örnekleri isabetli biçimde tahmin etmiştir. Bazı sınıflar arasında gözlemlenen sınırlı hata payı, meslek gruplarının teknik benzerliklerinden kaynaklanmış ve bu durum da modelin tahmin mantığının mantıklı bir temele dayandığını göstermektedir

Modelin sadece doğruluk düzeyiyle değil aynı zamanda şeffaflık ve yorumlana bilirlik düzeyiyle de değerlendirilmesi amacıyla açıklanabilir yapay zekâ yaklaşımları olan SHAP ve LIME analizlerinden faydalanılmıştır. SHAP analizi ile her bir meslek sınıfı için modelin hangi teknik özniteliklere ne ölçüde ağırlık verdiği ortaya konmuştur. Örneğin veri bilimci tahminlerinde yapay zekâ ve veri tabanı yönetimi bilgileri, yapay zekâ araştırmacısı sınıfında ise sinir ağı ve takviyeli öğrenme gibi ileri düzey teknikler yüksek SHAP değerleriyle öne çıkmıştır. LIME analizleri ise modelin bireysel örnekler üzerindeki kararlarını lokal düzeyde açıklayarak tahminin ardında yatan nedenleri kullanıcıya anlaşılır şekilde sunduğu görülmüştür.

Bu çalışmada elde edilen sınıflandırma modeli sonuçları, literatürde benzer amaçlarla kullanılan makine öğrenmesi algoritmalarıyla karşılaştırıldığında oldukça dikkat çekici bir düzeydedir. Gradient Boosting algoritması, %97.5 doğruluk oranıyla tüm modeller arasında en iyi performansı göstermiştir. Bentéjac vd. (2020) tarafından yapılan kapsamlı çalışmada, Gradient Boosting'in genel olarak %90 üzeri doğruluk sağladığı ifade edilmiştir [3]. Bu bağlamda, çalışmamızda elde edilen sonuçlar, literatürdeki başarı oranlarını aşarak modelin veri kümesine yüksek uyum sağladığını göstermektedir. Decision Tree ve Naive Bayes algoritmaları sırasıyla %95.2 ve %94.2 doğruluk oranı ile güçlü performans sergilemiş, bu oranlar Akgün (2019) gibi benzer sınıflandırma çalışmalarında elde edilen %85-90 aralığındaki sonuçları geçmiştir [1]. Logistic Regression %91.4 doğruluk ile beklenenin üzerinde bir başarı sağlamış, KNN algoritması ise %80 doğrulukla literatürde Uzbaş vd., (2024) %78 gibi değerlerle uyumlu bir sonuç vermiştir. Buna karşın, AdaBoost (%48.6) ve SVC (%45.7) modelleri, sınıf dağılımı dengesizliği ve yüksek boyutlu veri uzayının etkisiyle düşük doğruluk göstermiştir. Literatürde bu iki algoritmanın benzer veri yapılarında sınırlı performans sunduğu ve bu sonuçların olağan olduğu bildirilmektedir [4].Sonuç olarak, çalışmada elde edilen sınıflandırma başarıları, literatür bulgularıyla genel olarak uyumlu olmakla birlikte, bazı modeller açısından daha üstün doğruluk oranlarına ulaşılmıştır.

Sonuç olarak bu çalışma ile öğrencilerin teknik ve akademik yeterliliklerine göre bilgisayar mühendisliği alanındaki yönelimlerinin başarıyla tahmin edilebildiğini göstermekte ve bu sürecin şeffaf bir biçimde yorumlanabileceğini ortaya koymaktadır. Geliştirilen model yüksek doğruluk oranının yanı sıra güçlü açıklana bilirlik özellikleriyle de dikkat çekmekte olup, öğrencilerin kariyer planlaması, insan kaynakları süreçleri ve eğitim danışmanlığı gibi uygulamalarda etkili biçimde kullanılabilecek potansiyele sahip olduğu görülmektedir.

5. Kaynaklar (References)

[1] M. Akgün, Kariyer planlama için karar destek sistemi, Yüksek Lisans Tezi, Hacettepe Üniversitesi Fen Bilimleri Enstitüsü, 2019.

[2] H. Al-Dossari, F.A. Nughaymish, Z. Al-Qahtani, M. Alkahlifah, A. Alqahtani, A machine learning approach to career path choice for information technology graduates, Eng. Technol. Appl. Sci. Res. 10 (6) (2020) 6589–6596.

[3] E.F. Çelik, Makine öğrenmesi ile lise öğrencilerine mesleki alan seçim rehberi, Yüksek Lisans Tezi, Manisa Celal Bayar Üniversitesi Fen Bilimleri Enstitüsü, 2022.

[4] P. Guleria, M. Sood, Explainable AI and machine learning: performance evaluation and explainability of classifiers on educational data mining inspired career counseling, Educ. Inf. Technol. 28 (1) (2023) 1081–1116.

*Sorumlu Yazar/Corresponding Author: yl2230696009@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 133 – 149, 2025
DOI: 10.62301/usmtd.1720980

[5] D.R.D. Haritha, Smart career guidance and recommendation system, Int. J. Eng. Dev. Res. 7 (3) (2019) 633–638.

[6] J. Nizar, R. Sharmilla, K.U. Jaseena, Prediction and assessment of software engineering skillset among computer science students using convolutional neural networks through explainable AI, J. Theor. Appl. Inf. Technol. 102 (21) (2024).

[7] T. Saeed, M. Sufian, M. Ali, A.U. Rehman, Convolutional neural network based career recommender system for Pakistani engineering students, in: Proc. 2021 Int. Conf. Innov. Comput. (ICIC), 2021, pp. 1–10.

[8] Q. Wang, Support vector machine algorithm in machine learning, in: Proc. 2022 IEEE Int. Conf. Artif. Intell. Comput. Appl. (ICAICA), 2022, doi:10.1109/icaica54878.2022.9844516.

***Sorumlu Yazar/Corresponding Author:** yl2230696009@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 150 – 158, 2025
DOI:10.62301/usmtd.1718661

Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi
International Journal of Sustainable Engineering and Technology
ISSN: 2618-6055 / Cilt:9, Sayı:1, (2025), 150 – 158
Araştırma Makalesi – Research Article

BAZALT AGREGASININ TAŞ MASTİK ASFALT KARIŞIMLARDA KULLANIMI

*Hande VAROL MOROVA[1]

[1]Isparta Uygulamalı Bilimler Üniversitesi, Teknoloji Fakültesi, İnşaat Mühendisliği Bölümü

(Geliş/Received: 12.06.2025, Kabul/Accepted: 29.06.2025, Yayınlanma/Published: 30.06.2025)

ÖZ

Günümüzde artan trafik yükleri ve çevresel etkenler, asfalt kaplamaların daha dayanıklı, uzun ömürlü ve yüksek performanslı olmasını zorunlu kılmaktadır. Bu gereksinimleri karşılamak amacıyla geliştirilen Taş Mastik Asfalt (TMA), özellikle ağır taşıt trafiğine maruz kalan yol kesimlerinde tercih edilen bir üstyapı kaplama türüdür. TMA'nın başarısı, içeriğinde kullanılan agrega ve bağlayıcı malzemelerin kalitesi ile doğrudan ilişkilidir. Bu bağlamda, agrega bileşenlerinin özellikleri, karışımın performans kriterlerini önemli ölçüde etkilemektedir.

Bu çalışmada, TMA kaplamalarda ince agrega olarak kullanılan agrega türünün karışım performansına etkisi araştırılmıştır. Çalışmada bazalt agrega yerine daha yaygın ve ekonomik bir alternatif olan kireçtaşı, ince agrega olarak kullanılmıştır. Çalışmanın ilk aşamasında bazalt ve kireçtaşı agregalar kullanılarak TMA numuneler hazırlanmıştır. Hazırlanan karışımların optimum bitüm değerleri Marshall dizayn yöntemi ile belirlenmiştir. Deneysel çalışmalarda, optimum bitüm içeriklerinde hazırlanan bazalt ve kireçtaşı TMA numunelere ait boşluk oranı, agregalar arası boşluk oranı, pratik özgül ağırlık, agregalar arası bitümle dolu boşluk yüzdeleri belirlenerek dolaylı çekme mukavemeti deneyi (ITS) yapılmış ve numunelerin neme karşı hassasiyetleri (TSR) belirlenmiştir. Elde edilen bulgular, TMA karışımlarda kireçtaşının ince agrega olarak kullanılmasının bazı yönlerden üretim kolaylığı ile ekonomik avantajlar sağlayabileceğini göstermiştir.

Anahtar kelimeler: Taş Mastik Asfalt, Bazalt, Kireçtaşı, Marshall Metodu, Nem Hasarı.

USE OF BASALT AGGREGATE IN STONE MASTIC ASPHALT MIXTURES

ABSTRACT

Today, increasing traffic loads and environmental factors require asphalt pavements to be more durable, long-lasting and high-performance. Stone Mastic Asphalt (SMA), which was developed to meet these requirements, is a type of pavement coating preferred especially in road sections exposed to heavy vehicle traffic. The success of SMA is directly related to the quality of the aggregate and binding materials used in its content. In this context, the properties of the aggregate components significantly affect the performance criteria of the mixture.

In this study, the effect of the aggregate type used as fine aggregate in SMA pavements on the mixture performance was investigated. In the study, limestone, which is a more common and economical alternative to basalt aggregate, was used as fine aggregate. In the first stage of the study, SMA samples were prepared using basalt and limestone aggregates. The optimum bitumen values of the prepared mixtures were determined by the Marshall design method. In experimental studies, the void ratio, void in mineral agregate, bulk specific gravity, void filled with bitumen were determined for basalt and limestone SMA samples prepared with optimum bitumen content, indirect tensile strength test (ITS) was performed and the samples sensitivity to moisture (TSR) was determined. The findings obtained showed that the use of limestone as fine aggregate in SMA mixtures could provide economic advantages in some aspects due to ease of production.

Keywords: Stone Mastic Asphalt, Basalt, Limestone, Marshall Method, Moisture Damage.

***Sorumlu Yazar/Corresponding Author:** handevarol@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 150 – 158, 2025
DOI:10.62301/usmtd.1718661

1. Giriş (Introduction)

Artan trafik hacmi, asfalt kaplamalarda erken yaşlanma, deformasyon ve yapısal bozulmalar gibi ciddi performans kayıplarına yol açmaktadır. Bu sorunların önüne geçebilmek ve yol kaplamalarının hizmet ömrünü uzatmak amacıyla, bilim insanları bitümlü karışımların performans özelliklerini iyileştirmeye yönelik yenilikçi yöntemler ve malzeme yapıları geliştirmiştir. Bu doğrultuda öne çıkan uygulamalardan biri de Taş Mastik Asfalt (TMA) tasarımıdır. TMA, özellikle yüksek dayanım, uzun ömür ve deformasyona karşı direnç gerektiren yol yüzeyleri için tercih edilmektedir. TMA'nın karışım yapısı, performans özelliklerini doğrudan etkileyen özel bir oranlamaya sahiptir. Toplam agrega bileşiminin yaklaşık %70-80'i iri daneli agregalardan oluşur; bu yapı, karışım içinde dayanıklı bir taş iskeleti meydana getirir. Karışıma %6-7 oranında bağlayıcı (bitüm), %8-12 oranında mineral dolgu malzemesi ve yaklaşık %0.3-0.5 oranında stabilizasyon amacıyla elyaf (genellikle selülozik) eklenir. İri agregaların sıkı kenetlenmesi sonucu oluşan boşluklar, ince agrega, filler, polimer modifiye bitüm ve elyaftan oluşan mastik harç ile doldurularak yapının hem mekanik dayanımı hem de bağlayıcı stabilitesi artırılır [1]. Agrega derecelendirmesi, asfalt karışımının performansı üzerinde doğrudan etkili olup, yeterli düzeyde esneklik, stabilite, işlenebilirlik ve optimum hava boşluğu içeriği sağlanması açısından kritik bir rol oynar. Uygun şekilde tasarlanmış bir agrega gradasyonu, karışımın hem mekanik dayanımını artırmakta hem de uzun vadeli deformasyonlara karşı direncini güçlendirmektedir. Ayrıca, bu gradasyon, asfaltın yerinde serilmesi ve sıkıştırılması sırasında homojen bir yapı oluşmasına katkıda bulunarak, kalıcı bozulmaların önlenmesine ve servis ömrünün uzamasına yardımcı olur [2].

