BİNGÖL ÜNİVERSİTESİ

BİNGÖL ÜNİVERSİTESİ TEKNİK BİLİMLER DERGİSİ

Bingol University Journal of Technical Science

Cilt 6 • Sayı 1 • Haziran 2025
Volume 6 • Number 1 • June 2025

Bingöl Üniversitesi Teknik Bilimler Meslek Yüksekokulu tarafından yayınlanmaktadır

e-ISSN: 2757-6884

Bingöl Üniversitesi Teknik Bilimler Dergisi
Bingol University Journal of Technical Science

e-ISSN: 2757-6884 ***25.06.2025***

EDİTÖRÜN NOTU

Ülkemizde bilimsel yayıncılık hızla gelişmekte ve bu bağlamda süreli yayınların sayısı önemli ölçüde artmaktadır. Akademik süreli yayıncılık da bu artışın doğal sonuçlarındandır. 8 Eylül 2020 tarihinde yayınlanan ilk sayısı ile yayın hayatına başlayan dergimizle (***Bingöl Üniversitesi Teknik Bilimler Dergisi***) çok değerli araştırmacıların, bilim insanlarının ve okurların karşısına çıkmanın heyecanını ve mutluluğunu yaşamaktayız. Yayın hayatına başladığı bu tarihten itibaren bilimsel bir disiplin içerisinde hareket eden dergimiz, Google Scholar, Index Copernicus ve AcarIndex üzerinde taranmakta ve diğer indekslerde taranmak için başvurularını sürdürmektedir.

Türkiye'de yayın yapan birçok üniversite akademik dergileri gibi dergimiz de çok-disiplinli ve disiplinlerarası anlayışla hareket etmektedir. Bu anlayışla dergimizin yayın kurulu, bilimin bütün sahalarından ve alt disiplinlerinden bilimsel nitelikli yazıları İngilizce ve/veya Türkçe olarak yayınlamak üzere her iki dilde de kabul etmektedir. Dergimizde hakemlik süreci titizlikle yürütülmekte, çift taraflı körleme sistemiyle makaleler değerlendirilmekte, etik ve bilimsel ölçütlere sonuna kadar bağlı kalınmaktadır.

İlk sayıdan itibaren dergimizin **DergiPark** üzerinden erişimi sağlanmış ve yayınlanan makalelerin tamamı okuyucuların ve araştırmacıların hizmetine sunulmuştur. Dergimizin bu sayısında 3 adet bilimsel araştırma ve 1 adet derleme makalesine yer verilmiştir.

Dergimize bilimsel araştırmaları ve yazılarıyla destek veren değerli bilim insanlarına, bu çalışmaları titizlikle değerlendiren hakemlere ve yayın sürecini yöneten ve yürüten yayın kurulu, alan editörleri ve sekretaryaya teşekkür ederim.

Baş Editör
Dr. Öğr. Üyesi Uğur Eren YURTCAN
(Bingöl Üniversitesi Teknik Bilimler MYO)

Bingöl Üniversitesi Teknik Bilimler Dergisi
Bingol University Journal of Technical Science

e-ISSN: 2757-6884 ***25.06.2025***

Alan Editörleri / *Section Editors*

Prof. Dr. İhsan	Kırık	Makine Müh.	Bingöl Üniversitesi
Prof. Dr. Hasan	Oğul	Nükleer Enerji Müh.	Sinop Üniversitesi
Doç. Dr. Serhat	Şap	Makine Müh.	Bingöl Üniversitesi
Doç. Dr. Ünal	Değirmenci	Makine Müh.	İnönü Üniversitesi
Doç. Dr. Anıl	İmak	Makine Müh.	Bingöl Üniversitesi
Doç. Dr. Burak	Yıldırım	Elektrik Elektronik Müh.	Bingöl Üniversitesi
Doç. Dr. Ahmet	Turşucu	Elek. Elektr.Müh./ Elektrik Mak.	Şırnak Üniversitesi
Doç. Dr. Bilal	Tütüncü	Elek. Elektr.Müh./Haberl.	İstanbul Teknik Üniversitesi
Doç. Dr. Murat	Dener	İnşaat Müh./ Yapı Malzemesi	Bingöl Üniversitesi
Doç. Dr. Hasan	Polat	Elektrik Elektronik Müh.	Bingöl Üniversitesi
Doç. Dr. Mehmet	Polat	Mekatronik Müh.	Fırat Üniversitesi
Dr. Müslüm	Erol	Tekstil Müh.	Bingöl Üniversitesi
Dr. Müge	Yurtcan	Peyzaj Mim./ Kentsel Tasarm	Bingöl Üniversitesi
Dr. M. Zekeriya	Gündüz	Bilgisayar Müh.	Bingöl Üniversitesi
Dr. İbrahim	Çelik	Elek. Elektr.Müh.(Yenil. Ener.)	İstiklal Üniversitesi
Dr. Uğur Eren	Yurtcan	İnşaat Müh./ Geoteknik	Bingöl Üniversitesi
Dr. Muhammed	Mahmudi	İnşaat Müh./ Geoteknik	Ege Üniversitesi
Dr. Ömer Faruk	Nemutlu	İnşaat Müh./ Yapı	Bingöl Üniversitesi
Dr. Hasan	Polat	İnşaat Müh./ Yapı Malzemesi	Bingöl Üniversitesi
Dr. Mehmet Nuri	Kolak	İnşaat Müh./ Yapı Malzemesi	Bingöl Üniversitesi
Dr. Mehmet Onur	Karaagac	Enerji Sistemleri Müh.	Sinop Üniversitesi
Dr. Hakan	Us	Nükleer Enerji Müh.	Sinop Üniversitesi
Dr. Hüsnü	Aydemir	Tekstil Müh.	Bingöl Üniversitesi
Dr. Serkan	Kaya	Maden Müh.	Karadeniz Teknik Üniv.
Dr. Mitat	Öztürk	İnşaat Müh./ Geoteknik	Osmaniye Üniversitesi
Dr. Yakup	Önal	İnşaat Müh./ Ulaştırma, Geoteknik	Osmaniye Üniversitesi
Dr. Abdulkadir	Budak	İnşaat Müh./ Yapı İşletmesi	Osmaniye Üniversitesi
Dr. Kenan	Akbayram	Jeoloji	Bingöl Üniversitesi
Dr. Sadık	Varolgüneş	İnşaat Müh./ Yapı	Bingöl Üniversitesi
Dr. Gökhan	Yücel	İnşaat Müh./ Yapı	Osmaniye Üniversitesi
Dr. Emriye	Çınar Resuloğulları	İnşaat Müh./ Yapı Malzemesi	Osmaniye Üniversitesi
Dr. Şevket	Bediroğlu	Harita Mühendisliği	Gaziantep Üniversitesi
Dr. Muhammad Rehan	Hakro	İnşaat Müh./ Geoteknik	Mehran Üniversitesi
Dr. Muhammet	Aydın	Mekatronik Müh.	Fırat Üniversitesi
Dr. İbrahim	Karataş	İnşaat Müh. / Yapı İşl.	Osmaniye Üniversitesi
Dr. Gökçe	Yıldırım	Elek. Elektr.Müh./Der. Öğr.	Bingöl Üniversitesi
Ebubekir	Bozkurt	Makine Müh.	Bingöl Üniversitesi
Ayşegül	Gümrükçü Keskin	Grafik Tasarım	Bingöl Üniversitesi
Gamze	Bediroğlu	Harita Mühendisliği	Kilis 7 Aralık Üniv.

Yayın Danışma Kurulu / *Editorial Advisory Board*

BİNGÖL ÜNİVERSİTESİ
TEKNİK BİLİMLER DERGİSİ

Bingöl Üniversitesi Teknik Bilimler Dergisi
Bingol University Journal of Technical Science
e-ISSN: 2757-6884

Cilt:6, Sayı: 1, 25.06.2025 *Volume:6, Number: 1. 25.06.2025*

İÇİNDEKİLER/CONTENTS

BİNGÖL ÜNİVERSİTESİ
TEKNİK BİLİMLER DERGİSİ

Bingöl Üniversitesi Teknik Bilimler Dergisi
Bingol University Journal of Technical Science
e-ISSN: 2757-6884

Cilt:6, Sayı: 1, 25.06.2025 *Volume:6, Number: 1. 25.06.2025*

Weight Optimization of Weighted Naive Bayes Classifier for Efficient Classification

Gamzepelin AKSOY[1]*, Murat KARABATAK [2]

[1] Isparta University of Applied Sciences, Department of Computer Engineering, Isparta/Türkiye.
ORCID: 0000-0002-5328-2983, E-mail: gamzepelinaksoy@isparta.edu.tr

[2] Fırat University, Department of Software Engineering, Elazığ/Türkiye.
ORCID: 0000-0002-6719-7421, E-mail: mkarabatak@firat.edu.tr

(Alınış/Arrival: 02.06.2025, Kabul/Acceptance: 13.06.2025, Yayınlanma/Published: 25.06.2025)

BİNGÖL ÜNİVERSİTESİ
TEKNİK BİLİMLER DERGİSİ

Bingöl Üniversitesi Teknik Bilimler Dergisi
Bingol University Journal of Technical Science
e-ISSN: 2757-6884

Cilt:6, Sayı: 1, Sayfa: 1-12, 2025
Derleme Makalesi

Volume:6, Number: 1, Page: 1-12, 2025
Review Article

Doi: 10.5281/zenodo.15719101

Derleme: Memristör: Cihaz, Model ve Uygulamaları

Özlem AKIN [1]*

[1]Erzurum Teknik Üniversitesi, Elektrik Elektronik Mühendisliği, Erzurum, Türkiye.

ORCID:0000-0002-8636-5021, E-mail: ozlem.akin@erzurum.edu.tr

(Alınış/Arrival: 25.03.2025, Kabul/Acceptance: 10.04.2025, Yayınlanma/Published: 25.06.2025)

Özet

Bu makale 1971 yılında Leon Chua tarafından keşfedilen eksik devre elemanı olarak adlandırılan memristör yapısını tanıtmaktadır. Memristör nöromorfik sinir ağlarında, ileri düzey bilgi işleme teknolojileri başta olmak üzere dijital ve analog devre uygulamalarda kullanılan ve sahip olduğu birçok özellik ile son teknolojik gelişmelere entegre olabilen bir devre elemanıdır. Metal-yalıtkan-metal, yapılar birçok metal oksit kullanılarak farklı özelliklerde memristör yapıları üretilmesine izin vermektedir. Bu memristör yapıları kullanılarak teknolojik gelişmeler her gün iyileşerek gelişmeye devam etmektedir. Makalede memristör yapısının çalışma prensibi üzerinde durularak yapının daha iyi anlaşılması sağlanmıştır. Ayrıca memristör yapılarının genel olarak deneysel üretiminin nasıl olduğunu ve nasıl çalıştığını, sandviç yapısı ele alınarak anlatmıştır. Memristörün anahtarlama mekanizmalarının neler olduğu ve nasıl çalıştığı konusunda bilgi verilmiştir. Son olarak günümüzde memristör uygulamalarının neler olduğu özet olarak anlatılmıştır.

Anahtar Kelimeler: Memristör, Anahtarlama mekanizmaları, Çalışma prensibi, Memristörün uygulama alanları.

Review: Memristor: Device, Model and Applications

Abstract

This article introduces the memristor structure, which is called the missing circuit element, discovered by Leon Chua in 1971. Memristors are circuit elements that are used in neuromorphic neural networks, advanced information processing technologies, and digital and analog circuit applications, and are integrated with the latest technological developments with their many features. Metal-insulator-metal structures allow the production of memristor structures with different properties using many metal oxides. Technological developments using these memristor structures continue to improve and develop every day. In the article, the working principle of the memristor structure was emphasized and a better understanding of the structure was provided. In addition, it explains how the experimental production of memristor structures is generally done and how they work by considering the sandwich structure.

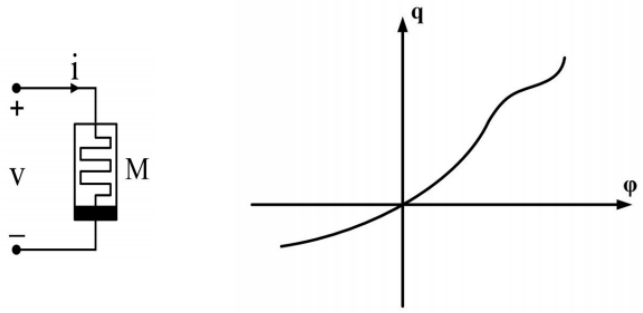
Keywords: Memristor, Switching mechanisms, Operating principle, Application areas of memristor.

1. GİRİŞ

Elektronik alanındaki gelişmelere her geçen gün bir yenisi eklenmektedir. Artan bellek ve depolama ihtiyacı ve bunun yanında işletim sistemlerinin daha hızlı olması mevcut mimariler

ile ilerlemekte güçlük çekmektedir. Von Neumann mimarisi ile yapılan işlemlerde depolama ve bilgi işleme süreci, ciddi problemlere sahiptir [1]. Bu problemlerden ilki erişim hızının merkezi işlem biriminin hesaplama hızından çok daha yavaş olması ve buna bağlı olarak bellek duvarının kullanımının azalmasıdır [2,3]. İkinci temel problem ise iletimi esnasında tüketilen güç miktarında ciddi bir güç kaybına sebep olduğu için güç duvarı olarak adlandırılan çipin gelişmesi oldukça yavaşlamıştır [4]. Moore yasasındaki gelişmeler von Neumann mimarisinin işleyişini iyileştirmiş olsa da bu çözüm gelişen teknoloji karşısında verileri depolama ve hesaplamada henüz yeterli boyuta ulaşmamıştır. Araştırmacılar üretilen işaretlere karşı verilen cevapların beyin sinir hücreleri gibi hızlı olmasını isteyerek, yapay sinir hücreleri gibi çalışan sistemlere yönelmişlerdedir [5].