Türkiye'de TMA uygulamaları ilk olarak 1998 yılında başlatılmış ve 1999 yılından itibaren özellikle yoğun trafik yüklerinin etkili olduğu yol kesimlerinde aşınma tabakası olarak kullanılmaya başlanmıştır. TMA, sahip olduğu yüksek performans özellikleri nedeniyle başta otoyollar, köprü üstü geçişleri, tünel giriş-çıkışları ve ağır taşıt trafiğine maruz kalan güzergâhlar olmak üzere, dayanıklılık ve uzun ömürlü yüzey kaplamalarının gerektiği stratejik noktalarda tercih edilmiştir. Bu uygulama süreci, yalnızca yeni bir kaplama tipi olarak kalmayıp, aynı zamanda asfalt teknolojisinde yenilikçi malzeme kullanımı, üretim teknikleri ve tasarım kriterlerinin araştırıldığı önemli bir mühendislik alanı haline gelmiştir. TMA ile ilgili yapım esasları, üretim koşulları ve kalite kriterleri ise, 2013 yılında yayımlanan Karayolları Teknik Şartnamesi (KTŞ) kapsamında tanımlanmış olup, bu düzenlemeler KTŞ'nin 408. Kısmı içerisinde ayrıntılı olarak yer almaktadır [3,4].

Literatürde TMA karışımlara ilişkin yapılan çalışmalarda Wu vd., (2007), TMA karışımlarda çelik cürufunun agrega olarak kullanımına ilişkin yaptıkları çalışmada düşük sıcaklıkta çatlamaya karşı direncinde iyileşme olduğu vurgulanmıştır [5]. Şengül vd. (2013), SBS polimer modifiye TMA kaplamaların performansını araştırmış, Marshall sonuçlarına göre SBS polimer modifiyeli numunelerin kontrol karışımlara göre daha iyi sonuçlar verdiği ortaya çıkmıştır [6]. Wang vd. (2019), SBS modifiyeli TMA karışımlarda bazalt elyafın performansını araştırarak bazalt elyafın bağlayıcı deney sonuçlarına göre temel özellikleri iyileştirdiği kanıtlanmıştır [7]. Alonso-Troyano vd., (2025) atık tekstil atıklarını mekanik performans ve sürdürülebilirliklerini artırmak için TMA karışımlarda uygulanabilirliğini test etmişlerdir. Aynı zamanda bu çalışma ile yerel olarak üretilen tekstil atıklarının yeniden kullanılmasının, depolama alanında birikimin azaltılmasının ve endüstriler arasında sinerji yaratılmasının çevresel ve ekonomik avantajlarını vurgulanmıştır [8].

Bu çalışmada iki farklı TMA numune serisi hazırlanmıştır. Her iki seri için %1.6 oranında 'ELVALOY™ 5160 Copolymer' (Elvaloy RET) ve %0.2 oranında polifosforik asit (PPA) kullanılarak bitüm modifikasyonu yapılmış, karışıma ise %0.35 oranında selüloz fiber eklenmiştir. İlk TMA karışım gradasyonunda kaba agrega bazalttan, ince agrega bazalttan, filler ise kireçtaşından hazırlanmıştır. İkinci TMA karışımda ise ince agrega kireçtaşı ile değiştirilerek karışım hazırlanmış ve değişen agrega tipinin karışımlar üzerindeki nem hasarına karşı dirençleri incelenmiştir.

2. Materyal ve Metot (Material and Method)

2.1. Agrega (Aggregate)

Çalışmada kullanılan kireçtaşı ve bazalt agrega Isparta Belediyesi'nden temin edilmiştir. Kullanılan kireçtaşı agregasına ait fiziksel özellikler Tablo 1'de, bazalt agregaya ait özellikler ise Tablo 2'de verilmiştir.

*Sorumlu Yazar/Corresponding Author: handevarol@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 150 – 158, 2025
DOI:10.62301/usmtd.1718661

Tablo 1. Kireçtaşı agrega özellikleri (Properties of limestone aggregate)

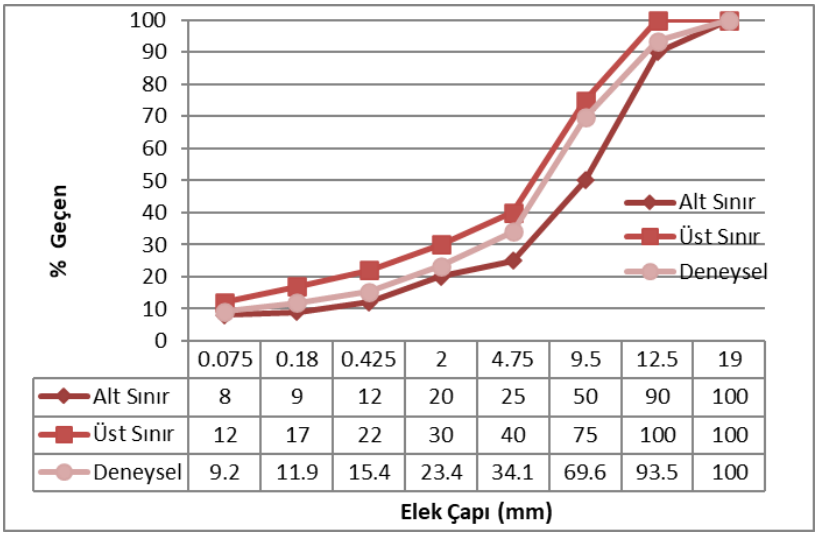
Deney	Değer	Standart
Özgül Ağırlık (gr/cm^3)		
İnce Agrega		
Zahiri Özgül Ağırlık	2.716	
Hacim Özgül Ağırlık	2.674	ASTM C128-88 [9]
Su Absorpsiyonu %	0.6	
Filler		
Hacim Özgül Ağırlık	2.730	ASTM C128-88 [9]
Los Angeles Aşınma (%)	22	ASTM C 131 [10]

Tablo 2. Bazalt agrega özellikleri (Properties of basalt aggregate)

Agrega Deneyleri	Sonuç	Standart
Karışımın Efektif Özgül Ağırlığı	2.61	ASTM D-2041 [11]

Agrega Deneyleri	Sonuç	KTŞ Değerleri [4] (TMA Aşınma Tabakası)	Standart
Parçalanma Direnci (Los Angeles Aşınma Kaybı), (% Kayıp)	20	≤25	AASHTO T-96 [12]
Aşınma Direnci (Mikro-Deval)	9.1		TS EN 1097-1 [13]
Hava Tesirlerine Karşı Dayanıklılık ($MgSO_4$, Donma Kaybı), (%)	1.6	≤14	TS EN 1367-2 [14]
Yassılık İndeksi, (%)	17	≤25	BS 812 [15]
Soyulma Mukavemeti, (%)	70-75	≥60	TS EN12697-11 [16]
Metilen mavisi, (g/kg) (Karışımın)	1.5	≤1.5	TS EN 933-9 [17]

Çalışmada kullanılan agrega gradasyonu Şekil 1'de, numune örnekleri Şekil 2'de gösterilmiştir.



	0.075	0.18	0.425	2	4.75	9.5	12.5	19
Alt Sınır	8	9	12	20	25	50	90	100
Üst Sınır	12	17	22	30	40	75	100	100
Deneysel	9.2	11.9	15.4	23.4	34.1	69.6	93.5	100

Şekil 1. TMA agrega gradasyonu (Gradation of TMA aggregate)

*Sorumlu Yazar/Corresponding Author: handevarol@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 150 – 158, 2025
DOI:10.62301/usmtd.1718661

Şekil 2. Karışım numune örnekleri (Mixture sample examples)

2.2. Bitüm (Bitumen)

Çalışmada Aliağa Rafinerisi'nden temin edilen 50-70 penetrasyonlu bitüm kullanılmıştır. 50-70 penetrasyonlu bitümün özellikleri Tablo 3'te, kullanılan modifiye edilmiş bitüm Şekil 3'te gösterilmiştir.

Tablo 3. 50-70 Bitümlü bağlayıcı özellikleri (Properties of 50-70 penetration bituminous binder)

Deney	Değer	Standart
Penetrasyon (25°C)	58	ASTM D5 [18]
Yumuşama Noktası(°C)	51	ASTM D36 [19]
Özgül Ağırlık (gr/cm^3)	1.037	ASTM D70 [20]

Şekil 3. Çalışmada kullanılan modifiye bitüm (Modified bitumen used in the study)

2.3. Karışım deneyleri (Mixture tests)

TMA numuneleri KTŞ'ye göre hazırlanmış, dizayn için Marshall yöntemi tercih edilmiştir [4]. 1150 gr olarak hazırlanan karışım numunelerine ağırlıkça sırasıyla %5, %5.5, %6, %6.5, %7 oranında bitüm ve %0.35 oranda selüloz fiber eklenmiş, her orandan da üçer adet numune hazırlanmıştır. Elde edilen numunelerin pratik özgül ağırlık (Dp), boşluk oranı (Vh), agregalar arası boşluk oranı (VMA), bitümle dolu boşluk miktarı (Vf) değerleri ve optimum bitüm miktarı (OBM) belirlenmiştir. Asfalt kaplamalar, hizmet süresi boyunca donma-çözülme, sıcaklık değişimleri gibi çeşitli çevresel etkilere maruz kalmaktadır. Bu etkiler, zamanla esnek kaplamaların mekanik özelliklerinde azalmaya neden olarak yapısal bozulmalara yol açabilmektedir. Dolaylı çekme mukavemeti (ITS) deneyi, silindirik asfalt

*Sorumlu Yazar/Corresponding Author: handevarol@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 150 – 158, 2025
DOI:10.62301/usmtd.1718661

numunelerinin çekme dayanımını belirlemek amacıyla uygulanmakta olup, bu yöntemde numuneler Marshall cihazı ile dikey çap doğrultusunda yüklemeye tabi tutulmaktadır [21]. İki TMA serisi için de belirlenen OBM değerleriyle numuneler hazırlanmış ve dolaylı çekme mukavemeti deneyi (ITS) yapılarak karışımların neme karşı hassasiyetleri (TSR) belirlenmiştir.

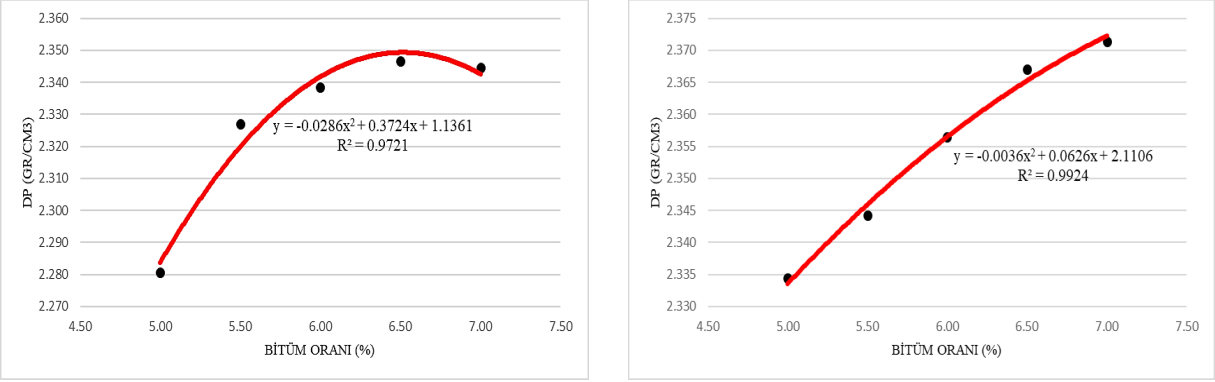
Çalışmada kullanılan serilere ait numune kodları Tablo 4'te verilmiştir.

Tablo 4. TMA Numune serileri ve agrega kombinasyonları (TMA sample series and aggregate combinations)

Seri	Kaba Agrega	İnce Agrega	Filler	Kod
1	Bazalt	Bazalt	Kireçtaşı	BBK
2	Bazalt	Kireçtaşı	Kireçtaşı	BKK

3. Araştırma Bulguları (Research Findings)

OBM belirlemek amacıyla yapılan Marshall Stabilite ve akma deneyi (MS) sonucunda numunelere ait Vh, Vf, VMA ve Dp değerleri elde edilmiştir. Şekil 4'te BBK ve BKK serilerinin pratik özgül ağırlık ile bitüm oranı ilişkisi görülmektedir.