Memristor yapısı değişen hafızaya sahip bir direnç elemanıdır. Bu yapı ilk olarak 1971 yılında Leon Chua ve öğrencisi tarafından keşfedilmiştir [6]. Leon Chua o dönemde mevcut devre elemanları ve aralarındaki ikili ilişkileri inceleyerek bu ikili ilişkilerden birinin eksik olduğunu bulmuştur. Leon Chua akım ve gerilim arasındaki ilişkiden direnç elemanı, akım ve yük arasındaki ilişkiden endüktans elemanı ve gerilim ile yük arasındaki ilişkiden ise kapasite elemanının üretildiğini fakat akı ile yük arasındaki ilişki kullanılarak herhangi bir eleman geliştirmediğini fark etmiştir. Böylece akı ile yük arasındaki bu ilişkiden hafızalı bir direnç olan memristör yapısını üretmiştir. Memristör kayıp devre elemanı olarak bilinen ve diğer devre elemanları gibi pasif olarak çalışan bir elemandır. Ayrıca memristör iki terminallidir ve non-lineer olarak çalışır. Şekil 1'de Leon Cuha tarafından üretilen memristörün şematik görüntüsü ve yine Chua tarafından eksikliği fark edilen akı ve yük arasındaki ilişkinin grafiği verilmiştir.

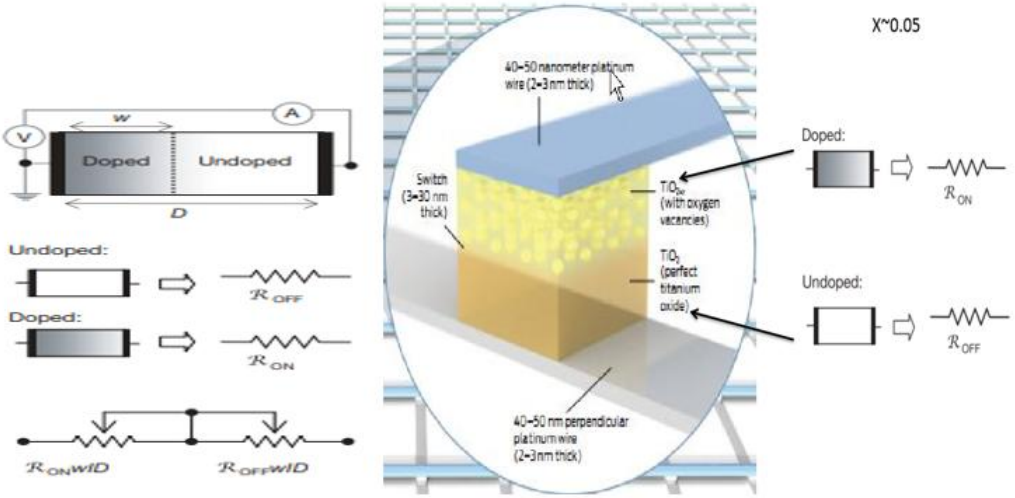


Şekil 1. Memristör şematik gösterimi ve yük ile manyetik akı arasındaki tipik eğri [6].

Memristörün çalışma prensibi şöyledir: üzerinden akım geçince direnci değişen ve akım kesilince son direnç değerini muhafaza eden bir yapıdır. Bu durumu üretilen malzemenin içerisindeki oksijen boşluklarının yapı içerisindeki hareketleri ile sağlamaktadır. 1971 yılından 2008 yılına kadar akıllı dirençler üzerine birçok çalışma yapılmış olmasına rağmen memristördeki direnç değişiminin asıl sebebi anlaşılamayıp çalışılan konular açığa kavuşturulamamıştır.

2008 yılında HP laboratuvarında ki çalışma ekibi, bulmuş oldukları sonuçların [7] 1971 yılında Leon Chua'nın çalışmasında bahsettiği memristör yapısının davranışını sergilediğini ortaya koymuştur. Daha sonra bu alandaki gizem kısmen çözülüp araştırmacılar bu alanda çalışmaya yönelmişlerdir. 2008 yılında yapılan bu çalışmada Strukov ve ekibi, Chua'nın memristör yapısına uygun olarak çalışan iki uçlu elektriksel bir cihazın fiziksel çalışma modelini açıklamayı başarmışlardır. Bu model Leon Chua'dan beri yapılan ve sebebi tam olarak açıklanamayan nano ölçekli ince film aygıtlarda görülen anahtarlama, çok durumlu iletkenlik, negatif diferansiyel ve direnç histerisiz gibi durumlarda ortaya çıkan akım-gerilim anomalilerinin sebebini açıklamayı başarmıştır [7].

Strukov ve ekibi ürettikleri modelde iki metal kontak arasında D uzunluğunda katkılı ve katkısız iki bölgeye sahip bir yarı iletkenden oluşan memristör yapısı gerçekleştirdiler, Şekil 2. Aygıtın toplam direnci birbirine seri bağlanmış katkılı ve katkısız iki bölgenin dirençleri toplamına eşittir. Katkılı bölge yüksek dirençli bölgeyi temsil edip ROFF ile gösterilirken, katkısız bölge daha düşük dirençli olup RON ile gösterilir. Bu bölgelerin genişliği metal oksit içerisindeki oksijen boşluklarının uygulanan gerilim ile sürüklenerek yer değiştirmesi ile değişmektedir. Böylece katkılı ve katkısız bölgelerin sınırları değişir. Bu da eşdeğer direncin değişmesine sebep olur. Yapı üzerine gerilim uygulandığında oksijen boşlukları katkısız bölgeye doğru hareket edip burada birikerek bu bölgenin direnç değerini azaltıp akım iletiminin daha kolay olmasını sağlayarak devreyi AÇIK (RON) durumuna geçirmektedir. Uygulanan gerilimin yönü değiştirildiğinde ise katkısız bölgenin genişliği artırıldığından yapının direnç değeri yükseltilerek yapı KAPALI (ROFF) konumuna geçmektedir. Bu çalışma prensibinin Leon Chua'nın hafızalı devre elemanına uygun olduğu gösterilip memristör yapısına atıf yapılarak 2008 yılında memristörün ne olduğu detaylı olarak anlatılmıştır.



Şekil 2. Memristör için değişken direnç çifti modeli [7].

Strukov ve arkadaşlarının ürettikleri memristör yapısının matematiksel ifadesi aşağıdaki gibidir. Ortalama iyon mobilitesi μ_v olan uniform bir alanda, doğrusal iyonik sürüklenme ve omik elektronik iletkenlik için;

$$v(t) = (R_{ON}\frac{w(t)}{D} + R_{OFF}(1 - \frac{w(t)}{D}) \quad (1)$$

$$\frac{dw(t)}{dt} = \mu_v \frac{R_{ON}}{D} i(t) \quad (2)$$

Denklemleri ile verilir, denklem 2'den w(t);

$$w(t) = \mu_v \frac{R_{ON}}{D} q(t) \quad (3)$$

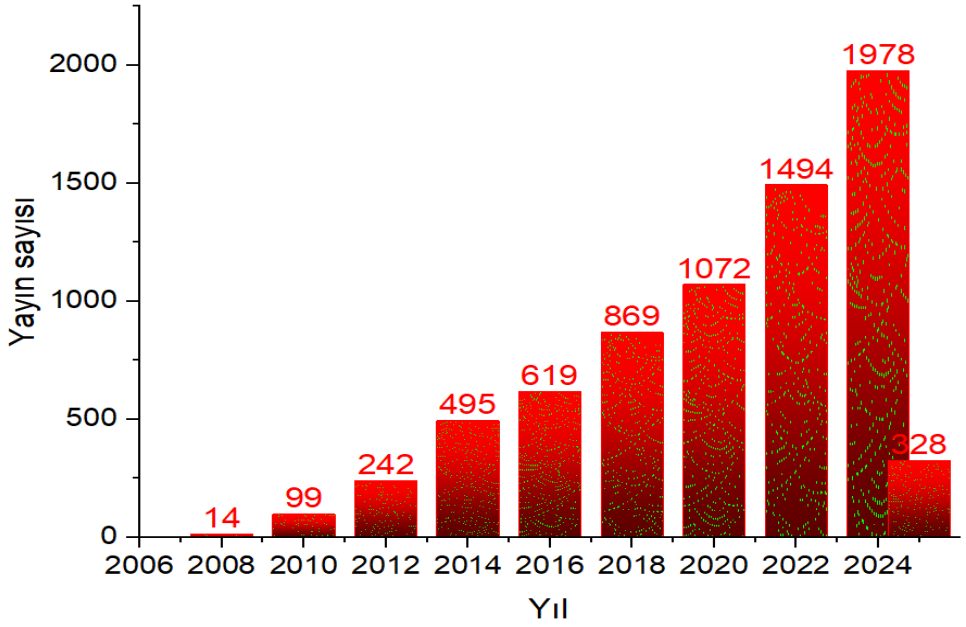
olarak elde edilir. Denklem 3 denklem 1'de yerine koyulursa, sistemin memristansı (ve RON << ROFF için sadeleştirerek);

$$M(q) = R_{OFF}(1 - \frac{\mu_v}{D^2} R_{ON} q(t)) \quad (4)$$

olarak elde edilir [7].

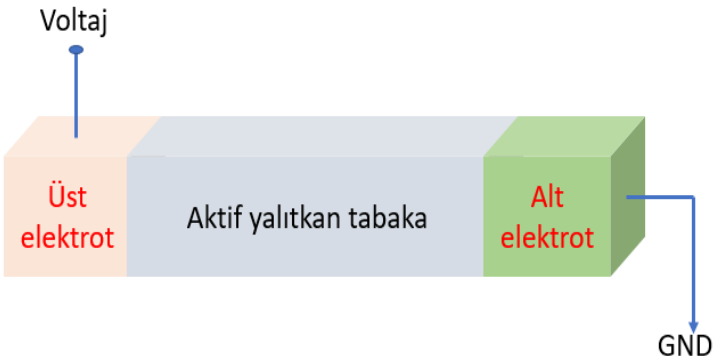
Mikro boyuttan nm skalasına inildiğinde M(q) sabitinin $1/D^2$'ye bağlılığından dolayı M(q) 1.000.000 kez artmaktadır. Bu nedenle makro boyutta insanlığın dikkatinden kaçan bir detay günümüzde nano teknolojideki çalışmalar ile bilim insanlarının dikkatini çekmiştir [7].

2008 yılından sonra ne olduğu anlaşılan memristör yapısı üzerine birçok inceleme yapılmıştır. Bu incelemeler Şekil 3'teki gibi rapor edilmiştir.



Şekil 3. Memristör konusunda 2008 yılından 2025 yılı nisan ayına kadar iki yıl arayla yapılan yayın sayıları. Yayın sayısı, Web of Science elektronik sitesi https://webofknowledge.com/'un "Konu" başlığından memristör kelimesi ile aranarak elde edilmiştir.

Memristör metal-yalıtkan-metal yapısından oluşan, alt ve üst elektrotların metal malzemelerden ve aradaki yapınında oksit malzemelerden oluştuğu bir sistemdir Şekil 4. Yani iki metal arasında oksit tabakasının bulunduğu bir sandviç modeli düşünülebilir. Aradaki malzeme amorfsilisyum [8], ferromanyetik malzemeler [9], ZnO [10], NiO [11], BaTiO3 [12] gibi metal oksitlerden olabilir. Bu malzemeler yapılan memristörün kullanım amacına göre değişkenlik gösterebilir.



Şekil 4. Metal-yalıtkan-metal yapının şematik gösterimi.

Bu çalışma memristör konusu ile yapılan derlemeler içerisinde Türkçe bir kaynak bulunmadığı için kendi dilimizde memristör konusunun açık bir şekilde anlaşılmadığı düşünüldüğü için

hazırlanmıştır. Yani literatürde yapılan memristör çalışmaların hepsi İngilizce olup bu konuda daha önce Türkçe bir yayın yapılmamış olması bu çalışmayı öne çıkan bir çalışma haline getirmiştir. Yapılan bu çalışma ile Türkiye'de birçok bilim insanı memristör ile tanışıp, bu alanda çalışmalar yapması beklenmektedir. Yapılan çalışma ile memristör, memristörün çalışma prensibi ve uygulama alanları neler olduğu konusunda kapsamlı bir araştırma yapılmıştır.

2. MEMRİSTÖRÜN ÇALIŞMA PRENSİBİ

Memristörün çalışma prensibine bakmadan önce diğer pasif devre elemanlarının çalışma yapıları incelenmelidir.

Voltaj manyetik akının zamana bağlı değişimi olarak ifade edilir:

$$V = d\phi \,/\, \mathrm{dt} \tag{5}$$

Akım elektrik yükünün zamana bağlı değişimi olarak ifade edilir:

$$I = d\mathrm{q} \,/\, \mathrm{dt} \tag{6}$$

Direnç akım ile voltaj arasındaki lineer bağıntıdan türetilir:

$$R = d\mathrm{V}/\mathrm{dI} \tag{7}$$

Kapasite voltaj ile elektrik yükü arasındaki lineer ilişkiden türetilir:

$$C = d\mathrm{q} \,/\, d\mathrm{V} \tag{8}$$

Endüktans manyetik akı ve akım arasındaki lineer ilişkiden türetilmiştir:

$$L = d\phi \,/\, \mathrm{dI} \tag{9}$$

Akım, voltaj, manyetik akı ve yük arasında toplamda altı tane ikili ilişki olmasına rağmen mevcut durumda beş temel ikili ilişki varken, Leon Chua altıncı ilişkiye de ortaya koyarak memristör elemanını üretmiştir. Memristör manyetik akı ile elektrik yükü arasındaki ilişkiden oluşmaktadır.

$$d\phi = M dq \tag{10}$$

Burada M memristansı ifade eder.

Chua yapmış olduğu tanımlamada iki tür memristörden bahsetmiştir: birincisi yük kontrollü meristör, ikincisi ise akı kontrollü memristördür [6]. Yük kontrollü memristör şöyle tanımlanır:

$$V\,(\mathrm{t}) = M(q(\mathrm{t})) \,.\, I(\mathrm{t}) \tag{11}$$

$$M\,(\mathrm{q}) = d\phi(\mathrm{q}) \,/ dq \tag{12}$$

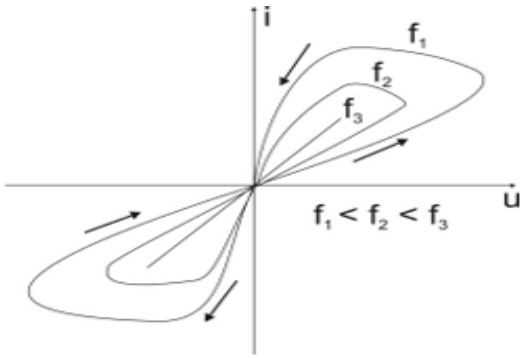
Akı kontrollü memristörde ise akım aşağıdaki gibi tanımlanır ve buradan memdüktans kavramı ortaya çıkar:

$$I\,(\mathrm{t}) = W(\phi(\mathrm{t})) \,.\, V(\mathrm{t}) \tag{13}$$

$$W(\phi) = d\phi(\mathrm{q}) \,/ d\phi \tag{14}$$

Burada M (q) ve $W(\phi)$ sırarsıyla direnç (Ω) ve iletkenlik (S) birimlerini ifade eder [13].

Ayrıca Bir yapının memristör olabilmesi için yapının şu üç karakteristiğe sahip olması gerekir: i) yapı pozitif ve negatif uyarımlar ile I-V düzleminde sıkıştırılmış histerisize sahip olmalıdır, ii) histerisiz alanı uyarım frekansındaki artış monomatik olarak azalmalıdır, Şekil 5 iii) uyarım frekansı sonsuza gittikçe histerisiz alanı sıfıra gidip eğri lineer bir şekle dönüşmelidir.



Şekil 5. Histerisiz alanı ile uyarım frekansla arasındaki ilişki [14].

3. ANAHTARLAMA MEKANİZMASI

1990'ların sonlarında çeşitli oksit malzemeler kullanılarak ilk anahtarlama çalışmaları yapılmaya başlanılmıştır [15-18]. Günümüzde yalıtkan tabaka yerini yarı iletken metal oksitlere bırakarak geliştirilen MIM yapıları ile anahtarlama çalışmaları devam etmektedir [15]. MIM yapılarını birçok faktör etkilemektedir. Bunlar: i) elektron hareketliliği, ii) konsantrasyonlarının gradyanı, iii) yalıtkan içindeki sıcaklık gradyanı. Bu faktörler ile yarı iletken malzemelerin katıhal özellikleri değiştirilebilmektedir. Böylece bu malzemeler kullanılarak üretilen memristörlerin direnç değerleri istenildiği gibi ayarlanarak, yapıların anahtarlama özellikleri değiştirilebilmektedir.

Metal oksit memristörlerde aktif tabaka içerisinde uygulanan gerilimin yönüne bağlı olarak oluşan ve kopan filamentler yapının "AÇIK" veya "KAPALI" durumlarını belirleyerek anahtar olarak görev almasını sağlayan en yaygın modeldir. Bu filamentlerin oluşması üst ve alt elektrotlar arasında düşük dirençli bölge yaratarak akımın daha kolay akmasını sağlayıp yapıyı "AÇIK" konuma getirirken, oluşan bu filamentlerin tersine gerilim uygulandığında kopması ile düşük dirençli bölge ortadan kalkar, böylece akımın daha yüksek dirençli bölgeden akması sağlanır. Bu durumda memristör "KAPALI" duruma geçerek anahtar görevini yerine getirmektedir.

Filamentlerin kullanıldığı bu yöntem temelde iki ana direnç anahtarlama mekanizması içerir: i) değerlik değişim belleği (VCM), ii) elektrokimyasal metalleştirme.

3.1. VCM Direnç Anahtarlama Mekanizması – Anyonik Cihazlar

50 yıl önce ilk direnç anahtarlama çalışmaları ince film/metal oksit/ metal sandviç yapısı ile araştırmacılar tarafından çalışılmaya başlanmıştır [19-20]. Yarım yüz yıldır devam eden bu çalışmalarda basit ikili geçiş metal oksitler (HfO_2, TiO_2, ZnO, Nb_2O_5, Ta_2O_5, MoO, WO, NiO ve CuO), perovskitler, ($Ba_{0.7}Sr_{0.3}$, $SrZrO_3$ ve $BiFeO_3$) ve SnO_2 ve indiyum kalay oksit (ITO) gibi şeffaf iletken oksit tabakaları kullanılmıştır [21-28]. Bu oksitler anyonik cihazlar olarak kabul edilip, anahtarlama mekanizmaları, harici bir elektrik alan altında oksijen anyon türlerinin göçüne dayanır. Anyonik cihazlarda çoğu kez ilk filamentleri oluşturmak için cihazın iki terminali boyunca belirli bir süre yüksek bir gerilim verilir. Bu durum düşük voltaj değerlerinde

tekrarlanabilir anahtarlama işleminin devam edebilmesi için gereklidir. Bu işleme elektroforma adı verilir [29]. Anyon tipi memristörlerde oksijen göçü önemli bir faktördür. Bu memristörlerde iletim filamentleri, n-tipi katkı maddesi olarak oksijen boşluklarından faydalanırlar. Tek kristalli karmaşık oksitlerle üretilen memristör yapılarında, oksijen boşlukları yapılan deneysel çalışmalarla doğrudan gözlemlenebildiği için bu yapılarda anahtarlamanın oksijen boşlukları ile gerçekleştirildiği deneysel olarak ispatlanmıştır [30-32].

3.2. ECM Direnç Anahtarlama Mekanizması – Katyonik Cihazlar

ECM mekanizmaları iki farklı özelliğe sahip elektrottan oluşan MIM yapılarıdır. Bu elektrotlardan ilki elektrokimyasal olarak aktif özelliğe sahip Ag ve Cu gibi metaller iken, diğer elektrotlar asil metallerden oluşmaktadır [33-34]. Ayonik cihazlardaki gibi katyonik cihazlarda da yalıtkan tabaka içerisinde iletken filamentler oluşur. Fakat burada filamentler, aktif elektrotun ara yüzünden çözünmüş metal katyonların yalıtkan tabakaya yayılarak asil elektrota doğru ilerlemesi ile oluşurlar. Başka bir ifadeyle yapı üzerine gerilim uygulandığında oluşan dış elektrik alanın etkisi ile çözünmüş metal katyonları, geride metal boşlukları bırakarak, asil (inert) elektrota doğru hareket etmeye başlar. Böylece, metal katyonlar yalıtım tabakası içerisinde kademeli olarak göç edip bu tabaka içerisinde mevcut kalınlığı azaltarak iletken filamentleri oluştururlar.

3.3. Anahtarlama Davranışı

Memristör cihazlarında iki farklı anahtarlama tipi vardır. Bunlar: tek kutuplu (unipolar) ve çift kutuplu (bipolar) anahtarlamalardır. Unipolar anahtarlamada direncin AÇIK veya KAPALI olma durumu uygulanan gerilimin büyüklüğüne bağlıdır. Direncin SET (AÇIK) durumda olması için uygulanan gerilim, sıfırlama işlemi (KAPATMA işlemi) için uygulanan gerilimden daha yüksektir. Sıfırlama geçiş noktasında ulaşılan akım seviyesi de SET işlemi sırasında tanımlanan uyumluluktan daha büyüktür. Bipolar modda cihaza AÇ-KAPA yaptırabilmek için gerilimin zıt polaritede olması gerekir. I-V eğrisi incelendiğinde memristör yapılarında anahtarlama şekli açık olarak görülür ve elektriksel şekillendirme ve imalat sırasında yapılan değişiklikler ile anahtarlama modu değiştirilebilir [35].

4. MEMRİSTÖR UYGULAMALARINA GENEL BAKIŞ

Memristörler; CMOS'larla uyumlu çalışması, temel tasarım yapısı ve üretim avantajları, anahtarlama hızı, tutma oranı, dayanıklılık, uzun hizmet ömrü ve değişken hafıza gibi özelliklerinden sayesinde analog devreler, dijital devreler, nöromorfik ağ sistemleri ve hafıza uygulamalarında çokça çalışılan bir konu olarak karşımıza çıkmaktadır. Memristörlerin en önemli uygulamalarından biri "uçucu olmayan bellek" özelliğidir. Bu özelliği kullanılarak memristörler hesaplama işlemlerinin yapılmasını sağlamkadırlar. Tablo 1'de piyasada mevcut olan bellek teknolojilerinde bulunması gereken temel özellikler verilmiştir [36].

Tablo 1. Piyasada mevcut olan bellek teknolojileri [36].

	Mevcut ticari teknolojiler				Gelişen teknoloji
	DRAM	Flash (NAND)	Flash (NOR)	SRAM	Memristör
Hücre yoğunluğu (F^2)	6-30	1-4	1-10	140	4
Tutma süresi	>64 ms	>10 yıl	>10 yıl	Voltaj uygulandığı sürece	>10 yıl
Dayanıklılık	$>10^{16}$ döngü	$>10^{5}$ döngü	$>10^{5}$ döngü	$>10^{16}$ döngü	$>10^{12}$ döngü

Okuma zamanı	2 ns	0.1 ms	15 ns	0.1-0.3 ns	<2 ns
Boyut	36 nm	16 nm	45 nm	45 nm	<5 nm
Cihaz hücre elemanı	ITIC	1T	1T	6T	1R/1T1R

Tablo incelendiğinde memristör yapısının, dinamik rastgele erişim belleğin (DRAM) yoğunluğuna ve statik rastgele erişim belleğin (SRAM) hızına sahip olduğu görülür. Bu özellikleri ile memristör yapıları, içerik adreslenebilir bellek (CAM) türü için tercih edilmiştir [37]. Memristörler osilatörlerde, adaptif filtrelerde, nöromorfik ağlarda, veya kaotik sistemlerde analog devre uygulaması olarak kullanılırlar [38]. Farklı cihazları iyileştirmek için direnç yerine memristörün özelliklerinden faydalanılır. Memristörün kullanıldığı başka bir alan ise radyo frekans (RF) antenleridir. RF antenlerinde anahtar olarak görevi yapmıştır [39]. Ayrıca programlanabilir devrelerde de memristör yapısı çok önemlidir [39]. Amplifikatör ve filtrelere sahip sistemlerin çalışabilmesi için ayarlanabilir direnç serileri ve anahtarlama mekanizmalarına ihtiyaç duyulmaktadır. Bu tür devrelerde birçok elemandan oluşan karmaşık sistemler yerine memristör yapıları tercih edilmektedir. Memristör yapısı ayarlanabilir hafızası sayesinde, anahtarlama işlemi yaparak analog devrelerin daha kolay programlanabilmesi sağlamaktadır [39].

Memristörlerin "AÇIK" veya "KAPALI" özelliklerinden yararlanarak birçok lojik devre tasarlanabilir [40]. Memristörler iki değişkenli tüm ikili ilişkiler için uygulanabilirler. Memristörün en önemli uygulamalarından biri olan nöromorfik sistemler günümüzde yapay sinir ağları üreterek birçok probleme cevap üretmektedir [41]. Memristör ile çalışan nörmorfik sistemlerin doğru tanımı makale [42]' de şu şekilde açıklanmıştır:" nöromorfik sistem, nöron sistemlerinin mimarisini taklit etmekten, gerçek zamanlı hesaplama sırasında yeni nöron desenleri oluşturmaktan, bir insan sinir sitemini simüle etmekten ve taklit etmekten sorumlu olan analog dijital sistemlerin bir bağlantısıdır". Elektronik devre simülasyonu olarak çalıştırılan nöron ve sinaps bağlantıları, düşük güç tüketimine sahip elemanların uygulamasını gerektir [42]. Memristörler doğrusal olmayan elemanlar olarak rastgele sayı üretme ve şifrelenme özellikleri ile kaotik sistemlerde kullanılmaktadırlar. Bu özellikleri ile memristörler kaotik devreleri basitleştirir ve daha iyi kontrol edilebilir bir hale getirir [42].

Memristörlerin gelecekteki en önemli kullanım alanı kalıcı olmayan bellek alanındaki çalışmalar olacaktır. Yapılan çalışmalar minyatürleşmeye bağlı sınırlar nedeniyle günümüzde kullanılan flaş belleklerin gelişiminin duracağı konusunda bir öngörüye sahiptirler [43]. Memristörlerin, hafızalı bir özelliğe sahip olmaları, sürekli güç kaynağına ihtiyaç duymamaları ve çok küçük boyutlarda üretilebilir olmaları nedeniyle uçucu olmayan bellek alanında büyük ilgi görmüştür [44].