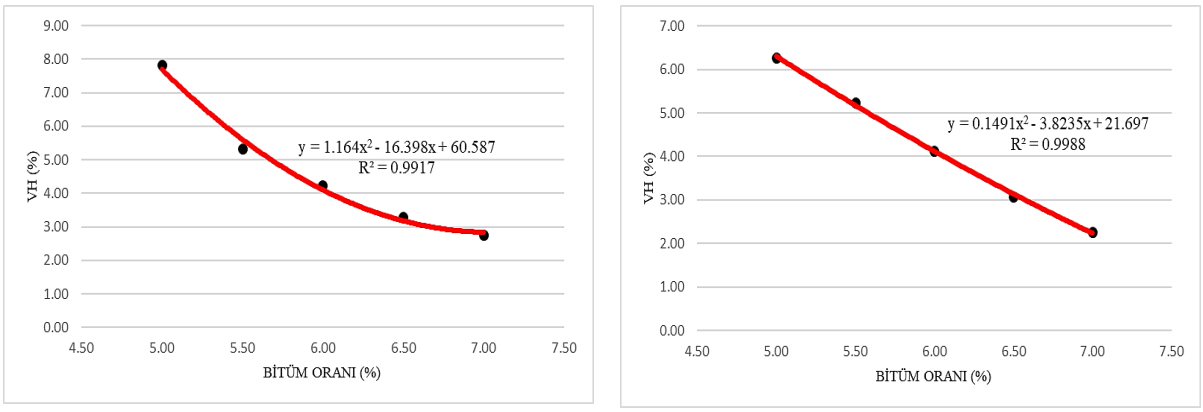
(a) (b)

Şekil 4. BBK ve BKK serilerine ait Dp - bitüm oranı ilişkisi (Relationship between Dp and bitumen content for BBK (a) and BKK (b) series)

BBK karışımına ait Şekil 4.a'da, Dp değeri bitüm oranı arttıkça önce yükselmekte, yaklaşık %6.5 seviyesinde maksimuma ulaştıktan sonra tekrar düşüş göstermektedir. Bu durum, optimum bitüm içeriğinin aşıldığı noktadan itibaren karışıma fazla bitüm eklendiğinde yapının yoğunluğunun azaldığını, dolayısıyla fazla bağlayıcının karışımda boşlukları doldurmaktan ziyade ayrışmaya ve yapısal zayıflamaya neden olduğunu göstermektedir. BKK serisine ait Şekil 4.b'de ise bitüm oranı arttıkça Dp değerinin sürekli yükseldiği ve yukarı yönlü bir parabol oluşturduğu gözlemlenmektedir. Bu durum, artan bitüm miktarının agrega boşluklarını daha iyi doldurarak karışımın yoğunluğunu artırdığını ve daha sıkı bir yapı sağladığını göstermektedir. Genel olarak, BKK karışımı daha doğrusal ve sürekli artan bir yoğunluk eğilimi gösterirken, BBK karışımı optimumdan sonra düşen eğrisiyle bitüm doygunluk noktasını daha belirgin şekilde ortaya koymaktadır. Şekil 5'te BBK ve BKK serilerinin Vh ile bitüm oranı ilişkisi gösterilmiştir.

*Sorumlu Yazar/Corresponding Author: handevarol@isparta.edu.tr

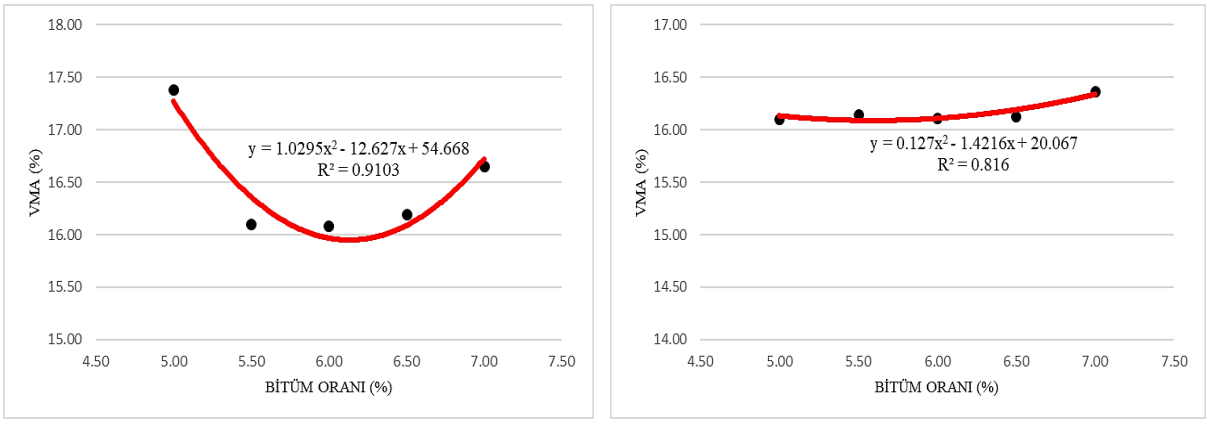
ISSN: 2618-6055 / 9, (1), 150 – 158, 2025
DOI:10.62301/usmtd.1718661



(a) (b)

Şekil 5. BBK ve BKK serilerine ait Vh - bitüm oranı ilişkisi (Relationship between Vh and bitumen content for BBK (a) and BKK (b) series)

BBK serisine ait Şekil 5.a'da, bağlayıcı oranı arttıkça hava boşluğu değerlerinin belirgin şekilde azaldığı görülmektedir. Bu düşüş, bağlayıcının agrega iskeletindeki boşlukları etkili biçimde doldurduğunu ve karışımın daha kompakt bir yapıya kavuştuğunu göstermektedir. BKK serisinde de benzer bir eğilim izlenmektedir. Bitüm oranı arttıkça hava boşluğu oranı sistematik şekilde düşmektedir. Bu düşüş, BKK karışımında da bitümün boşlukları doldurma kapasitesinin yüksek olduğunu ve karışımın daha yoğun bir yapıya ulaştığını göstermektedir. Şekil 6'da BBK ve BKK serilerinin VMA ile bitüm oranı ilişkisi gösterilmiştir.



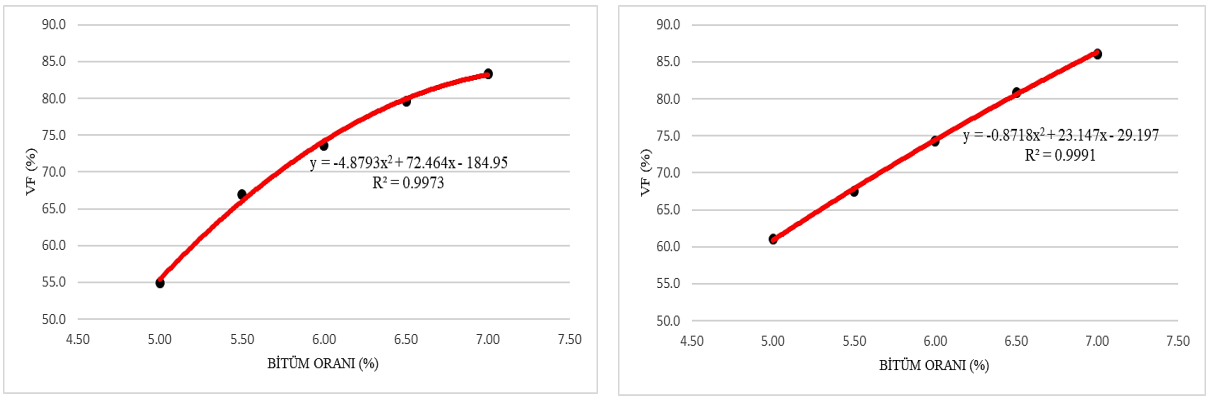
(a) (b)

Şekil 6. BBK ve BKK serilerine ait VMA - bitüm oranı ilişkisi (Relationship between VMA and bitumen content for BBK (a) and BKK (b) series)

BBK serisine ait Şekil 6.a'da, bitüm oranı arttıkça VMA değerinde daha belirgin bir eğilim gözlemlenmektedir. Eğri, yaklaşık %6.3 civarında minimum VMA değerine ulaşmakta, bu da karışımın optimum bitüm içeriğine yakın bir noktada en yoğun hâlini aldığını göstermektedir. Bu durum, agregaların daha etkin biçimde bağlayıcıyla kaplandığını ve boşlukların daha iyi doldurulduğunu ifade etmektedir. BKK serisine ait Şekil 6.b'de ise VMA değerinin bitüm oranına bağlı olarak daha sınırlı ve yatay bir değişim gösterdiği görülmektedir. Eğrinin belirgin bir minimum veya maksimum nokta içermemesi, BKK karışımının agregalar arası boşluk yapısının bitüm oranındaki değişimlerden önemli ölçüde etkilenmediğini göstermektedir. Şekil 7'de BBK ve BKK serilerinin Vf ile bitüm oranı ilişkisi gösterilmiştir.

*Sorumlu Yazar/Corresponding Author: handevarol@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 150 – 158, 2025
DOI:10.62301/usmtd.1718661



(a) (b)

Şekil 7. BBK ve BKK serilerine ait Vf - bitüm oranı ilişkisi (Relationship between Vf and bitumen content for BBK (a) and BKK (b) series)

BBK serisine ait Şekil 7.a'da Vf değeri, bitüm oranı arttıkça düzenli bir şekilde yükselmektedir. Bu artış, artan bağlayıcının agrega iskeleti içindeki boşlukları doldurma kapasitesinin oldukça yüksek olduğunu ve bağlayıcının büyük oranda boşluklara nüfuz ettiğini göstermektedir. BKK serisinde de benzer bir artış eğilimi gözlemlenmektedir. Bu artış, bağlayıcının boşlukları yüksek oranda doldurduğunu ve karışımın yoğunlaştığını göstermektedir. Şekil 7.a'nın daha kıvrımlı olması, BBK serisinin Vf açısından bitüm oranına daha hassas tepki verdiğini göstermektedir.

Tüm bu veriler ışığında BBK serisi için OBM değeri %6.29; BKK serisi için ise OBM değeri %6.31 olarak bulunmuştur. Çalışmada OBM değerleri esas alınarak hazırlanan yeni BBK ve BKK numunelerine ITS testi uygulanmış ve suyun neden olduğu kaplama bozulmasına karşı dirençleri belirlenmiştir. ITS ve TSR değerleri Tablo 5'te verilmiştir.

Tablo 5. BBK ve BKK numunelerinin ITS ve TSR sonuçları (ITS and TSR results of BBK and BKK samples)

	Koşullandırılmış-ITS (kPa)	**Koşullandırılmamış-ITS (kPa)**	**TSR**
BBK	869.8	959	0.91
BKK	813.9	914.5	0.89

Yapılan dolaylı çekme dayanımı (ITS) ve çekme dayanımı oranı (TSR) testleri, BBK ve BKK karışımlarının nem hasarına karşı gösterdiği direnci ortaya koymaktadır. BBK karışımı, koşullandırılmamış (kuru) durumda 959 kPa, koşullandırılmış (ıslak) durumda ise 869.8 kPa dayanım göstermiştir. Buna karşılık BKK karışımının kuru ve ıslak koşullardaki ITS değerleri sırasıyla 914.5 kPa ve 813.9 kPa olarak ölçülmüştür. Her iki karışımda da nem etkisiyle bir miktar dayanım kaybı yaşanmakla birlikte, TMA karışımları için KTŞ'de belirtilen minimum TSR sınırı olan 0.80 değerinin aşıldığı görülmektedir [4]. TSR değeri BBK için 0.91, BKK için ise 0.89'dur. Bu sonuçlara göre, BBK karışımı hem kuru hem de ıslak koşullarda daha yüksek mukavemet sergilemiş ve suya maruz kaldığında daha az performans kaybı yaşamıştır. Dolayısıyla, her iki karışım da teknik açıdan yeterli dayanımı sağlamakla birlikte, BBK karışımı nem hasarına karşı daha güvenilir ve dayanıklı bir yapı sunmaktadır.

4. Tartışma ve Sonuç (Discussion and Conclusion)

Bu çalışmada, TMA karışımlarda kullanılan ince agrega türünün karışım performansına etkisi değerlendirilmiş ve bu kapsamda BBK serisinde bazalt, BKK serisinde ise kireçtaşı ince agrega olarak kullanılmıştır. KTŞ'ye göre TMA karışımlarda OBM'nin en az %5.8 olması istenmektedir [4]. Bu bağlamda hem BBK hem de BKK serisi numunelere ait OBM değerleri şartname sınırını sağlamaktadır. KTŞ'ye göre TMA kaplamalarda VMA değerinin en az %16 olması istenmektedir [4]. BBK ve BKK serilerine ait VMA değerleri incelendiğinde her iki serinin de şartname değerlerini sağladığı görülmüştür. Bununla birlikte, karışımlar arasındaki performans farkı özellikle ITS ve TSR değerlerinde belirginleşmiştir.