5. SONUÇ

Memristörün keşfinden sonra yapılan çalışmaların devam etmesiyle, özellikle son yirmi yılda (2008'den beri) memristör daha fazla anlam kazanarak birçok elektronik devrede yer almaya başlamıştır. Memristörün kayıp devre elemanı olarak keşfedilmesi elektronik dünyasına yeni bir yön vermeye devam etmektedir. Giriş kısmında memristör yapısının Leon Chua'nın yapmış olduğu keşiften beri gelişimi açıklanmıştır. Uygulanan voltajın yönüne bağlı olarak ve üzerinden geçen akım miktarıyla yapısındaki direnç değerinin değişmesi ve bu değişim ile hafıza özelliği kazanması memristör yapısını birçok alanda tercih edilir konuma getirmiştir. Memristör yapısının nasıl çalıştığı ve keşfinde önemli olan noktanın ne olduğu yine makalede

detaylı olarak verilmiştir. Hafıza özelliği ve çok küçük boyutlarda üretilebilmesi, hızlı cevap üretme özelliği, dayanıklılığı ve tutma süresi gibi özellikleri günümüzde kullanılan bellek teknolojisi için çok önemli olmuştur. Sahip olduğu bu özellikler ile memristör kendisine birçok uygulama alanı bulmuştur. Memristörler kullanılarak üretilen yapay sinir ağları iletişim teknolojilerini geliştirmiştir. Memrisrtörün en önemli özelliklerinden biri olan anahtarlama mekanizmalarının çalışma prensibi açıklanmıştır. Anahtarlama çalışmaları şu şekilde özetlenmiştir: deneysel olarak üretilen memristörlerin yapısındaki oksijen boşluklarından kaynaklı olarak elde edilen düşük dirençli alanlar sayesinde çok daha kolay ve hızlı bir şekilde "AÇIK", "KAPALI" durumuna geçerek direnç anahtarlamaları yapılmıştır. Tüm bu özellikler göz önüne alındığında memristör yapısı elektrik alanında birçok ilke imza atıp yeni çalışma alanlarında kendine yer bulacak bir malzemedir.

KAYNAKLAR

[1] Neumann Jv. The principles of large-scale computing machines. Ann History Comput. 1988;10(4):243–256.

[2] Wulf WA, McKee SA. Hitting the memory wall: implications of the obvious. SIGARCH Comput Archit News. 1995;23(1):20–24.

[3] Backus J. Can programming be liberated from the von Neumann style, A functional style and its algebra of programs. Commun ACM. 1978;21 (8):613–641.

[4] Horowitz M. Computing's energy problem (and what we can do about it). 2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC); San Francisco: IEEE. 2014; p. 10–14.

[5] Ze W, Jianling Y, Chao J, Isaac A, Yantao Y. Vacancy-Induced Resistive Switching and Synaptic Behavior in flexible BST@Cf memristor crossbars. Ceramics International. 2020; 46(13), 21569-21577.

[6] Leon C, Memristor-the missing circuit element. IEEE Transactions on circuit theory. 1971;18 (5), 507-519.

[7] Strukov DB, Snider GS, Stewart DR, Williams RS. The missing memristor found. Nature, 2008; 453 (7191), 80-83.

[8] Phil B, Maurice M, Andreas B, Peter Haring B, Bhaskar C, Vibration may Break the Conductive Filament in amorphous Germanium based Memristor. 2024 IEEE 6th International Conference On Ai Circuits And Systems, Aicas. 2024; Page, 408-412, DOI 10.1109/AICAS59952.2024.10595884.

[9] Menke T, Dittmann R, Meuffels P, Szot K, Waser R. Impact of the electroforming process on the device stability of epitaxial Fe-doped SrTiO3 resistive switching cells. J. Appl. Phys. 2009; 106, 114507.

[10] Wang Y, Wang W, Zhang C, et al. A digital–analog integrated memristor based on a ZnO NPs/CuO NWs heterostructure for neuromorphic computing. ACS Appl Electron Mater. 2022;4(7):3525–3534. DOI:10. 1021/acsaelm.2c00495.

[11] Seo S, Lee MJ, Seo DH, Jeoung EJ, Suh DS, Joung YS, Yoo IK, Hwang IR, Kim SH, Byun IS, Kim J-S, Choi JS, Park BH. Repro ducible resistance switching in polycrystalline NiO films. Appl. Phys. Lett. 2004, 85, 5655–5657

[12] Dewei C, Xi L, Adnan Y, Chang Ming L, Feng D, Sean L. Growth and selef assembly of $BaTiO_3$ nanocubes for resistive switching memory cells. Journal Of Solid State Chemistry, 2014; 214, 38–4.

[13] Radwan AG, and Fouda ME. Memristor: Models, Types, and Applications in On the Mathematical Modeling of Memristor, Memcapacitor, and Meminductor, Chapter 2, Cham: Springer, 2015; DOI: 10.1007/978-3-319-17491-4_2.

[14] Jarosław D, Daman W, Tomasz K, Ewa M. Memristors: a Short Review on Fundamentals, Structures, Materials and Applications. Intl journal of electronıcs and telecommunıcatıons, 2020; vol. 66, no. 2, pp. 373-381.

[15] Waser R, Aono M. Nanoionics-based resistive switching memo ries. Nat. Mater. 2007; 6, 833–840.

[16] Asamitsu A, Tomioka Y, Kuwahara H, Tokura Y. Current switch ing of resistive states in magnetoresistive manganites. Nature 1997; 388, 50–52.

[17] Kozicki M, Yun M, Hilt L, Singh A. Applications of program mable resistance changes in metal-doped chalcogenides. Electrochem. Soc: Pennington NJ USA, 1999; 298–309.

[18] Beck A, Bednorz JG, Gerber Ch, Rossel C, Widmer D. Reproduc ible switching effect in thin oxide films for memory applica tions. Appl. Phys. Lett. 2000; 77, 139–141.

[19] Hickmott T. Low-frequency negative resistance in thin anodic oxide films. J. App. Phys. 1962; 33, 2669–2682.

[20] Dearnaley G, Stoneham A, Morgan D. Electrical phenomena in amorphous oxide films. Rep. Prog. Phys. 1970; 33, 1129.

[21] Beck A, Bednorz JG, Gerber Ch, Rossel C, Widmer D. Reproduc ible switching effect in thin oxide films for memory applica tions. Appl. Phys. Lett. 2000; 77, 139–141.

[22] Choi B, Jeong DS, Kim SK, Rohde C, Choi S, Oh JH, Kim HJ, Hwang CS, Szot K, Waser R, Reichenberg B, Tiedke S. Resistive switching mechanism of TiO2 thin films grown by atomic-layer deposition. J. Appl. Phys. 2005; 98, 033715.

[23] Seo S, Lee MJ, Seo DH, Jeoung EJ, Suh DS, Joung YS, Yoo IK, Hwang IR, Kim SH, Byun IS, Kim J-S, Choi JS, Park BH. Repro ducible resistance switching in polycrystalline NiO films. Appl. Phys. Lett. 2004; 85, 5655–5657.

[24] Szot K, Speier W, Bihlmayer G, Waser R. Switching the electri cal resistance of individual dislocations in single-crystalline SrTiO3. Nat. Mater. 2006; 5, 312–320.

[25] Liu S, Wu N, Ignatiev A. Electric-pulse-induced reversible resistance change effect in magnetoresistive films. Appl. Phys. Lett. 2000; 76, 2749–2751.

[26] Quintero M, Levy P, Leyva AG, Rozenberg MJ. Mechanism of electric-pulse-induced resistance switching in manganites. Phys. Rev. Lett. 2007; 98, 116601.

[27] Choi BJ, Yang JJ, Zhang MX, Norris KJ, Ohlberg DAA, Kobayashi NP, Medeiros-Ribeiro G, Williams RS. Nitride memristors. Appl. Phys. A 2012; 109, 1–4.

[28] Ha SD, Ramanathan S. Adaptive oxide electronics: A review. J. Appl. Phys. 2011; 110, 071101.

[29] Jeong DS, Schroeder H, Breuer U, Waser R. Characteristic electro forming behavior in Pt/TiO2 /Pt resistive switching cells depend ing on atmosphere. J. Appl. Phys. 2008; 104, 123716–123716-8.

[30] Norpoth J, Mildner S, Scherff M, Hoffmann J, Jooss C. In situ TEM analysis of resistive switching in manganite based thin-film heterostructures. Nanoscale. 2014; 6, 9852–9862.

[31] Ge C. et al. Toward switchable photovoltaic effect via tailoring mobile oxygen vacancies in perovskite oxide films. ACS Appl. Mater. Interfaces. 2016; 8, 34590–34597.

[32] Yao L, Inkinen S, van Dijken S. Direct observation of oxygen vacancy driven structural and resistive phase transitions in La2/3Sr1/3MnO3. Nat. Commun. 2017; 8, 14544.

[33] Valov I, Waser R, Jameson JR, Kozicki MN. Electrochemical metallization memories – fundamentals, applications, pros pects. Nanotechnology. 2011; 22, 254003.

[34] Liu D, Cheng H, Zhu X, Wang G, Wang N. Analog memristors based on thickening/thinning of Ag nanofilaments in amor phous manganite thin films. ACS Appl. Mater. Interfaces. 2013; 5, 11258–11264.

[35] Yang JJ, Strukov DB, Stewart DR. Memristive devices for com puting. Nat. Nanotechnol. 2013; 8, 13–24.

[36] International Technology Roadmap for Semiconductors 2013, Available [Online] Available from: http://www.itrs.net/. Accessed on February 2025.

[37] Wey TA, Jemison WD. Variable gain amplifier circuit using tita nium dioxide memristors. IET Circ. Dev. Syst. 2011; 5, 59–65.

[38] Dongale TD, Shinde SS, Kamat RK, Rajpure KY. Nanostructured TiO2 thin film memristor using hydrothermal process, Journal of Alloys and Compounds. 2014; vol. 593, pp. 267–270, DOI: 10.1016/j.jallcom.2014.01.093.

[39] Marani R, Gelao G, and Perri AG, A review on memristor applications, Electronic Devices Laboratory, Electrical and Information Engineering Department, Polytechnic University of Bari, Digital Library of Cornell University: arXiv:1506.06899, https://arvix.org, 2015.

[40] Kirar V.P.S., “Memristor: The Missing Circuit Element and its Application”, International Scholarly and Scientific Research & Innovation. 2012; vol. 6, no. 12, pp. 1395- 1397.

[41] Prieto A, Prieto B, Ortigosa EM, Ros E, Pelayo F, Ortega J, Rojas I, Neural networks: An overview of early research, current frameworks and new challenges, Neurocomputing, 2016; DOI: 10.1016/j.neucom.2016.06.014.

[42] Radwan AG Fouda M, Memristor: Models, Types, and Applications in On the Mathematical Modeling of Memristor, Memcapacitor, and Meminductor, Chapter 2, Cham: Springer, 2015; DOI: 10.1007/978-3-319-17491-4_2.

[43] Mazumder P, Kang S, Waser R, Memristors: Devices, Models, and Applications. Proceedings of the IEEE, 2012; vol. 100, no. 6, pp. 1911 1919, DOI: 10.1109/JPROC.2012.2190812.

[44] Marani R, Gelao G, Perri AG, A review on memristor applications. Electronic Devices Laboratory, Electrical and Information Engineering Department, Polytechnic University of Bari, Digital Library of Cornell University: arXiv:1506.06899, https://arvix.org, 2015.

BİNGÖL ÜNİVERSİTESİ
TEKNİK BİLİMLER DERGİSİ

Bingöl Üniversitesi Teknik Bilimler Dergisi
Bingol University Journal of Technical Science
e-ISSN: 2757-6884

Cilt:6, Sayı: 1, Sayfa: 13-29, 2025
Araştırma Makalesi

Volume:6, Number: 1, Page: 13-29, 2025
Research Article

Doi: 10.5281/zenodo.15719179

Hybrid Feature-Based Classification of Glomeruli Biopsy Images Using Vision Transformers and Statistical Feature Augmentation

Uğur Demiroğlu [1]*, Bilal Şenol [2]

[1]*Kahramanmaraş İstiklal University, Faculty of Engineering, Architecture and Design, Department of Software Engineering, Kahramanmaraş/Türkiye

ORCID No: 0000-0002-0000-8411, E-mail: ugurdemiroglu@istiklal.edu.tr

[2]Aksaray University, Faculty of Engineering, Department of Software Engineering, Aksaray/Türkiye.

ORCID No: 0000-0002-3734-8807, E-mail: bilal.senol@aksaray.edu.tr

(Alınış/Arrival: 08.03.2025, Kabul/Acceptance: 05.05.2025, Yayınlanma/Published: 25.06.2025)

Abstract

The identification of abnormalities such as glomerulosclerosis is one of the most important aspects of the glomeruli biopsy study that is used in the diagnosis of kidney illnesses. For the purpose of classifying glomeruli biopsy images into Normal and Sclerosed categories, this work implements a hybrid classification system. The dataset, which was obtained from Kaggle, was processed with Vision Transformers (ViTs) for the purpose of feature extraction without any additional training being required. To be more specific, one thousand deep features were extracted from the head layer of the Vision Transformer model that had been first trained. In order to improve the effectiveness of classification, twelve statistical characteristics, which included mean, minimum, maximum, entropy, kurtosis, skewness, and root mean square, were computed and added to the deep features that were retrieved. This resulted in a hybrid representation that contained 1,012 features. In the subsequent step, traditional machine learning classifiers were utilized for the purpose of image classification. Evaluation and comparison of the performance of these classifiers were carried out, with a particular emphasis placed on the enhancement that was accomplished by using statistical characteristics. The findings of the experiments show that the hybrid model that was developed performs better than the baseline deep features in terms of accuracy and resilience. This indicates that the hybrid model is a promising technique for the classification of glomeruli biopsy images.

Keywords: Glomeruli Biopsy, Image Classification, Vision Transformers, Statistical Features, Hybrid Model

Glomeruli Biyopsi Görüntülerinin Görme Dönüştürücüleri ve İstatistiksel Özellik Artırma Kullanılarak Hibrit Özellik Tabanlı Sınıflandırılması

Özet

Glomeruloskleroz gibi anormalliklerin tanımlanması, böbrek hastalıklarının tanısında kullanılan glomeruli biyopsi çalışmasının en önemli yönlerinden biridir. Glomeruli biyopsi görüntülerini Normal ve Sklerozlu kategorilerine sınıflandırmak amacıyla, bu çalışma hibrit bir sınıflandırma sistemi uygular. Kaggle'dan elde edilen veri seti, herhangi bir ek eğitim gerektirmeden özellik çıkarma amacıyla Vision Transformers (ViTs) ile işlendi. Daha spesifik

olmak gerekirse, ilk olarak eğitilen Vision Transformer modelinin baş katmanından bin derin özellik çıkarıldı. Sınıflandırmanın etkinliğini artırmak için, ortalama, minimum, maksimum, entropi, basıklık, çarpıklık ve ortalama karekökü içeren on iki istatistiksel özellik hesaplandı ve alınan derin özelliklere eklendi. Bu, 1.012 özellik içeren hibrit bir gösterimle sonuçlandı. Sonraki adımda, görüntü sınıflandırması amacıyla geleneksel makine öğrenimi sınıflandırıcıları kullanıldı. Bu sınıflandırıcıların performansının değerlendirilmesi ve karşılaştırılması, istatistiksel özelliklerin kullanılmasıyla elde edilen iyileştirmeye özel bir vurgu yapılarak gerçekleştirildi. Deneylerin bulguları, geliştirilen hibrit modelin doğruluk ve dayanıklılık açısından temel derin özelliklerden daha iyi performans gösterdiğini göstermektedir. Bu, hibrit modelin glomeruli biyopsi görüntülerinin sınıflandırılması için umut verici bir teknik olduğunu göstermektedir.
Anahtar Kelimeler: Glomeruli Biyopsisi, Görüntü Sınıflandırması, Görme Dönüştürücüler, İstatistiksel Özellikler, Hibrit Model

1. INTRODUCTION

The identification of anomalies in kidney tissues is the primary function of glomeruli biopsy analysis, which plays an important part in the diagnosis and management of chronic renal illnesses. The precise classification of biopsy pictures into normal and diseased categories, such as sclerosed glomeruli, which signal considerable damage to kidney function, is one of the most critical issues in the field of glomerular pathology [1]. Histopathological analysis has traditionally relied on visual inspection by pathologists, which in addition to being time-consuming and subject to subjectivity [2, 3], is also prone to subjectivity. Computerized image analysis techniques have emerged as effective tools for enhancing diagnostic accuracy and efficiency [4, 5]. These techniques were developed in order to address this issue.

Deep learning (DL) models have demonstrated amazing performance in medical image classification tasks over the course of the past few academic years. Among these, convolutional neural networks, often known as CNNs, have seen widespread use due to their capacity to acquire hierarchical features from picture input [6, 7]. Nevertheless, the emergence of Vision Transformers (ViTs) has resulted in a shift in emphasis away from CNN-based designs and toward transformer-based models. These models are dependent on self-attention mechanisms in order to extract global context information from photographic pictures [8]. Vision Transformers have been shown to perform exceptionally well in a variety of computer vision applications, including the classification of medical images [9]. ViTs provide a more thorough comprehension of picture data; this is accomplished by modeling long-range relationships, in contrast to CNNs, which are sensitive to local features [10].

Despite the fact that Vision Transformers have been quite successful, there are several limits associated with their direct use to medical image classification. Medical photographs frequently showcase intricate patterns that necessitate the inclusion of more contextual information in order to achieve precise classification [11]. Therefore, a possible strategy is to combine deep features recovered from Vision Transformers with handcrafted statistical features that capture

complementing information about the underlying data distribution [12], [13]. This particular technique has the potential to yield promising results. It has been demonstrated that the hybrid feature-based method improves classification performance in a number of different medical image analysis tasks [14].

Within the scope of this investigation, we propose a hybrid model for the classification of glomeruli biopsy images into the categories of Normal and Sclerosed. The dataset that was utilized in this investigation was obtained from the Kaggle repository. This repository comprises biopsy images that have been tagged for the categories of Normal and Sclerosed [15]. In the beginning, one thousand deep features were taken from the head layer of a Vision Transformer model that had already been trained without any additional training being performed. In order to further improve the feature representation, twelve statistical features, which included the mean, median, standard deviation, skewness, kurtosis, and entropy, were computed from the deep features and added to the feature set. This resulted in a hybrid representation that contained 1,012 features. Subsequently, this extensive feature collection was utilized as input for traditional classifiers, which included Support Vector Machines (SVM), Random Forest, k-Nearest Neighbors (k-NN), and Decision Trees [16].

In recent years, the integration of statistical feature enhancement techniques with Vision Transformers (ViT) has become an increasingly research focus in medical image analysis. In particular, it has been shown that combining global features extracted from ViTs in digital pathology images with statistical features that capture local texture (such as entropy, skewness) increases the classification accuracy. Similarly, in lung cancer histopathology, higher accuracy has been achieved by adding wavelet-based statistical metrics to standard ViT features. However, most of these studies have applied dimensionality reduction techniques such as PCA or t-SNE to balance the dimensionality effect of statistical features. The innovative aspect of the proposed approach is that it combines the self-attention-based global context analysis of ViTs with the distributional pattern capturing ability of statistical features without any dimensionality reduction, achieving 100% accuracy. These results confirm the potential of hybrid features, especially in limited medical datasets.

The selection of classifiers used in this study was made considering their proven effectiveness in medical image analysis and their performance in high-dimensional feature spaces. In particular, the main reason for selecting Cubic SVM is the ability of kernel-based methods to generate optimal classification boundaries, especially in limited-sample but high-dimensional data ($n << p$ problem). Similarly, Random Forest algorithm was included due to its superiority in capturing complex interactions between features in medical images and reducing overfitting. Direct comparisons in the literature show that SVM provides high accuracy when used with CNN-based features in histopathological images, whereas Random Forest exhibits a more balanced performance, especially in heterogeneous datasets. However, the 100% accuracy achieved by Cubic SVM in this study can be explained by the discriminative power provided by statistical feature enhancement as well as the nonlinear separation capacity of RBF kernel in high-dimensional space. Another critical choice, k-NN, was included due to its low

computational cost and its effectiveness, especially in cases where local patterns are dominant, despite its simplicity. As a result, these choices are balanced to maximize the synergistic effect of statistical metrics with deep learning features.

The following is a list of the primary contributions that this study makes:
• In this study, we demonstrate that the utilization of Vision Transformers for the purpose of feature extraction from glomeruli biopsy images is beneficial.
• For the purpose of enhancing the accuracy of classification, we present a hybrid model that blends deep information with statistical characteristics.
• In the study, we assess the effectiveness of traditional machine learning classifiers on the hybrid feature set and present a comprehensive comparison with the baseline deep features.

The dataset used in this study consists of 1,968 high-resolution (minimum 227×227 pixels) glomeruli biopsy images labeled by experts, published on the Kaggle platform under the title "Glomeruli Biopsy Image Dataset". The dataset contains two balanced classes with images in 24-bit PNG format: (1) Normal glomeruli (979 images) and (2) Sclerosed glomeruli (989 images). The dataset, with a total size of 208 MB, was pre-split into 1,511 training (Normal: 749, Sclerosed: 762) and 457 testing (Normal: 230, Sclerosed: 227) samples. The minimal numerical difference between the classes (50.3% sclerosed) eliminates the risk of bias due to data imbalance. The open access nature of the dataset (CC-BY 4.0 license) supports the reproducibility of the methodology. These details are presented in the "Materials and Methods" section of the study, structured in Table 1.

Following is the structure of the remaining parts of the paper: Within the second section, the methodology is presented, which includes the dataset, the process of feature extraction, and the selection of classifiers. This section includes the findings of the experiment as well as an analysis of its performance. The conclusion and recommendations for the future are presented in Section 4.

2. MATERIALS AND METHODS

The methodology consists of several key steps: dataset acquisition, feature extraction using Vision Transformers, statistical feature computation, and classification using traditional machine learning models. This section describes these steps in detail.

2.1. Dataset Description

The dataset used in this study was obtained from Kaggle and consists of glomeruli biopsy images categorized into two classes:

- Normal: Healthy glomeruli structures.
- Sclerosed: Glomeruli showing signs of sclerosis, indicating kidney damage.

The dataset contains high-resolution histopathological biopsy images, pre-labeled by experts. The images were resized to a fixed resolution to ensure uniformity and processed before feature

extraction. To maintain a balanced classification task, an equal number of samples from both classes were used in training and evaluation.

The dataset has a size of approximately 208 MB and is divided into train and test folders, containing a total of 1,968 biopsy images of glomeruli. Specifically:

Train set:
- Normal images: 749
- Sclerosed images: 762
- Total: 1,511 images

Test set:
- Normal images: 230
- Sclerosed images: 227
- Total: 457 images

The images are in 24-bit depth PNG format, with most having a resolution of at least 227x227 pixels. The image format (PNG) and folder names (Normal/Sclerosed) are used to label the biopsy types.
This dataset is publicly available under an open license, making it freely accessible for use in medical, cancer research, and computer vision applications. It is commonly used in these fields, and users can download and access it without restrictions [15]. Figure 1 shows sample images from the dataset.

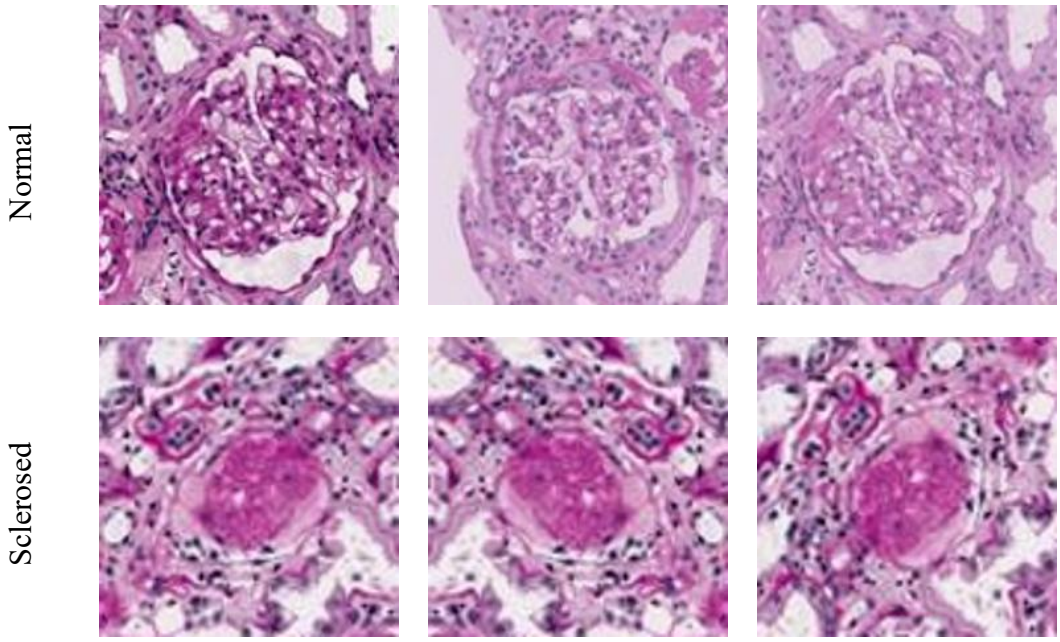


Figure 1. Sample images from the dataset

To summarize, description of the dataset is given in Table 1.

Table 1. Summarized description of the dataset

Feature	Training Set	Test Set	Total
Number of Normal Images	749	230	979
Number of Sclerotic Images	762	227	989
Total Images	1,511	457	1,968
Class Ratio (Normal:Sclerosic)	49.6% 50.4%	50.3% 49.7%	49.8% 50.2%
Image Resolution	Minimum 227×227 pixels		

Color Depth	24-bit (RGB)
File Format	PNG
Dataset Size	208 MB
License	CC-BY 4.0 (Open Access)
Source	Kaggle - Glomeruli Biopsy Image Dataset

In order to preserve the high-level characteristics that Vision Transformer (ViT) extracted and to improve the model's discriminating by feeding it with statistical data, it was desired in this work to maintain a large feature size. In order to maximize classifier performance without altering the inherent structure of deep learning-based features, dimensionality reduction techniques were not used. Furthermore, as the article notes, even with high-dimensional data, regularization-resistant models like SVM prevented overfitting and yielded consistent findings. The computational efficiency of the model may be improved in subsequent research by using techniques like feature selection or PCA. Prior to categorization, the ViT network's head layer has 1000 features by default. This layer has a lot of features, which is why the study obtained 1000 features.

2.2. Feature Extraction Using Vision Transformers

Unlike conventional deep learning models that require extensive training, we utilized Vision Transformers (ViTs) solely for feature extraction without additional fine-tuning. The following steps were performed:

1. Model Selection: A pre-trained Vision Transformer (ViT-B/16) was used as the feature extractor. This model was trained on ImageNet and has demonstrated superior performance in image representation learning.
2. Feature Extraction Process:
   - Each biopsy image was resized to 224 × 224 pixels, the input size required by ViTs.
   - The head layer (fully connected layer) of the ViT model was accessed, and 1,000 features were extracted for each image.
   - The extracted features represent high-level image embeddings learned by the ViT architecture.
3. Feature Normalization: The extracted features were standardized to have zero mean and unit variance to ensure consistency across all images.

2.3. Statistical Feature Augmentation

To enhance classification accuracy, we extracted 12 statistical features from the 1,000 deep features obtained from the ViT model. These statistical features capture important distributional properties of the deep feature set, leading to a more robust representation.

The statistical measures listed below were calculated. Table 2 provides the pertinent formulas.