*Sorumlu Yazar/Corresponding Author: handevarol@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 150 – 158, 2025
DOI:10.62301/usmtd.1718661

BBK karışımı hem kuru hem de ıslak numunelerde BKK'ye göre daha yüksek ITS değerleri vermiş ve 0.91'lik TSR değeri ile bitüm-agrega bağının nem etkisi altında daha kararlı olduğunu ortaya koymuştur. Buna karşın BKK karışımı 0.89 TSR değeri ile minimum şartname sınırı olan 0.80'in üzerinde kalsa da suya karşı dayanım açısından bazalt agregası kullanılan karışımdan daha fazla performans kaybı göstermiştir.

Sonuç olarak, TMA karışımlarda agrega türü performans üzerinde doğrudan etkili olmakta, bazalt kullanılan BBK serisi**,** kireçtaşı içeren BKK serisine kıyasla nem hasarına karşı daha dayanıklı bir yapı sunmaktadır.

5. Kaynaklar (References)

[1] S.S.M. Naser, M. Seyedi, S. Al-Busaltan, Enhancing Stone Mastic Asphalt through the integration of waste paper and Cement Kiln Dust, J. Civ. Hydraul. Eng. 1 (1) (2023) 23–37.

[2] A.M. Babalghaith, S. Koting, N.H.R. Sulong, M.Z.H. Khan, A. Milad, N.I.M. Yusoff, A.H.B.N. Mohamed, A systematic review of the utilization of waste materials as aggregate replacement in stone matrix asphalt mixes, Environ. Sci. Pollut. Res. 29 (24) (2022) 35557–35582.

[3] M. Aslan, Mastik Asfalt Üretiminde Pomza, Perlit ve Ahlat Taşı Agregasının Kullanılabilirliği, Master's Thesis, Bitlis Eren Univ. and Dicle Univ. Inst. of Science, Bitlis, 2018.

[4] Karayolları Genel Müdürlüğü, Karayolu Teknik Şartnamesi (KTŞ), Ankara, 2013.

[5] S. Wu, Y. Xue, Q. Ye, Y. Chen, Utilization of steel slag as aggregates for stone mastic asphalt (SMA) mixtures, Build. Environ. 42 (7) (2007) 2580–2585.

[6] C.E. Şengül, S. Oruç, E. İskender, A. Aksoy, Evaluation of SBS modified stone mastic asphalt pavement performance, Constr. Build. Mater. 41 (2013) 777–783.

[7] W. Wang, Y. Cheng, P. Zhou, G. Tan, H. Wang, H. Liu, Performance evaluation of styrene-butadiene-styrene-modified stone mastic asphalt with basalt fiber using different compaction methods, Polymers 11 (6) (2019) 1006.

[8] C. Alonso-Troyano, D. Llopis-Castelló, B. Olaso-Cerveró, Incorporating recycled textile fibers into stone mastic asphalt, Buildings 15 (8) (2025) 1310.

[9] ASTM C 128-88, Test Method for Specific Gravity and Adsorption of Fine Aggregate, Annual Book of ASTM Standards, USA, 1992.

[10] ASTM C 131-96, Standard Test Method for Resistance to Abrasion of Small Size Coarse Aggregate by Use of the Los Angeles Machine, Annual Book of ASTM Standards, USA, 1996.

[11] ASTM D2041–03, Standard Test Method for Theoretical Maximum Specific Gravity and Density of Bituminous Paving Mixtures, Annu. Book ASTM Stand. 4.03 (2003) 166–169.

[12] AASHTO T 96, Standard Method of Test for Resistance to Degradation of Small-size Coarse Aggregate by Abrasion and Impact in the Los Angeles Machine, American Association of State Highway and Transportation Officials, 2002.

[13] Türk Standartları Enstitüsü, TS EN 1097-1. Agregaların mekanik ve fiziksel özellikleri için deneyler – Bölüm 1: Aşınma direncinin tayini (microDeval), Ankara, Türkiye, 2011.

[14] Türk Standartları Enstitüsü, TS EN 1367-2. Agregaların termal ve bozunma özellikleri için deneyler – Bölüm 2: Magnezyum sülfat deneyi, Ankara, Türkiye, 2010.

[15] BS 812, Methods for Determination of Physical Properties, British Standards Institution, UK, 1975.

[16] Türk Standartları Enstitüsü, TS EN 12697-11. Bitümlü karışımlar – Deney metotları – Bölüm 11: Agrega ve bitüm arasındaki bağlanmanın tayini, Ankara, Türkiye, 2012.

[17] Türk Standartları Enstitüsü, TS EN 933-9. Agregaların geometrik özellikleri için deneyler – Bölüm 9: İnce tanelerin tayini – Metilen mavisi deneyi, Ankara, Türkiye, 2001.

***Sorumlu Yazar/Corresponding Author:** handevarol@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 150 – 158, 2025
DOI:10.62301/usmtd.1718661

[18] ASTM D5, Standard Test Method for Penetration of Bituminous Materials, Annual Book of ASTM Standards, USA, 1992.

[19] ASTM D36, Standard Test Method for Softening Point of Bitumen (Ring-and-Ball Apparatus), Annual Book of ASTM Standards, USA, 1992.

[20] ASTM D70, Standard Test Method for Density of Semi-Solid Bituminous Materials (Pycnometer Method), Annual Book of ASTM Standards, USA, 1992.

[21] S.E. Zoorob, L.B. Suparma, Laboratory design and investigation of the properties of continuously graded asphaltic concrete containing recycled plastics aggregate replacement (Plastiphalt), Cem. Concr. Compos. 22 (4) (2000) 233–242.

***Sorumlu Yazar/Corresponding Author:** handevarol@isparta.edu.tr

ISSN: 2618-6055 / 9, (1), 159 – 173, 2025
DOI:10.62301/usmtd.1710461

Uluslararası Sürdürülebilir Mühendislik ve Teknoloji Dergisi
International Journal of Sustainable Engineering and Technology
ISSN: 2618-6055 / Cilt:9, Sayı:1, (2025), 159 – 173
Araştırma Makalesi – Research Article

GAMIFICATION IN COMPUTER ENGINEERING EDUCATION: A BIBLIOMETRIC REVIEW OF GLOBAL RESEARCH TRENDS

*Gürcan ÇETİN, Saadet KURU ÇETİN

[1]Muğla Sıtkı Koçman University, Information Systems Engineering, Muğla
[2]Muğla Sıtkı Koçman University, Department of Educational Sciences, Muğla

(Geliş/Received: 31.05.2025 Kabul/Accepted: 30.06.2025, Yayınlanma/Published: 30.06.2025)

ABSTRACT

With the impact of digitalization, innovative approaches have emerged in the field of education, making learning processes more effective and engaging through gamification and digital games. Gamification aims to enhance individuals' motivation and participation by using game design elements (points, badges, levels, leaderboards, etc.) in non-game contexts. This approach is particularly important in disciplines with technical content, such as computer engineering, to capture students' attention and increase interaction. This study aims to examine scientific publications on gamification in computer engineering education using bibliometric analysis. A total of 343 publications indexed in the Web of Science (WoS) database between 2012 and 2025 were analyzed. The analysis covers publication trends by year, most cited works, influential authors, institutions, journals, and keyword clusters. This study reveals the intellectual structure of the field, identifies research trends, collaboration networks, and potential gaps, aiming to guide future studies.

Keywords: Bibliometric Analysis, Computer Engineering, Educational Technology, Gamification.

BİLGİSAYAR MÜHENDİSLİĞİ EĞİTİMİNDE OYUNLAŞTIRMA: KÜRESEL ARAŞTIRMA EĞİLİMLERİNE YÖNELİK BİBLİYOMETRİK BİR İNCELEME

ÖZ

Dijitalleşmenin etkisiyle eğitim alanında da yenilikçi yaklaşımlar öne çıkmakta, oyunlaştırma ve dijital oyunlar öğrenme süreçlerini daha etkili ve motive edici hâle getirmektedir. Oyunlaştırma, oyun dışı bağlamlarda oyun tasarım öğeleri (puan, rozet, seviye, liderlik tablosu vb.) kullanılarak bireylerin motivasyonunu ve katılımını artırmayı hedeflemektedir. Bu yaklaşım, özellikle teknik içerikli bilgisayar mühendisliği gibi disiplinlerde öğrencilerin dikkatini çekmek ve etkileşimi artırmak açısından önemlidir. Bu çalışma, bilgisayar mühendisliği eğitiminde oyunlaştırma konusundaki bilimsel yayınları bibliyometrik analiz yöntemiyle incelemeyi amaçlamaktadır. Web of Science (WoS) veri tabanında 2012–2025 yılları arasında yayımlanmış 343 çalışma analiz edilmiştir. Analiz kapsamında yıllara göre yayın eğilimleri, en çok atıf alan çalışmalar, etkili yazarlar, kurumlar, dergiler ve anahtar kelime kümelenmeleri değerlendirilmiştir. Bu çalışma, alanın entelektüel yapısını ortaya koyarak araştırma eğilimlerini, iş birliği ağlarını ve potansiyel boşlukları tanımlamaktadır. Böylece, gelecek çalışmalara yön vermeyi amaçlamaktadır.

Anahtar kelimeler: Bibliometrik naliz, Bilgisayar Mühendisliği, Eğitim Teknolojisi, Oyunlaştırma.

***Corresponding Author:** gcetin@mu.edu.tr

ISSN: 2618-6055 / 9, (1), 159 – 173, 2025
DOI:10.62301/usmtd.1710461

1. Introduction

The rapid advancement of technology and the pervasive influence of digitalization in all areas of life have led to profound transformations in the field of education. Alongside traditional teaching methods, innovative approaches that render learning processes more effective, motivating, and interactive have come to the forefront. In this context, gamification and digital games have emerged as prominent pedagogical tools, particularly aligned with the learning dynamics of the digital age.

Gamification is defined as the use of game design elements in non-game contexts to motivate individuals, enhance engagement, and encourage problem-solving [1, 2]. Initially adopted in the digital media industry, the concept has rapidly expanded into various domains since 2010, including healthcare [3], human resources [4], marketing [5], environmental protection [6], and most notably, education [7].

The dynamics of the digital age have transformed educational environments, necessitating the adoption of innovative methods capable of capturing learners' attention. In the educational context, gamification is an innovative approach that involves the integration of game design elements—such as points, badges, level progression, and leaderboards—into learning environments to enhance learner motivation and engagement [8-10]. Kalogiannakis et al. [10], through an analysis of 24 studies published between 2012 and 2020, highlighted the positive effects of gamification on learning outcomes, social interaction, and motivation.

Traditional instructional methods, particularly in computer engineering courses that involve technical and abstract content, often fall short in ensuring sufficient interaction and motivation, leading to a search for new pedagogical approaches [9]. As such, the gamification of educational content has become especially important in disciplines such as computer engineering, which are heavily focused on technical and abstract concepts [11]. Courses such as algorithms, programming, software testing, and data structures are often perceived by students as complex and demotivating. Therefore, gamification in computer engineering education is considered an effective pedagogical tool for increasing motivation, supporting experiential learning, and fostering critical thinking and problem-solving skills [11-13].

Although various studies have examined the effects of gamification in computer engineering education, these studies have predominantly focused on pedagogical outcomes or practical implementations [10, 13]. However, there is a notable lack of studies that systematically analyze the gamification literature through bibliometric criteria such as publication trends, author network relationships, citation analyses, and thematic intensities. This gap hinders the structural mapping of the knowledge base and the generation of guiding insights for future research.

In this context, the aim of the present study is to systematically examine gamification-based research in computer engineering education by uncovering the intellectual structure, thematic development, collaborative networks, and citation hotspots within the field. Accordingly, this study seeks not only to describe the current state of the literature but also to provide guidance for future academic inquiries. In line with these objectives, the following research questions will be addressed:

1. What is the annual distribution of publications on gamification in computer engineering education?
2. Which studies on gamification in the context of computer engineering have received the highest number of citations, and what is their intellectual impact within the literature?
3. Who are the most prolific authors in the field of gamification in computer engineering education, and what is the level of their citation impact?
4. Which institutions have published the highest number of studies on this theme?
5. Which academic journals publish the most research on gamification in computer engineering education?
6. Based on keyword frequencies and word cloud analysis, which conceptual trends and thematic clusters stand out in the gamification literature?
7. During which periods has research on gamification in computer engineering education become more prominent?