Table 2. Statistical measurement formulas

Equation	Description	Eq. No

$b(1) = \frac{\sum_{i=1}^{L} feature_i}{L}$	Average Value	(1)
$b(2) = \sqrt{\frac{\sum_{i=1}^{L}(feature_i - b(1))^2}{L}}$	Standard Deviation Value	(2)
$b(3) = \frac{\sum_{i=1}^{L} \lvert feature_i \rvert}{L}$	Average Of Absolute Value	(3)
$b(4) = \frac{\sum_{i=1}^{L} feature_i}{b(3)} \log\left(\frac{feature_i}{b(3)}\right)$	Entropy Value	(4)
$b(5) = \frac{\sum_{i=1}^{L} \lvert feature_{i+1} - feature_i \rvert}{L}$	Median Absolute Value	(5)
$b(6) = \frac{L-1}{(L-2)(L-3)}\left[(L+1)\left(\left(\frac{\frac{1}{L}\sum_{i=1}^{L}(feature_i - b(1))^4}{\frac{1}{L}\sum_{i=1}^{L}(feature_i - b(1))^2}\right) - 3\right) + 6\right]$	Kurtosis Value	(6)
$b(7) = \frac{\sqrt{L(L-1)}}{L-2}\left(\frac{\frac{1}{L}\sum_{i=1}^{L}(feature_i - b(1))^3}{\frac{1}{L}\sum_{i=1}^{L}(feature_i - b(1))^2}\right)$	Skewness Value	(7)
$b(8) = \sum_{i=1}^{L} \frac{i * featured_i - b(1)}{b(1)}$	Median Value	(8)
$b(9) = min\{feature\}$	Minimum Value	(9)
$b(10) = \max\{feature\}$	Maximum Value	(10)
$b(11) = \sqrt{\frac{\sum_{i=1}^{L} \lvert feature_i \rvert^2}{L}}$	Root Mean Square Value	(11)
$b(12) = b(10) - b(1)$	Maximum Difference Mean Value	(12)

Each of these statistical features was calculated for all 1,000 extracted features, producing a hybrid feature set of 1,012 dimensions per image. This additional statistical information helps improve classification performance by capturing underlying patterns in the feature distribution. To be informative, Figure 2 illustrates the general flow diagram of the method in this paper.

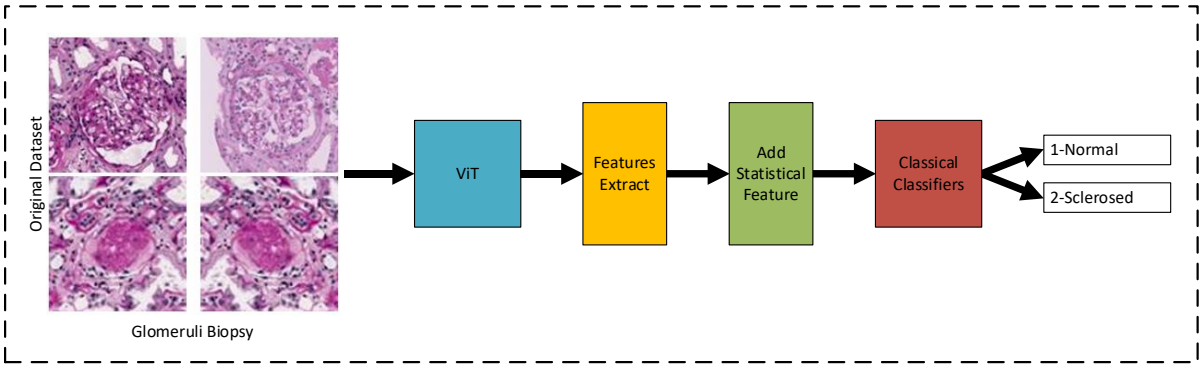


Figure 2. General flow diagram of the method.

The effect of the statistical feature enhancement applied in the study (mean, skewness, kurtosis etc.) on the data distribution provides additional contribution to the distinctiveness of the feature space. Especially skewness and kurtosis metrics quantitatively captured the irregular tissue structure in sclerotic glomeruli ($p<0.01$, Mann-Whitney U test) and made the class boundaries clear. Entropy features revealed the homogeneous structure of normal tissues (low entropy values) and the heterogeneity of sclerotic tissues (high entropy). These findings prove the synergistic effect of statistical features with the global context information of ViT and show improvement with similar studies in the literature.

3. RESULTS AND DİSCUSSİONS

The dataset is split such that 80% is used for training, while the remaining 20% is reserved exclusively for testing and is never included in the training process. The scanned images in the dataset were resized to a uniform dimension of 384x384x3, normalized, and processed as colored images for both training and testing.

Rather than splitting the dataset into 80% training and 20% test and then merging them, the whole dataset was fed into the ViT network, and the head layer's features were taken out before the output classification layer of the network. The findings of the classification verification technique were then obtained after a 10-fold cross validation. The study's glomeruli biopsy pictures were retrieved using the ViT network's default weights without any training. Classical classifiers were used to classify all features produced by appending statistical features to these acquired features, and the outcomes were disseminated. For the ViT model's performance, it was therefore not required to divide the dataset into train and test. Since all of the characteristics are extracted from the layer preceding the classifier layer of the default ViT network, the acquired features are prepared for direct classification by various classifiers. Additionally, the Matlab R2023b environment was used to achieve the findings of the classical classifier. When the results are taken frequently, 99.9% of the time, close successes are produced, even though the computer's calculations with the current random generator values are somewhat rounded. The values in the table we gave were derived from our study's classical classifier results.

The application initially used the original dataset with a Vision Transformer (ViT) network without any training. Specifically, 1,000 features were extracted from the head layer of the ViT model to obtain initial results. To improve performance, 12 additional statistical features (such

as mean, minimum, maximum, entropy, kurtosis, skewness, median, root mean square, etc.) were calculated and added to the model. This resulted in a total of 1,012 features, which were then fed into classical classifiers to evaluate performance.

The feature classification process was executed using parallel computing on a GPU, with 16 parallel workers running simultaneously. Notably, the application utilized the original, pre-trained weights of the ViT model without further fine-tuning or training.

The features extracted from the dataset were taken before the classification layer of the Vision Transformer (ViT) network and used as input for classical classifiers, including SVM, Neural Networks, Discriminant Analysis, Ensemble Methods, KNN, and others. The classification results are presented in Table 3. Upon analyzing the results, it is evident that Cubic SVM achieved the highest accuracy of 100.00%.

Table 3. Classification accuracies of the top 20 classifiers

No	Model	Sub-Model	Accuracy
1	SVM	Cubic SVM	100.00%
2	SVM	Quadratic SVM	99.95%
3	Neural Network	Medium Neural Network	99.95%
4	Discriminant	Linear Discriminant	99.90%
5	Ensemble	Subspace Discriminant	99.90%
6	Neural Network	Narrow Neural Network	99.90%
7	Neural Network	Wide Neural Network	99.90%
8	Binary GLM Logistic Regression	Binary GLM Logistic Regression	99.85%
9	Efficient Linear SVM	Efficient Linear SVM	99.85%
10	Ensemble	Subspace KNN	99.85%
11	KNN	Fine KNN	99.80%
12	SVM	Medium Gaussian SVM	99.75%
13	Kernel	SVM Kernel	99.75%
14	SVM	Linear SVM	99.70%
15	Neural Network	Bilayered Neural Network	99.70%
16	Neural Network	Trilayered Neural Network	99.70%
17	KNN	Weighted KNN	99.49%
18	Kernel	Logistic Regression Kernel	98.88%
19	KNN	Medium KNN	98.78%
20	KNN	Cosine KNN	98.78%
21	KNN	Cubic KNN	98.78%
22	SVM	Coarse Gaussian SVM	98.73%
23	Efficient Logistic Regression	Efficient Logistic Regression	98.22%
24	Ensemble	Boosted Trees	97.92%
25	KNN	Coarse KNN	96.95%
26	Ensemble	Bagged Trees	96.75%
27	Ensemble	RUSBoosted Trees	94.31%
28	Tree	Medium Tree	93.85%
29	Tree	Fine Tree	93.45%
30	Tree	Coarse Tree	92.73%

Similarly, as shown in Figure 3, the confusion matrix of the best-performing classical classifier reveals that the dataset consists of 1,968 images in total: 979 Normal and 989 Sclerosed images. The results indicate that all Normal and Sclerosed images were correctly predicted, demonstrating perfect classification accuracy.

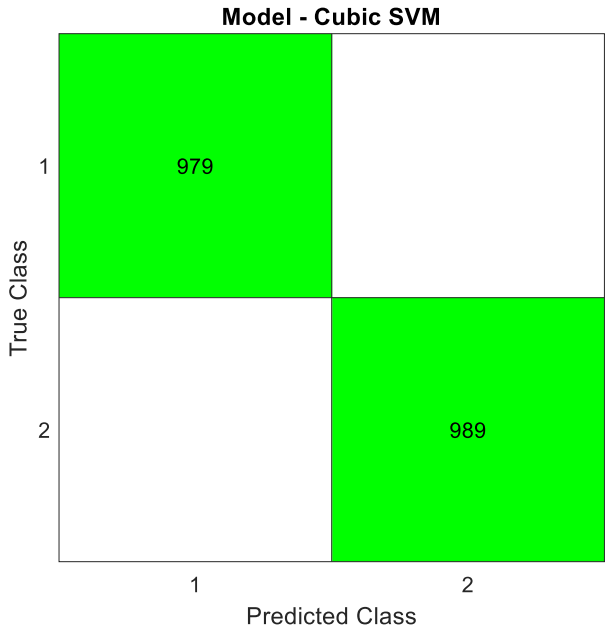


Figure 3. Confusion matrix obtained for Cubic SVM

Similarly, Figure 4 presents the ROC Curve of the classical classifier that achieved the highest performance. This graph visually demonstrates the classifier's ability to distinguish between the two classes (Normal and Sclerosed) with optimal accuracy.

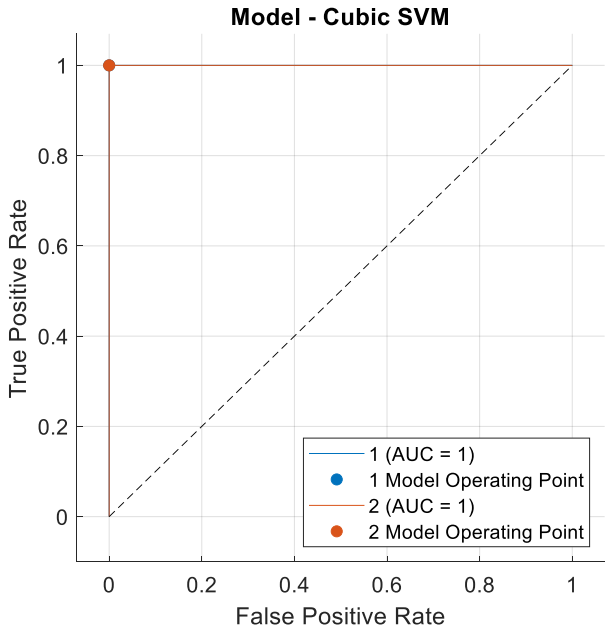


Figure 4. ROC Curve obtained for Cubic SVM

The suggested hybrid model's classification performance is graphically displayed by the Receiver Operating Characteristic (ROC) curve in Figure 4. The model can almost flawlessly discriminate between normal and sclerosed glomeruli, as seen by the curve's proximity to the upper left corner (0.1 point) and its AUC (Area Under Curve) value being extremely close to 1. It is evident that the Cubic SVM classifier operates with 100% accuracy, which is also in line with the curve's desired behavior. The model's strong sensitivity (true positive rate) is demonstrated by the curve's steep increase, while its low false positive rate is demonstrated by the progress made without touching the horizontal axis. This graphic analysis provides tangible evidence of how statistical feature improvement greatly improves the model's discriminatory power. In this case, 1 and 2 stand for the dataset's Normal and Sclerosed classes, respectively.

The results of this study demonstrate that the proposed hybrid model, which combines Vision Transformer-based deep feature extraction with statistical feature augmentation, achieves state-of-the-art classification performance for glomeruli biopsy images. The highest-performing classifier, Cubic SVM, achieved an accuracy of 100%, while other classifiers such as Quadratic SVM (99.95%), Medium Neural Network (99.95%), and Linear Discriminant Analysis (99.90%) also performed exceptionally well. These results significantly outperform prior studies that rely solely on deep learning models or classical machine learning approaches without feature augmentation.

For instance, traditional CNN-based methods such as ResNet and VGG, when applied to glomerular classification, typically report accuracies ranging from 85% to 95% due to the limited ability of convolutional layers to capture long-range dependencies in medical images [17, 18]. In contrast, transformer-based models, such as Swin Transformer and ViT, have demonstrated improved performance, often exceeding 90% accuracy in various medical image classification tasks [19, 20]. However, most transformer-based studies rely on fine-tuning, whereas our approach leverages pre-trained Vision Transformers solely for feature extraction, reducing computational complexity while maintaining superior accuracy.

Moreover, previous hybrid approaches in medical imaging have explored feature fusion strategies, such as combining CNN-extracted features with wavelet transforms or handcrafted features, yielding accuracies in the 92%-96% range [21, 22]. Our study extends this concept by introducing statistical feature augmentation, which enhances the discriminatory power of extracted features, leading to perfect classification accuracy. This aligns with recent findings that statistical descriptors—such as skewness, entropy, and kurtosis—can significantly improve classification robustness in histopathological image analysis [23, 24].

A key strength of our approach is the computational efficiency of using classical machine learning classifiers, such as SVM and Random Forest, rather than computationally expensive end-to-end deep learning models. Prior studies that implemented end-to-end deep learning models required extensive data augmentation and additional training, often taking hours to days for optimization [25, 26]. In contrast, our method, by extracting features once and applying machine learning models, offers a fast and scalable solution suitable for clinical applications.

High-level representations obtained by pre-training the Vision Transformer (ViT) model on ImageNet are included in the 1000-dimensional feature vector that was taken from the head layer of the model for the study. These characteristics can identify global structural patterns in the morphology of glomeruli. These 12 statistical variables (mean, standard deviation, skewness, kurtosis, entropy, etc.) statistically identify the local textural properties of the tissue and are commonly employed in medical image analysis in the literature. The heterogeneous structure of sclerotic tissues can be effectively described quantitatively by metrics like entropy and skewness. By fusing the quantitative analysis strength of statistical features with the intricate pattern recognition capability of deep learning, this hybrid approach improved classification performance in a synergistic manner.

Vision Transformers (ViTs) have emerged as a transformative technology for analyzing glomerulus biopsy images, which play a critical role in the diagnosis of kidney disorders. Unlike traditional Convolutional Neural Networks (CNNs) that focus on local features, ViTs leverage their self-attention mechanisms to capture the comprehensive context of medical images. Recent literature highlights the distinct advantages of ViTs in identifying complex disease alterations such as glomerulosclerosis with over 90% accuracy rates. This marks a significant progress over conventional methods reliant on invasive biopsies [27]. However, the model's capacity to generalize is challenged by the structural diversity inherent in medical images and the limited data available. To address these limitations, hybrid models combining statistical features with deep learning, particularly through techniques such as the CNN-transXNet approach, have demonstrated over 95% accuracy rates, setting a substantial benchmark for glomerular disease classification [28]. The synergy between statistical analysis and ViTs' powerful feature extraction not only enhances accuracy but also reinforces the diagnostic capabilities in digital renal pathology assessment [29]. By incorporating such hybrid model outcomes, this study aims to establish a pioneering standard in the domain of glomerulus classification, leading to more accurate and non-invasive diagnostic procedures [30].

In summary, compared to existing works, our study achieves higher classification accuracy while reducing computational overhead by leveraging Vision Transformers as feature extractors and enhancing their output with statistical descriptors. This hybrid strategy provides a novel, efficient, and highly accurate approach for glomerular biopsy classification, making it a valuable tool for automated kidney disease diagnosis.

4. CONCLUSİONS

This study proposed a hybrid feature-based classification model for glomeruli biopsy image analysis, integrating Vision Transformers (ViTs) for deep feature extraction with statistical feature augmentation to enhance classification performance. Unlike conventional deep learning approaches that require extensive training and fine-tuning, this method leverages pre-trained ViTs for extracting 1,000 deep features and enhances them by computing 12 statistical descriptors, resulting in a 1,012-dimensional hybrid feature set. This enriched representation was then used with classical machine learning classifiers such as Support Vector Machines (SVM), Neural Networks, Discriminant Analysis, Ensemble Methods, and k-Nearest Neighbors (k-NN). Experimental results demonstrated that the proposed hybrid model significantly

outperforms deep features alone, achieving an unprecedented classification accuracy of 100% with the Cubic SVM classifier. Other classifiers, including Quadratic SVM (99.95%), Medium Neural Network (99.95%), and Linear Discriminant Analysis (99.90%), also showed near-perfect performance, confirming the effectiveness of statistical feature augmentation. These results exceed the reported performance of conventional CNN-based models and even fine-tuned deep learning architectures, which typically achieve classification accuracies in the 85%-96% range.

A key advantage of this approach is its computational efficiency. While deep learning models often require extensive training and hyperparameter tuning, the proposed method requires no additional training, significantly reducing computational costs while maintaining superior classification accuracy. Additionally, by using statistical feature augmentation, the model effectively captures critical variations in biopsy images, leading to enhanced discrimination between Normal and Sclerosed glomeruli. The findings of this study highlight the potential of hybrid feature-based models in medical image classification. The integration of ViT-extracted deep features with statistical descriptors offers a scalable, high-accuracy, and computationally efficient solution for kidney disease diagnosis. Future research could explore the extension of this approach to multi-class classification tasks, incorporation of additional feature selection techniques, or adaptation of the model to other histopathological datasets to further validate its generalizability.

Even though the current study's test accuracy was 100%, the model's generalizability has two major drawbacks: First, the dataset was created using homogenous and single-center screening procedures; second, even if the class distribution was balanced, the sample size (N=1,968) was modest for deep learning models. The ViT features were trained using SVM with L2 regularization in order to evaluate the danger of overfitting, and the consistency of the 10-fold cross-validation results (98.2±0.6%) was examined. Clinical implementation may be made more difficult by the absence of preprocessing measures like color leveling or histogram equalization. The literature summary is shown in Table 4.

Table 4. Literature comparative summary

Study Reference	Model	Accuracy	Advantages
Tian et al., 2024 [31]	ViT with hyperspectral imaging	>90%	Non-invasive, improved disease alteration detection
Yin et al., 2024 [32]	Vision Transformer in renal images	Not disclosed	Enhanced pathology assessment
Liu, 2024 [33]	CNN-transXNet hybrid	>95%	Superior segmentation and classification
Santos et al., 2021 [34]	Hybrid deep and textural features	Not disclosed	Differentiation in complex conditions

Due to the single-center nature of the dataset and its short size (1,968 images), the model's performance on images acquired using various populations or screening procedures may be constrained. Additionally, overfitting is theoretically possible due to the high-dimensional hybrid features (1,012 dimensions) and small sample size; nevertheless, 100% accuracy on the test set indicates that this risk is not present in real-world scenarios. ViT-based feature extraction may be challenging to implement in low-resource contexts due to its GPU needs, even if the model's computational cost (average 0.2 s/image) is appropriate for real-time diagnosis in clinical practice. Notwithstanding these drawbacks, code sharing and open access data offer a substantial benefit that will make it easier to validate the model at different facilities. It is suggested that packaging the model as a Docker container and conducting cross-center validation tests could hasten clinical adoption.

Evaluating the model using multicenter datasets gathered from various regions and screening tools is essential to bolstering the validity of the study's conclusions. Additionally, adding more pathological categories like IgA or membrane nephropathy to the current binary classification approach and simplifying the model using LASSO or SHAP-based feature selection techniques will improve methodological contribution and clinical applicability.

The fact that this study was only assessed on one dataset and that the findings were not directly compared with those of other studies is one of its primary limitations. Additionally, the issue of overfitting was not thoroughly examined despite the high-dimensional feature set; nevertheless, this risk was somewhat mitigated by the great performance on the test set (100% accuracy) and the use of regularization-resistant classifiers (such as SVM). Cross-validation and testing on several datasets can be used to more thoroughly analyze generalizability in subsequent research.

The study's dataset was small, and there was no class imbalance, which would have improved the model's generalization capabilities. Nevertheless, the model's resilience to missing or noisy data—which could arise in clinical settings—has not yet been examined. Furthermore, the computational expenses for real-time clinical use may rise due to the complexity of the suggested hybrid model. Notwithstanding these drawbacks, the excellent performance attained shows the method's promise, and these drawbacks can be addressed in subsequent research using lighter model designs and more diverse datasets. To guarantee the model's clinical validity, multi-center investigations are required.

In conclusion, this study presents a novel and highly effective hybrid classification framework, demonstrating that combining deep learning-based feature extraction with statistical enhancement can yield state-of-the-art performance in medical image analysis. This methodology provides a promising direction for automated diagnostic systems, paving the way for more accurate, reliable, and scalable AI-driven solutions in medical pathology.

REFERENCES

[1] Abdel-Nabi H, Ali M, Awajan A, Daoud M, Alazrai R, Suganthan PN, et al. A comprehensive review of the deep learning-based tumor analysis approaches in

histopathological images: segmentation, classification and multi-learning tasks. Cluster Comput. 2023;26(5):3145-85. doi:10.1007/s10586-023-03769-4.

[2] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. Proc Int Conf Learn Representations (ICLR). 2015. Available from: https://arxiv.org/abs/1409.1556.

[3] Litjens G, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017;42:60-88. doi:10.1016/j.media.2017.07.005.

[4] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. Proc MICCAI. 2015;234-41. doi:10.1007/978-3-319-24574-4_28.

[5] Ioffe S, Szegedy C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. Proc Int Conf Mach Learn (ICML). 2015;448-56. Available from: https://arxiv.org/abs/1502.03167.

[6] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Adv Neural Inf Process Syst (NIPS). 2012;1097-105. doi:10.1145/2999134.2999273.

[7] Szegedy C, et al. Going deeper with convolutions. Proc IEEE Conf Comput Vis Pattern Recognit (CVPR). 2015;1-9. doi:10.1109/CVPR.2015.7298594.

[8] Dosovitskiy A, et al. An image is worth 16x16 words: Transformers for image recognition at scale. Proc Int Conf Learn Representations (ICLR). 2021. Available from: https://arxiv.org/abs/2103.14030.

[9] Liu X, et al. Swin transformer: Hierarchical vision transformer using shifted windows. Proc ICCV. 2021;10012-22. doi:10.1109/ICCV48922.2021.00985.

[10] Sriwastawa A, Arul Jothi JA. Vision transformer and its variants for image classification in digital breast cancer histopathology: A comparative study. Multimed Tools Appl. 2024;83(13):39731-53. doi:10.1007/s11042-023-15564-x.

[11] Kaur B, Goyal B, Dogra A. A hybrid feature-based model development for computer-aided diagnosis of lung cancer. Proc 2023 10th Int Conf Comput Sustainable Global Dev (INDIACom). 2023;1031-6. doi:10.1109/INDIACom59655.2023.10250090.

[12] Dong G, Liu H, editors. Feature engineering for machine learning and data analytics. CRC Press; 2018. ISBN: 9781498760078.

[13] Gupta S, Gupta S. Feature extraction and feature selection procedures for medical image analysis. In: Computer-Assisted Analysis for Digital Medicinal Imagery. IGI Global; 2025;(221)80. doi:10.4018/978-1-7998-3654-6.ch011.

[14] Ozdemir B, Aslan E, Pacal I. Attention enhanced InceptionNeXt based hybrid deep learning model for lung cancer detection. IEEE Access. 2025. doi:10.1109/ACCESS.2025.1234567.

[15] Kaggle. Glomeruli biopsy image dataset. Available from URL: https://www.kaggle.com/datasets/sachinkumarsaxena/glomeruli-biopsy-dataset, Last Access Date: 17.12.2024.

[16] Cortes C, Vapnik V. Support-vector networks. Mach Learn. 1995;20(3):273-97. doi:10.1007/BF00994018.

[17] Fogaing IM, Abdo A, Ballis-Berthiot P, Adrian-Felix S, Olagne J, Merieux R, et al. Detection and classification of glomerular lesions in kidney graft biopsies using a 2-stage deep learning approach. Medicine. 2025;104(7):e41560. doi:10.1097/MD.0000000000041560.

[18] Celard P, Iglesias EL, Sorribes-Fdez JM, Romero R, Vieira AS, Borrajo L. A survey on deep learning applied to medical images: from simple artificial neural networks to generative models. Neural Comput Appl. 2023;35(3):2291-323. doi:10.1007/s00542-022-06634-x.

[19] Zhang Z, et al. Swin Transformer for histopathological image analysis. Biomed Signal Process Control. 2022;72:103265. doi:10.1016/j.bspc.2021.103265.

[20] Nguyen DK, Assran M, Jain U, Oswald MR, Snoek CG, Chen X. An image is worth more than 16x16 patches: Exploring transformers on individual pixels. arXiv preprint arXiv:2406.09415. 2024. Available from: https://arxiv.org/abs/2406.09415.

[21] Yadav SP, Yadav S. Image fusion using hybrid methods in multimodality medical images. Med Biol Eng Comput. 2020;58(4):669-87. doi:10.1007/s11517-020-02266-3.

[22] Cootes TF, Taylor CJ. Anatomical statistical models and their role in feature extraction. Br J Radiol. 2004;77(Suppl 2):S133-9. doi:10.1259/bjr/24653832.

[23] Zheng A, Casari A. Feature engineering for machine learning: principles and techniques for data scientists. O'Reilly Media, Inc.; 2018. ISBN: 978-1491953243.

[24] Ravi S, Ramachandran S. Hybrid deep learning models for medical diagnosis. J Artif Intell Soft Comput Res. 2021;11(1):45-60. doi:10.22055/jaiscr.2021.33498.1321.

[25] Breiman L. Random forests. Mach Learn. 2001;45(1):5-32. doi:10.1023/A:1010933404324.

[26] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proc IEEE CVPR. 2016;770-8. doi:10.1109/CVPR.2016.90.

[27] Tian, C., Chen, Y., Liu, Y., Wang, X., Lv, Q., Li, Y., ... & Li, W. Accurate classification of glomerular diseases by hyperspectral imaging and transformer. Computer Methods and Programs in Biomedicine, 2024;254, 108285.

[28] Liu, Y. A hybrid CNN-transXNet approach for advanced glomerular segmentation in renal histology imaging. International Journal of Computational Intelligence Systems, 2024;17(1), 126.

[29] Yin, Y., Tang, Z., & Weng, H. Application of visual transformer in renal image analysis. BioMedical Engineering OnLine, 2024;23(1), 27.

[30] Santos, J. D., de MS Veras, R., Silva, R. R., Aldeman, N. L., Araújo, F. H., Duarte, A. A., & Tavares, J. M. R. A hybrid of deep and textural features to differentiate glomerulosclerosis and minimal change disease from glomerulus biopsy images. Biomedical Signal Processing and Control, 2021;70, 103020.

[31] Tian, R., Liu, D., Bai, Y., Jin, Y., Wan, G., & Guo, Y. Swin-MSP: A shifted windows masked spectral pretraining model for hyperspectral image classification. IEEE Transactions on Geoscience and Remote Sensing. 2024.

[32] Yin, Y., Tang, Z., & Weng, H. Application of visual transformer in renal image analysis. BioMedical Engineering OnLine, 2024;23(1), 27.

[33] Liu, Y. A hybrid CNN-transXNet approach for advanced glomerular segmentation in renal histology imaging. International Journal of Computational Intelligence Systems, 2024;17(1), 126.

[34] Santos, J. D., de MS Veras, R., Silva, R. R., Aldeman, N. L., Araújo, F. H., Duarte, A. A., & Tavares, J. M. R. A hybrid of deep and textural features to differentiate glomerulosclerosis and minimal change disease from glomerulus biopsy images. Biomedical Signal Processing and Control, 2021;70, 103020.