*Corresponding Author: gcetin@mu.edu.tr

ISSN: 2618-6055 / 9, (1), 159 – 173, 2025
DOI:10.62301/usmtd.1710461

2. Methodology

In this study, the quantitative research method of bibliometric analysis was employed to identify the publications and research trends related to the concept of gamification in computer engineering education. The bibliometric method is particularly suitable for evaluating the number and significance of empirical contributions in a given field, identifying similarities and differences within the literature, and constructing a research map [14]. In the study, descriptive bibliometrics was used to measure the productivity of the identified publications, while evaluative bibliometrics was utilized to assess the use of literature.

Bibliometric analysis is a method that reveals the relationships among publications and different authors. This approach allows researchers to base their work on the bibliographic contributions of other scholars and to express their ideas through writing and collaboration. When such data are collected and analyzed, insights into social networks, current areas of interest, and the structural characteristics of the field can be obtained [15]. The bibliometric method can be both descriptive and evaluative in nature. Common data sources are widely used in most bibliometric analyses. While Web of Science (WoS) and Scopus databases are widely preferred in bibliometric analyses [16], the results of bibliometric analyses may vary depending on the database used [17]. WoS and Scopus handle bibliographic metadata differently. WoS processes and reformats references to include details such as first author, year, journal, issue, and DOI. Whereas Scopus retains all APA-style citations provided by authors [14]. This means that although the tools have been perfected, limitations still prevent some analyses from being performed by merging WoS and Scopus data [18]. Therefore, considering the problems related to the merging of the data presented in different cell formats and the ease of classification of information retrieved from WoS in a research database, in this study, bibliometric data in WoS (including all indexes) core collection were included in the review. Another reason to choose WoS over SCOPUS was that WoS is a collection of databases that index the world's most authoritative scholarly literature in the social sciences, arts, and humanities [19].

Table 1. The dataset

Main Information About Data	
Timespan	2012-2025
Sources (Journals, Books, etc.)	178
Documents	343
Annual Growth Rate %	24,9
Document Average Age	4,13
Average citations per doc	27,62
References	15989
Document Contents	
Keywords Plus (ID)	523
Author's Keywords (DE)	1243
Authors	
Authors	1276
Authors of single-authored docs	33
Authors Collaboration	
Single-authored docs	33
Co-Authors per Doc	3,93
International co-authorships %	21,87
Document Types	
article	343

*Corresponding Author: gcetin@mu.edu.tr

ISSN: 2618-6055 / 9, (1), 159 – 173, 2025
DOI:10.62301/usmtd.1710461

2.1. Creating Dataset

Articles related to gamification in computer engineering education were examined by accessing the Web of Science (WoS) Core Collection databases. Following a comprehensive literature review on publications in this field, relevant keywords were identified, and the final search query was formulated through consultations with experts in bibliometric research. The information about the dataset is presented in Table 1.

Table 1 above shows general information about 343 studies obtained with the topic Gamıfıcatıon In Computer Engıneerıng Educatıon. Within the scope of WoS, studies between 2012 and 2025 were analyzed. These studies have an average citation rate of 27,62 per publication. We understand that the number of single-author studies, n=33, is significantly lower than the total number of studies.

In this study, bibliometric analyses were conducted using biblioshiny, a web-based application built on the R package bibliometrix (Aria & Cuccurullo, 2017). Biblioshiny is a shiny app that provides an interactive graphical interface for performing bibliometric analysis, facilitating user-friendly and comprehensive exploration of bibliographic data.

2.2. Planning, Selection, Extraction and Execution Process in Bibliometric Analysis

This section covers the planning, selection, data extraction, and implementation stages of the analysis process. During the inclusion/exclusion phase, as illustrated in Figure 1, the articles retrieved through the search query were assessed according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) standards.

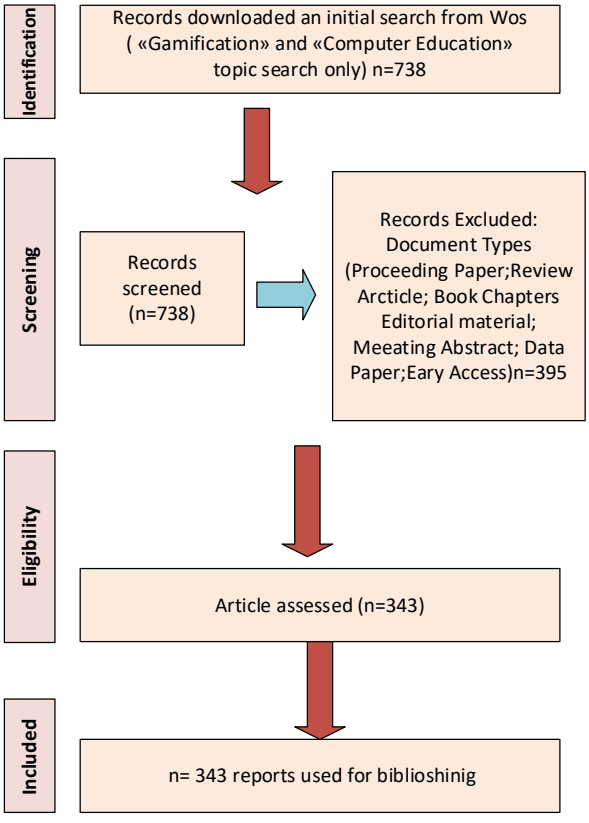


Figure 1. PRISMA flow chart of the article screening process

In the first stage, the selection of internationally recognized scientific databases was carried out during the planning phase. The Web of Science (WoS) database was chosen for this study due to its broad indexing of scientific journals and its distinction from other databases. The second stage involved the identification of keywords related to the research topic or question. This stage is a cornerstone of any successful bibliometric study. The search query was limited to articles containing the terms "computer

*Corresponding Author: gcetin@mu.edu.tr

ISSN: 2618-6055 / 9, (1), 159 – 173, 2025
DOI:10.62301/usmtd.1710461

education" and "gamification" in the topic field. As for document type, only journal articles were selected. In the third stage, inclusion criteria were applied. No time restrictions were imposed in order to include the most recent and relevant studies. Only peer-reviewed articles were included in the analysis, as they undergo rigorous review processes and meet established quality standards. In the fourth and final stage, the analysis process was carried out based on the records of 343 potentially relevant studies. At this point, a detailed evaluation was conducted to determine whether these studies were truly related to the topic of interest. Ultimately, 343 studies were thoroughly analyzed.

3. Research Findings

This section first presents the publication outputs related to gamification in computer engineering education, followed by an analysis of the sources, the most prolific researchers in the field, the most relevant institutions, and the most highly cited publications. Additionally, co-citation analysis of the articles, examination of the source journals, keyword trends, and the evolution of publications and citations are explored to reveal overall research trends.

3.1. Annual Number of Publications

As part of the annual literature output, the average yearly number of publications related to the keyword gamification in computer engineering education indexed in the Web of Science (WoS) is presented in Figure 2.

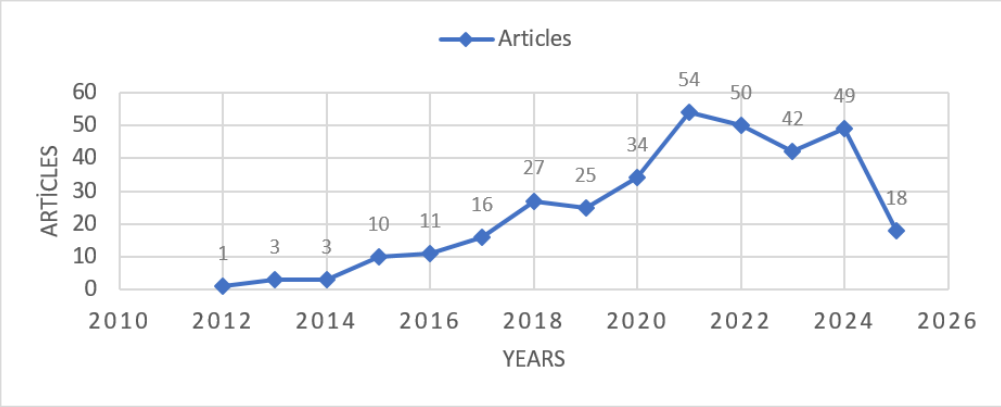


Figure 2. Average annual number of publications

The annual distribution of the articles examined within the scope of the bibliometric analysis is presented in Table 2. Between 2012 and 2015, the number of publications on gamification in computer engineering education remained quite limited, with only 1 to 3 articles published per year during this period. From 2016 onwards, a steady increase in publication output was observed, with a notable surge in 2018, reaching 27 articles. The most significant rise occurred in 2021, when the number of publications peaked at 54, marking the highest point in the literature.

In the following two years, 2022 and 2023, 50 and 42 articles were published, respectively. In 2024, the number increased again to 49. However, in 2025, a sharp decline was observed, with the publication count dropping to 18. This drop may be attributed to the fact that data for 2025 were collected before the year had concluded.

This trend indicates that academic interest in gamification-related studies has significantly increased, particularly after 2018, highlighting gamification as an emerging research theme in the field. At the same time, sudden declines in publication counts may not only reflect limitations in data coverage but also suggest temporary saturation or shifts in research focus within the academic community.

*Corresponding Author: gcetin@mu.edu.tr

ISSN: 2618-6055 / 9, (1), 159 – 173, 2025
DOI:10.62301/usmtd.1710461

3.2. Most Globally Cited Publications

Citations to studies can be analyzed using bibliometric methods at both local and global levels [14]. When examining publications related to gamification in computer engineering education, Figure 3 presents the ten most globally cited studies on gamification in computer engineering education

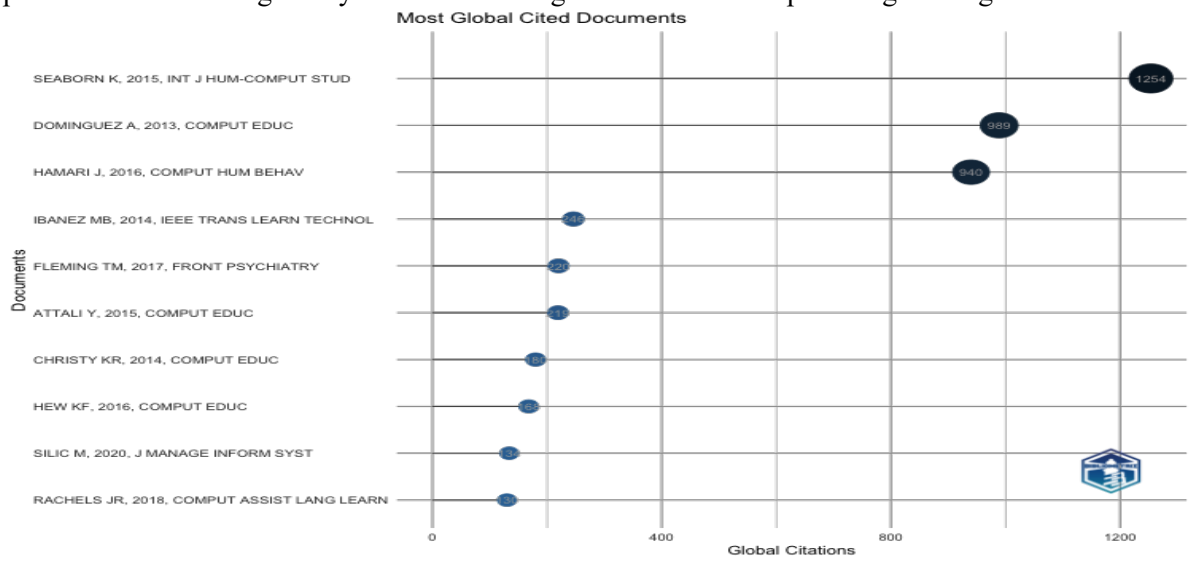


Figure 3. Top ten most globally cited publications

As shown in Figure 3, the most cited publications on gamification in the context of computer engineering education represent the foundational and influential works in the literature. The most highly cited document is the conceptual review by Seaborn et al. [8], which has received a total of 1,254 citations. This study provides a comprehensive conceptual framework for understanding gamification and has had a significant impact on subsequent literature.

The second most cited work is an experimental study by Domínguez et al. [20], published in Computers & Education, with 980 citations. This article is notable for empirically demonstrating the effects of gamification on learning outcomes. The third most cited publication is by Hamari et al. [21], published in Computers in Human Behavior, which received 940 citations. This study offers systematic findings on the effects of gamification on user behavior and is frequently cited in both educational and digital media research.

The remaining publications on the list generally contribute to the fields of educational technology, computer-assisted learning, and cognitive sciences. Noteworthy contributions include studies by Banerjee et al. [22], Fleming et al. [23], and Attali et al. [24], among others. These documents, due to their high citation counts and publication across various disciplines, reflect the interdisciplinary nature of research on gamification.

This analysis reveals that the most influential studies on gamification were predominantly published between 2013 and 2015, a period during which the theoretical and empirical foundations of the field were solidified. These publications continue to serve as key citation hubs and knowledge clusters for researchers in this domain.

*Corresponding Author: gcetin@mu.edu.tr

ISSN: 2618-6055 / 9, (1), 159 – 173, 2025
DOI:10.62301/usmtd.1710461

3.3. Analysis of the Source Journals

Journals that have published articles on gamification in computer engineering education were examined, and based on the number of publications, the top ten sources are presented below.

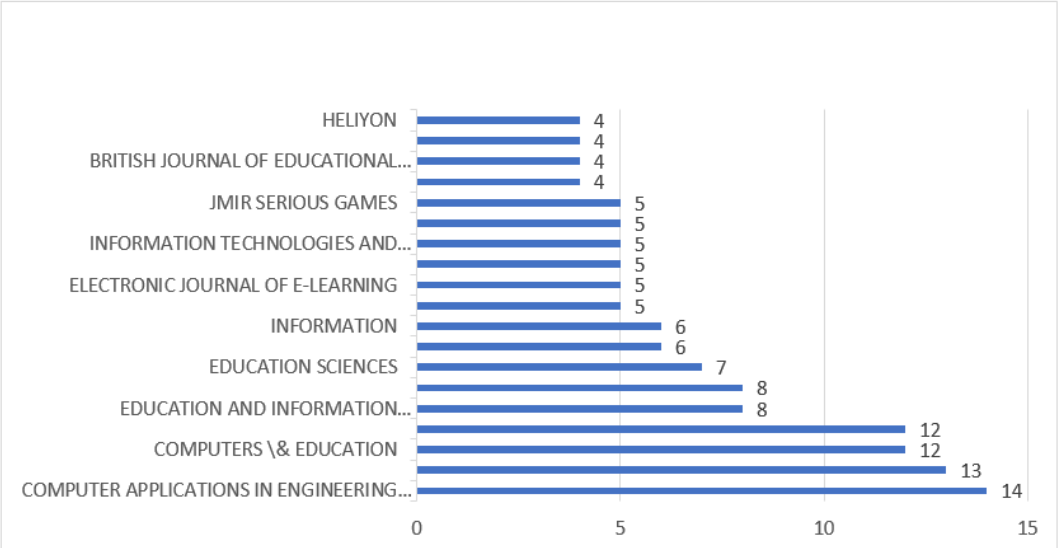


Figure 4. Top ten journals publishing gamification-themed articles (by subject area)

As shown in Figure 4, the journals that most frequently publish articles on gamification in computer engineering education are identified. Among them, Computer Applications in Engineering Education stands out as the leading source, hosting 14 publications. It is followed by Computers & Education with 13 articles, and both Education and Information Technologies and Education Sciences with 8 articles each.

This distribution indicates that research on gamification is concentrated in journals that focus both on engineering education and on educational technology, reflecting the interdisciplinary nature of the field.

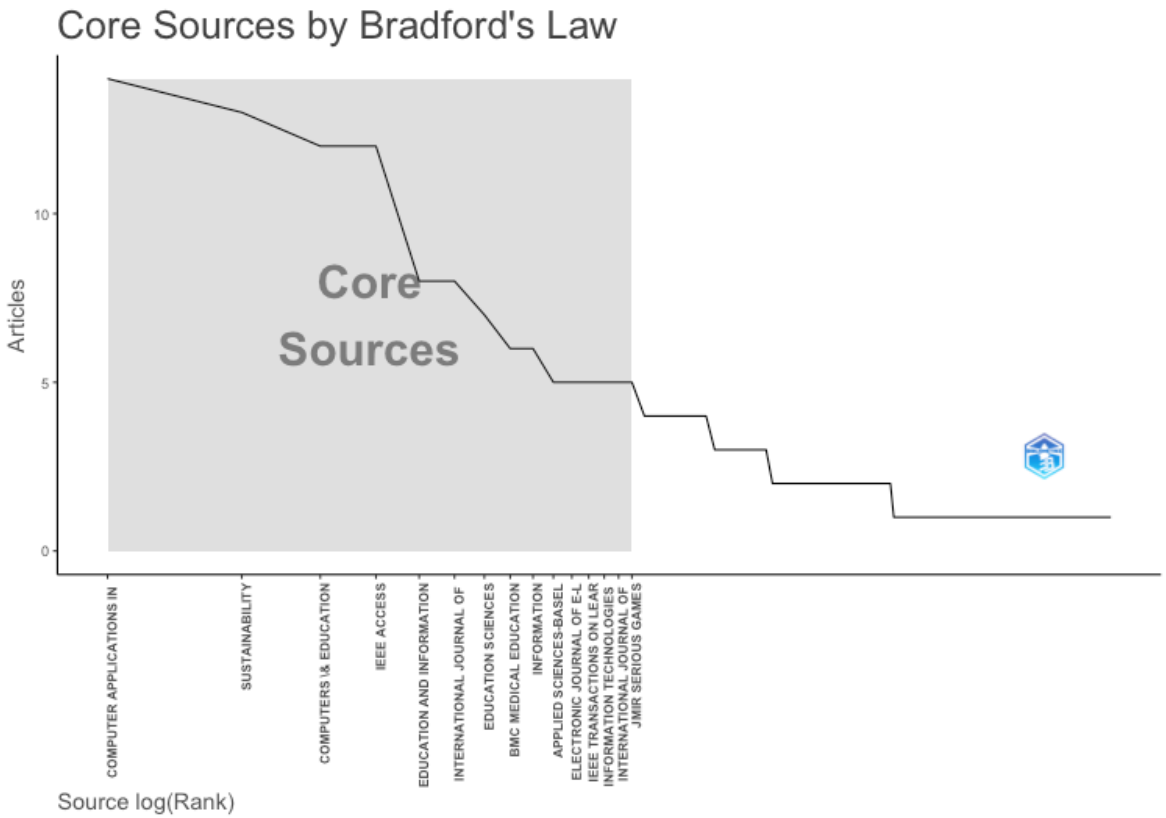


Figure 5. Core journals according to Bradford's Law Analysis

*Corresponding Author: gcetin@mu.edu.tr

ISSN: 2618-6055 / 9, (1), 159 – 173, 2025
DOI:10.62301/usmtd.1710461

As shown in Figure 5, the distribution is based on Bradford's Law, which is used in bibliometric analysis to identify the core journals within a research field. According to Bradford's Law, publications are divided into three zones based on their impact, with the first zone representing the "core sources." In this analysis, the journals located within the grey-shaded core zone are those with the highest scientific productivity.

Within this zone, journals such as Computer Applications in Engineering Education, Sustainability, Computers & Education, IEEE Access, and Education and Information Technologies have been identified as playing a central role in the knowledge production related to gamification.

When Figure 4 and Figure 5 are considered together, it becomes evident that research on gamification is concentrated across three main axes: engineering education, educational technologies, and multidisciplinary domains such as sustainability and information systems. This pattern highlights the interdisciplinary nature of the field and demonstrates how the topic of gamification in computer engineering education evolves through interaction with multiple knowledge domains.

Moreover, the inclusion of SSCI-indexed journals within this core cluster further strengthens the academic depth and visibility of the field in the broader scholarly literature.

3.4. Prolific Authors

As shown in Figure 6, based on Web of Science (WoS) data, the most prolific authors in the field of gamification in computer engineering education were analyzed.

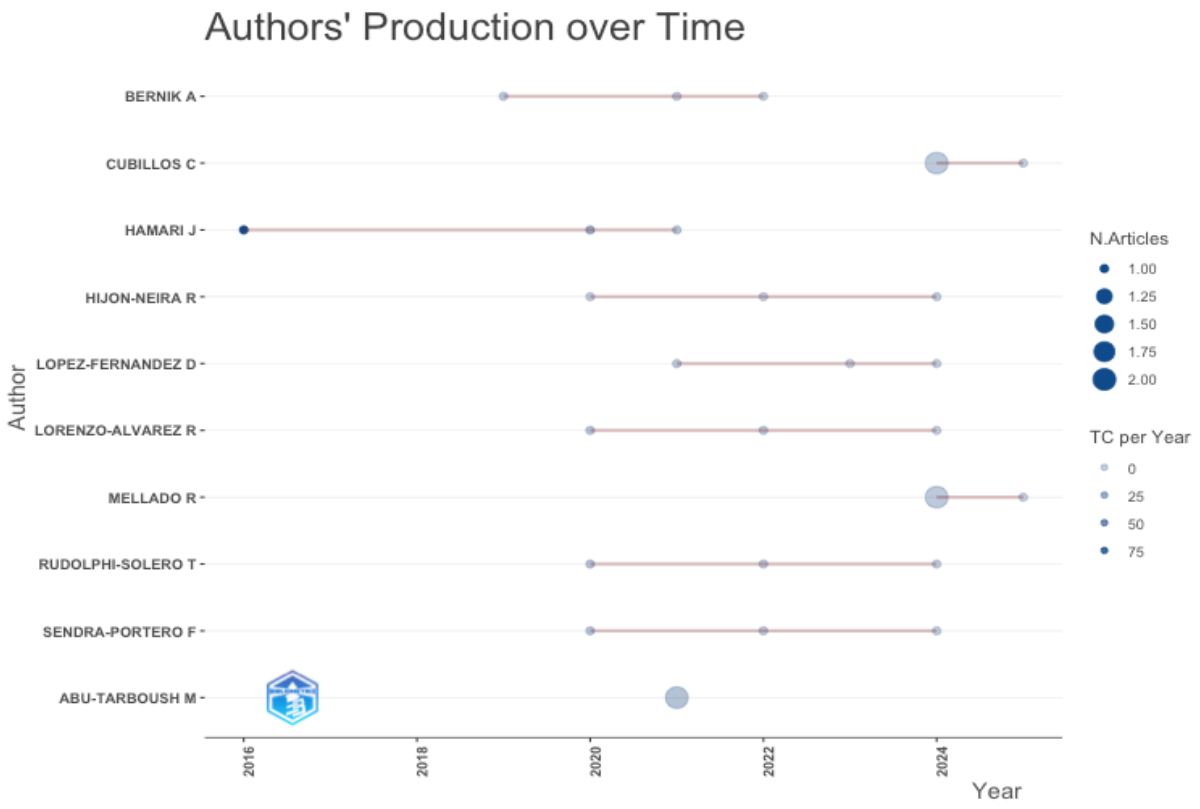


Figure 6. Most prolific authors in the field

According to the data presented in Figure 6, the most prolific authors in the domain of gamification in the context of computer engineering include Hamari J., Cubillos C., Mellado R., Hijon-Neira R., and Lopez-Fernandez D. The figure visualizes not only the number of publications but also the annual citation impact of each author's work. The size of the dots represents the total citations per year (TC/year) of the authors' publications.

Among these, Hamari J. stands out as the most influential and productive scholar, demonstrating consistent publication activity since 2016, along with the highest average annual citation rate (indicated

*Corresponding Author: gcetin@mu.edu.tr

ISSN: 2618-6055 / 9, (1), 159 – 173, 2025
DOI:10.62301/usmtd.1710461

by the largest dot). His work has laid the theoretical foundations of gamification and significantly influenced its application in educational technologies.
Cubillos C. and Mellado R. have emerged as highly productive authors in the recent period (2023–2024); however, their citation impact is currently lower than Hamari's, likely due to the recency of their publications. Authors such as Hijon-Neira R., Lorenzo-Alvarez R., Sendra-Portero F., and Rudolph-Solero T. exhibit a more stable, moderate level of productivity and citation impact.

This analysis reveals that academic output on gamification tends to cluster around certain key authors, who play a central role in shaping the field. Moreover, the continuity of scholarly production over time reflects the dynamic nature of the topic and the sustainability of academic interest. In particular, the works of highly cited authors serve as foundational references for future research by establishing the theoretical groundwork of the field.

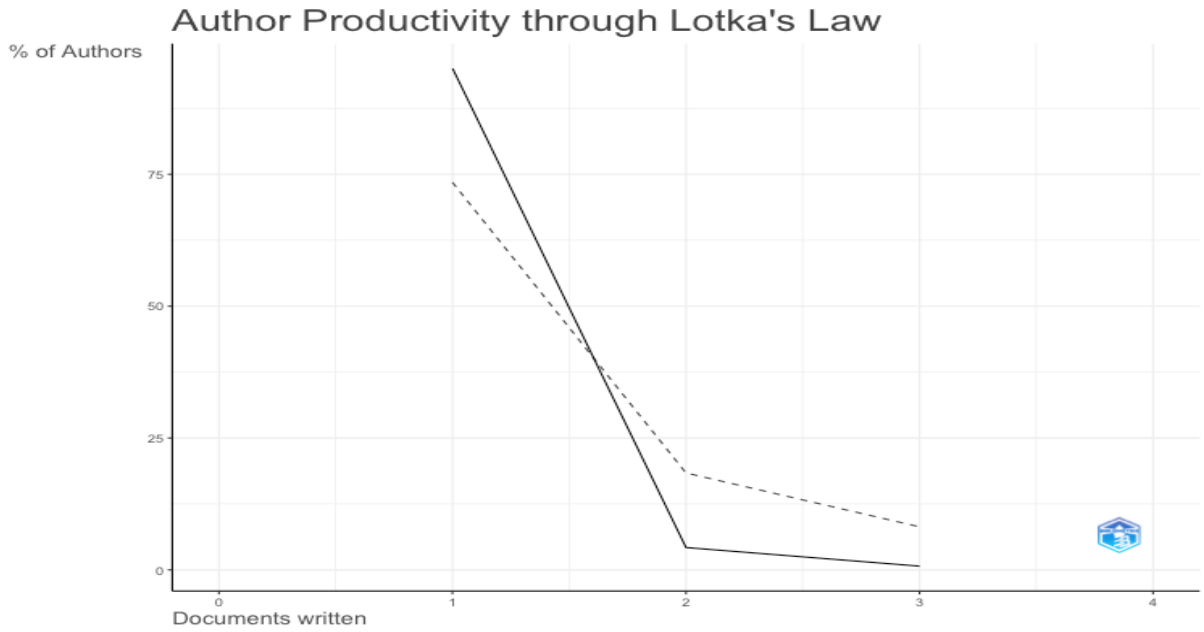


Figure 7. Author distribution according to Lotka's Law

Figure 7 illustrates the distribution of author productivity within the literature on gamification in computer engineering education, analyzed through the framework of Lotka's Law. According to Lotka's principle, the majority of authors in a scientific field produce only one publication, while a much smaller number are responsible for multiple publications. This distribution is commonly referred to in bibliometric literature as an inverse proportional productivity model.

In the graph, the solid line represents the observed distribution, while the dashed line indicates the theoretical distribution predicted by Lotka [25]. As seen in the analysis, approximately 80% of authors in the examined literature have published only one article. This indicates that the field includes a large number of one-time contributors, whereas the number of authors with continuous and high productivity remains quite limited. The proportion of authors with two or more publications falls below 20%.

This finding supports the earlier "Authors' Production over Time" analysis, which identified a few key authors—such as Hamari, Cubillos, and Mellado—as consistently productive, while most others have contributed only once. The Lotka distribution makes this structure more systematically visible.

This pattern suggests that the field is still in a developmental stage, with scholarly output concentrated around a limited number of authors. At the same time, it points to an expanding research landscape through the involvement of new scholars. To ensure sustainable knowledge production and to enhance the field's theoretical depth, greater co-authorship, interdisciplinary collaboration, and theme-focused research should be encouraged.

*Corresponding Author: gcetin@mu.edu.tr

ISSN: 2618-6055 / 9, (1), 159 – 173, 2025
DOI:10.62301/usmtd.1710461

3.5. Most Relevant Institutions

In terms of research on gamification in computer engineering education, the institutions with the highest number of publications indexed in the Web of Science (WoS) database were identified. The results are presented in Figure 8.

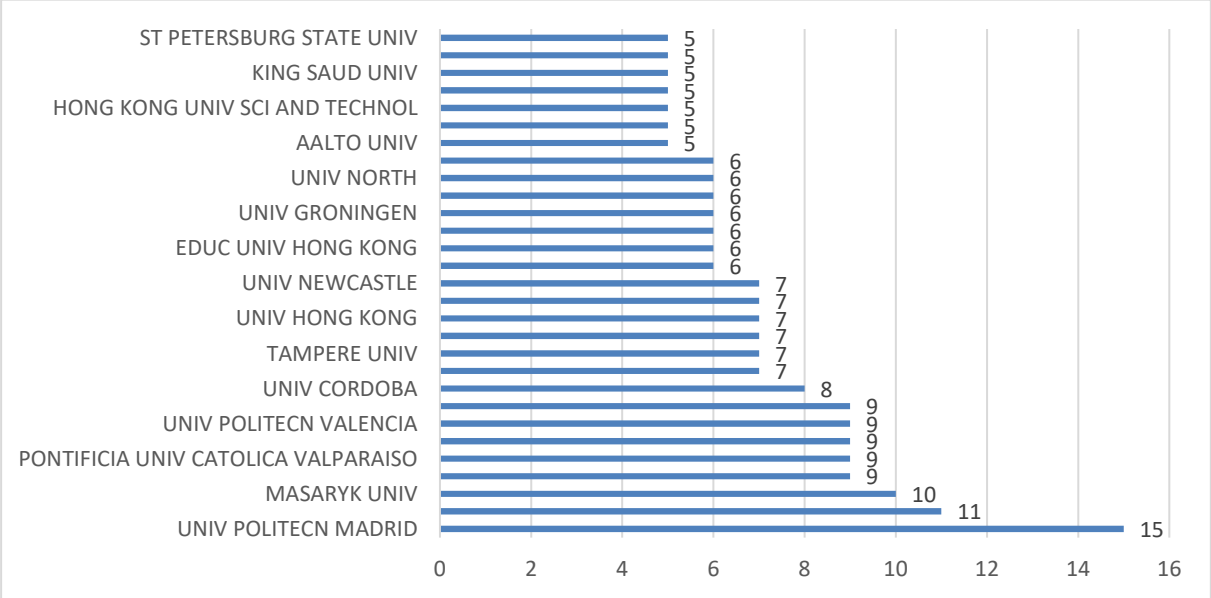


Figure 8. Top Institutions Publishing on Gamification in the Computer Engineering Education

Figure 8 visualizes the institutions that have made the most significant contributions to publications on gamification in computer engineering education. The institution with the highest contribution is Universidad Politécnica de Madrid (Spain), with 15 publications. This institution's leadership in the field of gamification underscores Spain's global influence in educational technology research. In second place is Universidad de Málaga (Spain), which has demonstrated notable productivity with 11 publications. The prominence of these two institutions highlights the significant role of Spain-based research networks and grant-supported projects in advancing gamification studies. They are followed by Masaryk University (Czech Republic, 10 publications) and, with nine publications each, Jilin University (China), Pontificia Universidad Católica de Valparaíso (Chile), Sichuan University (China), Universidad Politécnica de Valencia (Spain), Universidad Rey Juan Carlos (Spain), and Universidad de Córdoba (Spain). The presence of four different universities from Spain among the top ten further illustrates the country's strong interdisciplinary and international contributions to the field. Tabriz University of Medical Sciences (Iran) appears on the list with 7 publications, as the only medical university represented. This indicates that gamification is not limited to engineering and education domains but is also finding applications in diverse fields such as health education.

This analysis reveals that institutions from Spain, China, and Latin America are particularly prominent in the global gamification research landscape. Moreover, it demonstrates that Europe and Asia are strong actors in the race for leadership in this field. The concentration of institutional activity in specific countries also suggests a fertile ground for international collaboration.

3.6. Word Cloud Generated from Studies

Keywords often provide the first insight into the concepts and focus of a research study. Therefore, word cloud analysis holds an important place in bibliometric research. This method counts the frequency of keywords used within a body of literature, highlights the most frequently occurring terms, and generates a visual representation in the form of a word cloud [26].

Visually, the largest word in the center of the cloud typically represents the most frequently used keyword across the analyzed studies. Accordingly, Figure 5 illustrates the size and frequency of appearance of words or word phrases related to the theme of gamification in computer engineering education, using varying colors to distinguish different frequency levels.

*Corresponding Author: gcetin@mu.edu.tr

ISSN: 2618-6055 / 9, (1), 159 – 173, 2025
DOI:10.62301/usmtd.1710461

Figure 9. Word cloud of gamification-themed publications

According to the analysis results presented in Figure 9 and Figure 10, the most frequently emphasized key concepts in gamification-themed publications in computer engineering education have been identified. Among the most commonly used keywords are “computer science” and “game-based learning”, each appearing 16 times. This indicates that the studies predominantly focus on the application of gamification strategies within ICT-related disciplines, particularly in the field of computer science.

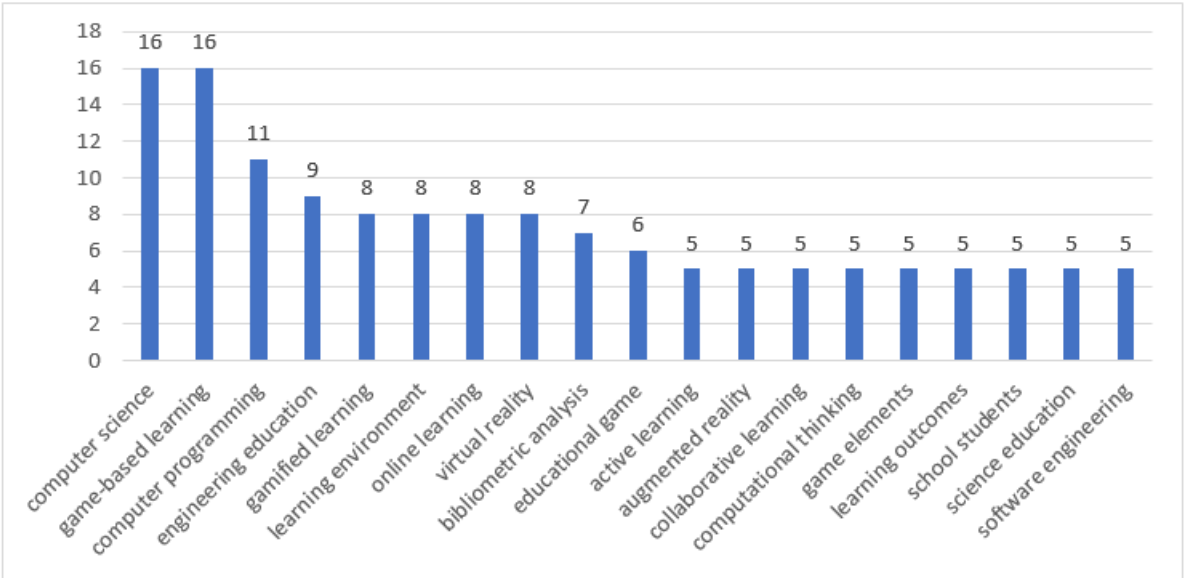


Figure 10. Keyword Frequency Table for Gamification-Themed Publications

Among the high-frequency terms following these top keywords are “computer programming” (11 occurrences) and “engineering education” (9 occurrences). This suggests that gamification is positioned not only as a pedagogical approach, but also as a didactic tool that effectively supports the teaching of programming and technical content. Additionally, terms such as “gamified learning”, “online learning”, “virtual reality”, “learning environment”, and “bibliometric analysis”, each appearing 7–8 times, reflect a strong focus on technological integration and learning environment design as core themes in the literature. Mid-frequency terms (5–6 occurrences), including “educational game”, “active learning”,

*Corresponding Author: gcetin@mu.edu.tr

ISSN: 2618-6055 / 9, (1), 159 – 173, 2025
DOI:10.62301/usmtd.1710461

"computational thinking", "collaborative learning", and "learning outcomes", point to a learner-centered perspective, with emphasis on the cognitive and interactive benefits of gamification. This suggests that gamification is closely associated not only with pedagogical advantages but also with cognitive and social development. Less frequent but thematically important keywords—such as "achievement badges", "gamification framework", "academic performance", "medical students", and "chemical design"—demonstrate the breadth of the field. These terms reflect connections to both application contexts and assessment methodologies, indicating that gamification is an interdisciplinary approach applied not only in computer engineering but also in medicine, chemistry, and general science education.

Overall, the keyword distribution reveals that gamification research is clustered around themes such as the teaching of technical content, the design of digital learning environments, and the transformation of learner behaviors. This demonstrates the theoretical and practical potential of gamification in both educational technology and engineering education literature.

3.7. Research Trends

Using bibliometric analysis, it is possible to identify research trends based on authorship, citations, journals, and subtopics, and to monitor the evolution of scholarly interest (Secinaro et al., 2020). Accordingly, the trridetyoends in gamification-themed studies in computer engineering education are illustrated in Figure 11.

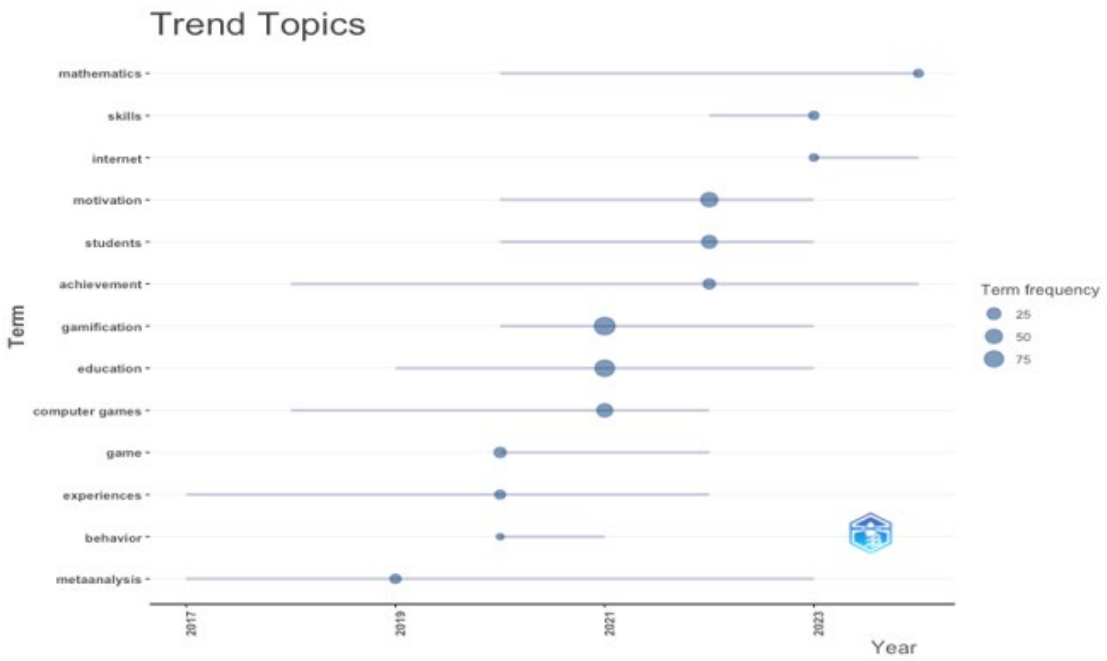


Figure 11. Trend Analysis: Temporal distribution of keywords (2017–2024)

Figure 11 presents a trend analysis showing the emergence and prominence of thematic keywords in publications on gamification in computer engineering education from 2017 to 2024. The size of each term represents its frequency of occurrence in the literature.

Early Period Themes (2017–2019):

- During this phase, dominant keywords included "meta-analysis", "experiences", "behavior", "game", and "computer games."
- These findings suggest that early studies on gamification primarily focused on theoretical evaluations, learning experiences, and the behavioral impact of games.
- The literature also explored how gamification tools relate to game-based pedagogical models.

Period of Intensification and Pedagogical Focus (2020–2022):

*Corresponding Author: gcetin@mu.edu.tr

ISSN: 2618-6055 / 9, (1), 159 – 173, 2025
DOI:10.62301/usmtd.1710461

- After 2020, the most frequent keywords in the visualization were “gamification”, “education”, “motivation”, and “students.”
- This reflects the surge in interest in digital education during the COVID-19 pandemic, where gamification was increasingly discussed in relation to enhancing student engagement, fostering motivation, and facilitating interaction in remote learning environments.
- Pedagogically oriented, applied research appears to have become more prominent during this period.

Emerging Themes (2023–2024):

- In 2023 and beyond, keywords such as “skills”, “internet”, and “mathematics” have gained visibility.
- This shift indicates that gamification is expanding into skills-based education (particularly digital skills), internet-supported learning environments, and interdisciplinary applications, especially in mathematics education.
- The increasing integration of gamification into STEM education and sustainable digital pedagogies is also suggested.

The time-series trend reveals a developmental trajectory: beginning with theoretical and experiential focuses, moving toward pedagogical applicability, and finally diversifying into skill development and digital learning integration. Keywords such as “motivation,” “education,” and “students” remain central to the discourse, while newer themes like “mathematics,” “internet,” and “skills” show strong potential for shaping future research directions in this domain.

4. Results and Discussion

This study employed a bibliometric approach to examine the international literature on gamification in computer engineering education, aiming to uncover the structural characteristics, research trends, thematic focuses, and intellectual production centers of the field. Based on data from the Web of Science database, a total of 343 articles were analyzed, and the implementation of gamification in engineering education was evaluated through a multidimensional lens.

The findings reveal that gamification has gained significant academic momentum, particularly after 2018, with the number of publications peaking in 2021. This surge appears to be closely related to the increased interest in digital learning environments during the COVID-19 pandemic [28]. The fact that the most cited studies were published between 2013 and 2015 suggests that the theoretical foundations of the field were laid during this period and have guided subsequent research.

Prominent authors such as Hamari, Cubillos, Mellado, and Hijon-Neira were found to exhibit high levels of productivity and scholarly impact. However, in line with Lotka’s Law, the vast majority of authors have contributed with only a single publication, indicating that the field is still in a developmental stage. This highlights the need to support consistently productive research groups and foster interdisciplinary collaborations.

Journal analysis results show that studies on gamification are most frequently published in journals such as Computer Applications in Engineering Education, Sustainability, and Computers & Education, which focus on both engineering and educational technologies. According to Bradford’s Law, these journals constitute the core knowledge sources of the field and play a central role in the dissemination of research on gamification.

According to keyword and thematic analyses, terms such as “game-based learning”, “computer science”, “engineering education”, and “virtual reality” are among the most frequently repeated and form the thematic core of the literature. Time-series analysis indicates that from 2017 to 2019, research on gamification primarily focused on theoretical and behavioral themes, while the post-2020 period saw a shift towards pedagogical applications, student motivation, and content integrated with digital environments. These findings are supported by previous studies [29, 30].

*Corresponding Author: gcetin@mu.edu.tr

ISSN: 2618-6055 / 9, (1), 159 – 173, 2025
DOI:10.62301/usmtd.1710461

In light of these results, there is a clear need for more in-depth qualitative analyses, interdisciplinary collaborations, and application-oriented research in the field of gamification. Future studies are encouraged to move beyond the mere use of game elements and instead focus on models that holistically address learners' cognitive, affective, and social development. Additionally, further exploration into sustainable digital pedagogies, the role of gamification in online learning environments, and the use of learning analytics is recommended.

Overall, this study provides researchers with a structural roadmap and increases the visibility of knowledge clusters within the literature, firmly establishing gamification in computer engineering education as a growing and dynamic field of inquiry. Moreover, this study identifies emerging themes such as interdisciplinary applications and digital skill development.

5. References

[1] S. Deterding, D. Dixon, R. Khaled, L.Nacke From game design elements to gamefulness: Defining gamification, Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments (2011) 9-15.

[2] K. M. Kapp, The gamification of learning and instruction: Game-based methods and strategies for training and education, Pfeiffer (2012).

[3] R. Damaševičius, R. Maskeliūnas, T. Blažauskas, Serious games and gamification in healthcare: A meta-review, Information, 14 (2) (2023).

[4] S. Mohanty, B. P. Christopher, The role of gamification research in human resource management: A PRISMA analysis and future research direction, SAGE Open, 14(2) (2024).

[5] G. Malik, D. Pradhan, B. K. Rup, Gamification and customer brand engagement: A review and future research agendas, Marketing Intelligence & Planning, 43(1) (2025) 210-239.

[6] O. Inbar, N. Tractinsky, O. Tsimhoni, T. Seder, Driving the scoreboard: Motivating eco-driving through in-car gaming, Workshop Gamification: Using Game Design Elements in Non-Game Contexts, (2011) 7-12.

[7] F. Cassano, A. Piccinno, T. Roselli, V. Rossano, Gamification and learning analytics to improve engagement in university courses, Methodologies and Intelligent Systems for Technology Enhanced Learning, 8 (2019) 156-163.

[8] K. Seaborn, D. I. Felds, Gamification in theory and action: A survey, International Journal of Human-Computer Studies, 74 (2015) 14–31. Gamifying education: What is known, what is believed and what remains uncertain: A critical review, International Journal of Educational Technology in Higher Education, 14(9) (2017).

[9] C. Dichev, D. Dicheva, Gamifying education: What is known, what is believed and what remains uncertain: A critical review, International Journal of Educational Technology in Higher Education, 14 (9) (2017).

[10] M. Kalogiannakis, S. Papadakis, A. Zourmpakis, Gamification in science education: A systematic review of the literature, Education Sciences, 11(1) (2021).

[11] G. Lampropoulos, Kinshuk, Virtual reality and gamification in education: A systematic review, Educational Technology Research and Development, (2024) 1-95.

[12] S. Tonhão, R. R. da Silva, T. U. Conte, Gamification in software engineering education: A tertiary study, SBES 2023: Brazilian Symposium on Software Engineering, (2023) 202–211.

[13] N. Zeybek, E. Saygı, Gamification in education: Why, where, when, and how?—A systematic review. Games and Culture, 19(2) (2024) 237–264.

[14] M. Aria, C. Cuccurullo, Bibliometrix: An R-tool for comprehensive science mapping analysis, Journal of Informetrics, 11(4) (2017) 959–975.

[15] I. Zupic, T. Čater, Bibliometric methods in management and organization, Organizational Research Methods, 18(3) (2015) 429–472.

*Corresponding Author: gcetin@mu.edu.tr

ISSN: 2618-6055 / 9, (1), 159 – 173, 2025
DOI:10.62301/usmtd.1710461

[16] J. Zhu, W. Liu, A tale of two databases: the use of web of science and Scopus in academic papers, Scientometrics, 123 (2020) 321-335.

[17] P. Mongeon, A. Paul-Hus, The journal coverage of Web of Science and Scopus: a comparative analysis. Scientometrics, 106 (2016) 213-228.

[18] M. Kumpulainen, M. Seppänen, Combining Web of Science and Scopus datasets in citation-based literature study. Scientometrics, 127(10) (2022) 5613-5631.

[19] G. Spinaci, G. Colavizza, S. Peroni, A map of Digital Humanities research across bibliographic data sources, Digital Scholarship in the Humanities, 37(4) (2022) 1254-1268.

[20] A. Domínguez, J. Saenz-de-Navarrete, L. De-Marcos, L. Fernández-Sanz, C. Pagés, J. J. Martínez- Herráiz, Gamifying learning experiences: Practical implications and outcomes. Computers & Education, 63 (2013) 380-392.

[21] J. Hamari, J. Koivisto, H. Sarsa, Does gamification work? – A literature review of empirical studies on gamification, Proceedings of the 47th Hawaii International Conference on System Sciences, (2014) 3025–3034.

[22] A. V. Banerjee, S. Cole, E. Duflo, L. Linden, Remedying education: Evidence from two randomized experiments in India. The quarterly journal of economics, 122(3) (2007) 1235-1264.

[23] T. M. Fleming, L. Bavin, K. Stasiak, E. Hermansson-Webb, S. N. Merry, S. N., C. Cheek, S. Hetrick, Serious games and gamification for mental health: current status and promising directions, Frontiers in Psychiatry, 7 (2017) 215.

[24] Y. Attali, M. Arieli-Attali, Gamification in assessment: Do points affect test performance?, Computers & Education, 83 (2015) 57-63.

[25] A. J. Lotka, The frequency distribution of scientific productivity. Journal of the Washington Academy of Sciences, 16 (1926) 317–323.

[26] W. J. Lee, A study on word cloud techniques for analysis of unstructured text data, The Journal of the Convergence on Culture Technology, 6(4) (2020) 715–720.

[27] M. M. Alhammad, P.M. Moreno-Marcos, Approaches and game elements used to tailor digital gamification for learning: A systematic literature review. Computers & Education, 208 (2024) 104861.

[28] E. G. Rincón-Flores, B. N. Santos-Guevara, Gamification during Covid-19: Promoting active learning and motivation in higher education, Australasian Journal of Educational Technology, 37(5) (2021) 43–59.

[29] S. Secinaro, V. Brescia, D. Calandra, P. Biancone, Employing bibliometric analysis to identify suitable business models for electric cars, Journal of Cleaner Production, 264 (2020) 121503.

[30] M. Videnovik, T. Vold, L. Kiønig, A. Madevska Bogdanova, V. Trajkovik, Game-based learning in computer science education: A scoping literature review. International Journal of STEM Education, 10(1) (2023) 54.

*Corresponding Author: gcetin@mu.edu.tr