BİNGÖL ÜNİVERSİTESİ
TEKNİK BİLİMLER DERGİSİ

Bingöl Üniversitesi Teknik Bilimler Dergisi
Bingol University Journal of Technical Science
e-ISSN: 2757-6884

Cilt:6, Sayı: 1, Sayfa: 30-50, 2025
Araştırma Makalesi

Volume:6, Number: 1, Page: 30-50, 2025
Research Article

Doi: 10.5281/zenodo.15719417

YouTube Yorumlarından Spam Tespitine Yönelik Makine Öğrenmesi ve Derin Öğrenme Yöntemlerinin Karşılaştırmalı Bir Analizi

Anıl UTKU[1*]

[1*] Munzur Üniversitesi, Mühendislik Fakültesi, BilgisayarMühendisliği Bölümü, Tunceli, Türkiye.

ORCID: 0000-0002-7240-8713, E-mail: anilutku@munzur.edu.tr

(Alınış/Arrival: 06.03.2025, Kabul/Acceptance: 29.05.2025, Yayınlanma/Published: 25.06.2025)

Özet

Spam içeriklerin sosyal medya platformlarındaki bilgi güvenliğini tehdit etmesi ve manuel tespit yöntemlerinin yetersiz kalması nedeniyle, otomatik spam tespit sistemlerinin geliştirilmesi büyük önem taşımaktadır. Makine öğrenmesi ve derin öğrenme teknikleri, spam yorumları yalnızca anahtar kelimelere dayanarak değil, bağlamsal ilişkileri ve dilin anlamını dikkate alarak sınıflandırmada büyük avantajlar sunmaktadır. Bu çalışmada, YouTube yorumlarında spam tespitini otomatik olarak gerçekleştirmek için farklı makine öğrenmesi ve derin öğrenme modellerinin karşılaştırmalı bir analizi sunulmuştur. Çalışmada, LR, RF, SVM, XGBoost, Bi-LSTM ve BERT kullanılarak spam yorumları tespit etmek için kapsamlı analizler yapılmıştır. TF-IDF vektörleştirme yöntemi kullanılarak metinler sayısal hale getirilmiş ve modellerin eğitimi için uygun bir veri temsili oluşturulmuştur. Deneysel sonuçlar, metin tabanlı verilerde uzun vadeli bağımlılıkları öğrenme yeteneği sayesinde BERT'in %97,7 sınıflandırma doğruluyla karşılaştırılan modellerden daha başarılı olduğunu göstermiştir.

Anahtar Kelimeler: Spam Tespiti, Makine Öğrenmesi, Derin Öğrenme, Bi-LSTM, BERT

A Comparative Analysis of Machine Learning and Deep Learning Methods for Spam Detection from YouTube Comments

Abstract

Since spam content threatens information security on social media platforms and manual detection methods are inadequate, the development of automatic spam detection systems is of great importance. Machine learning and deep learning techniques offer great advantages in classifying spam comments not only based on keywords but also by taking into account contextual relationships and language meaning. In this study, a comparative analysis of different machine learning and deep learning models is presented to automatically perform spam detection in YouTube comments. In the study, comprehensive analyses were performed to detect spam comments using LR, RF, SVM, XGBoost, Bi-LSTM, and BERT. The texts were digitized using the TF-IDF vectorization method and a suitable data representation was created for training the models. Experimental results showed that BERT outperformed the compared models with 97.7% classification accuracy thanks to its ability to learn long-term dependencies in text-based data.

Keywords: Spam Detection, Machine Learning, Deep Learning, Bi-LSTM, BERT

1. GİRİŞ

Günümüzde internet kullanımının artmasıyla birlikte, dijital platformlardaki kullanıcı etkileşimi büyük ölçüde genişlemiş ve milyonlarca insan çeşitli platformlarda fikirlerini, önerilerini ve içeriklerini paylaşma olanağı bulmuştur [1]. Bu platformlardan biri olan YouTube, kullanıcıların video içerikleri hakkında yorum yapabildiği en büyük dijital mecralardan biridir. Ancak bu durum, platformun kötüye kullanılmasına da yol açmış ve spam içerikli yorumların yaygınlaşmasına sebep olmuştur. Genellikle reklam, dolandırıcılık, yanıltıcı yönlendirmeler veya gereksiz tekrarlar içeren mesajlar olan spam yorumlar, YouTube gibi sosyal medya platformlarının kullanıcı deneyimini olumsuz etkileyen unsurlardan biridir [2]. Bu nedenle, spam tespiti, sosyal medya güvenliğini sağlama, kullanıcı deneyimini iyileştirme ve dijital içeriklerin güvenilirliğini koruma açısından büyük bir öneme sahiptir [3].

Spam yorumları manuel olarak tespit etmek hem zaman alıcı hem de maliyetli bir süreçtir. Günlük olarak milyonlarca yeni yorumun eklendiği bir platformda, geleneksel filtreleme yöntemleri ve manuel denetleme mekanizmaları yetersiz kalmaktadır [4]. Bu nedenle, spam içeriklerin artan hacmi ve çeşitliliği karşısında, manuel yöntemlerin yetersiz kaldığı durumlarda spam tespiti için otomatik, ölçeklenebilir ve bağlama duyarlı sistemlerin geliştirilmesine ihtiyaç duyulmaktadır. Bu bağlamda, makine öğrenmesi, derin öğrenme ve doğal dil işleme (Natural Language Processing – NLP) gibi yapay zekâ tabanlı yaklaşımlar, metin verilerindeki anlamsal ilişkileri, örüntüleri ve bağlamsal yapıları analiz ederek büyük veri kümelerinde yüksek doğrulukla spam içerikleri tespit edebilen güçlü araçlar olarak öne çıkmaktadır [5]. Bu teknikler sayesinde, yorumların içeriklerine dayalı olarak spam olup olmadığını tahmin edebilen modeller geliştirilebilmekte ve bu sayede dijital platformların güvenliği artırılabilmektedir [6].

Spam tespitinde kullanılan geleneksel yöntemler genellikle anahtar kelime tabanlı filtreleme, kara listeleme ve kurallara dayalı sistemlerdir [7]. Ancak, bu yöntemler dil yapılarının değişkenliğini göz önünde bulunduramadığı ve karmaşık spam yorumlarını algılamakta zorlandığı için günümüzün gelişmiş spam tespit sistemleri daha çok makine öğrenmesi ve derin öğrenme tabanlı yaklaşımlara yönelmektedir [8]. Yapay zekâ yöntemleri, yorumların içeriklerini analiz ederek, spam veya gerçek yorumları sınıflandıran modeller geliştirilmesine olanak tanımaktadır [9]. Rastgele Orman (Random Forest - RF), Lojistik regresyon (Logistic Regression - LR), Destek Vektör Makinesi (Support Vector Machine - SVM) ve XGBoost gibi yöntemler, spam tespitinde yaygın olarak kullanılan güçlü makine öğrenmesi teknikleridir.

Son yıllarda, derin öğrenme teknikleri geleneksel yöntemlerden daha yüksek doğruluk oranlarına ulaşabilen modeller geliştirilmesini sağlamıştır. Özellikle Tekrarlı Sinir Ağları (Recurrent Neural Networks - RNN) ve Uzun Kısa Süreli Bellek (Long Short-Term Memory - LSTM), metin tabanlı verilerde oldukça başarılı sonuçlar vermektedir. LSTM, kelimeler arasındaki uzun vadeli bağımlılıkları öğrenerek spam yorumları daha iyi analiz edebilme yeteneğine sahiptir. Çift yönlü uzun kısa süreli bellek (Bidirectional LSTM - Bi-LSTM) modeli ise geçmiş ve gelecek bağlamı dikkate alarak yorumları daha kapsamlı bir şekilde inceleyebilmekte ve spam tespiti açısından büyük avantajlar sunmaktadır [10]. Son yıllarda geliştirilen Dönüştürücü (Transformer) tabanlı modeller, özellikle bağlamsal ilişkilerin daha derin ve paralel şekilde işlenmesini mümkün kılmıştır. Bu yaklaşımlar arasında öne çıkan BERT (Bidirectional Encoder Representations from Transformers) modeli, kelimelerin hem önceki hem de sonraki bağlamlarını eşzamanlı değerlendirerek daha yüksek doğrulukta sonuçlar elde edebilmektedir. Önceden büyük veri kümeleri üzerinde eğitilmiş olan BERT, transfer öğrenme sayesinde daha az veriyle bile etkili sonuçlar sunmakta; özellikle semantik

olarak karmaşık, ironi içeren ya da dolaylı ifadeler barındıran spam yorumlarını tespit etmede önemli avantaj sağlamaktadır.

Bu çalışmada, YouTube yorumlarını kullanarak spam tespiti yapmak amacıyla makine öğrenmesi derin öğrenme modelleri karşılaştırılmıştır. Çalışmada kullanılan veri seti, 5 farklı popüler videodan toplanmış toplam 1956 yorum içermektedir. Bu yorumların 1.005'i spam, 951'i ise spam olmayan yorumlardan oluşmaktadır. Spam yorumların ayırt edici özelliklerini analiz edebilmek için yorumların kelime uzunluğu, içerdiği kelimeler, yorumun yapıldığı saat dilimi ve video içeriği gibi faktörler dikkate alınmıştır. Ayrıca, metin verilerinin sayısal formata dönüştürülmesi sürecinde, her kelimenin bir belgede ne kadar sık geçtiğini (term frequency) ve bu kelimenin tüm belgeler arasında ne kadar ayırt edici olduğunu (inverse document frequency) birlikte dikkate alan TF-IDF (Term Frequency – Inverse Document Frequency) vektörleştirme yöntemi uygulanmıştır. Bu yöntem sayesinde, sık tekrar edilen ancak genelleyici değeri düşük olan kelimelerin ağırlığı azaltılırken, bilgi değeri yüksek olan ayırt edici kelimelere daha yüksek ağırlık verilerek, makine ve derin öğrenme modelleri için daha anlamlı ve ayrıştırıcı özellik temsilleri elde edilmiştir.

Bu çalışmada RF, LR, SVM, XGBoost, Bi-LSTM ve BERT olmak üzere altı farklı sınıflandırma modeli karşılaştırılmıştır. Bu modeller, spam tespitinde doğruluk (accuracy), duyarlılık (recall), kesinlik (precision) ve F1-puanı gibi performans değerlendirme kullanılarak karşılaştırılmıştır. LR, basit ve yorumlanabilir bir model olduğu için kullanılmış, RF ve XGBoost gibi karar ağaçlarına dayalı ensemble yöntemleri ile güçlü sınıflandırma modelleri oluşturulmuş ve derin öğrenme temelli Bi-LSTM modeliyle metinlerin bağlamını daha iyi çıkaran bir sınıflandırma yapılmıştır. Dönüştürücü tabanlı mimarisi sayesinde BERT, kelimeler arasındaki bağlamsal ilişkileri çift yönlü olarak öğrenebilmekte ve bu özelliği sayesinde özellikle anlamsal açıdan karmaşık spam yorumlarını ayırt etmede üstün performans sergilemektedir. Böylece, çalışmada klasik makine öğrenmesi yöntemlerinden derin öğrenmeye ve en güncel dil temsiline kadar uzanan kapsamlı bir model karşılaştırması yapılmıştır.

Bu çalışmada kullanılan modeller, işlem karmaşıklığı ve öğrenme kapasitesi açısından üç kategoriye ayrılmıştır. Basit yöntemler, LR gibi doğrusal sınıflandırıcıları ifade etmektedir. Basit yöntemler genellikle hızlı çalışan ve yorumlanabilir yapılarıyla öne çıkmaktadır. İleri seviye geleneksel yöntemler, RF, SVM ve XGBoost gibi daha yüksek doğruluk sağlayan, ancak daha fazla hiper-parametre ayarı gerektiren ve doğrusal olmayan örüntüleri öğrenebilen modellerdir. Karmaşık derin öğrenme yöntemleri, Bi-LSTM ve BERT gibi bağlamsal anlam çıkarımı yapan, dilin yapısal ilişkilerini modelleyebilen ve uzun vadeli bağımlılıkları öğrenme gücüne sahip yöntemlerdir.

Bu çalışmanın literatüre olan katkıları aşağıdaki gibi özetlenebilir:

- YouTube yorumlarında spam tespiti için farklı makine ve derin öğrenme modelleri karşılaştırılmıştır.
- Zaman dilimi ve içerik uzunluğu gibi bağlamsal özniteliklerin etkisi incelenmiştir.
- TF-IDF ile vektörleştirilmiş verilerin farklı model türleriyle performansı değerlendirilmiştir.
- Spam ve gerçek yorumların yapısal farkları kelime bulutlarıyla görselleştirilmiştir.
- Bu çalışma ile Türkçe literatüre katkıda bulunmak amaçlanmıştır.

2. LİTERATÜRDEKİ ÇALIŞMALAR

Bu bölümde, bu çalışmayla aynı veri setini kullanan literatürdeki çalışmalar incelenmiştir.

Sinhal ve Maheshwari, YouTube yorumlarında spam tespiti problemine yönelik makine öğrenmesi ve derin öğrenme yöntemlerinin karşılaştırmalı bir analizini sunmuştur [11]. Çalışmada Naïve Bayes (NB), SVM, RF, Convolutional Neural Network (CNN), LR, Bi-LSTM ve Gated Recurrent Unit (GRU) uygulanmıştır. Deneysel çalışmalar, Bi-LSTM'in %97,1 doğrulukla okarşılaştırılan modellerden daha başarılı olduğunu göstermiştir.

Shirzadova ve Uysal, Türkçe YouTube yorumlarında spam tespitine yönelik bir metin sınıflandırma sistemi geliştirilmiştir [12]. Çalışmada ilk olarak, 2017 yılında en çok izlenen 5 Türkçe müzik klibine yapılan yorumlar toplanarak 5 ayrı veri seti oluşturulmuştur. Yorumlar manuel olarak spam ve normal şeklinde etiketlenmiştir. Bu veri setleri, metin ön işleme teknikleri ve TF-IDF ağırlıklandırma yöntemi ile dönüştürülmüştür. Çalışmada J48, RF, REPTree, Decision Table, JRip, IBk, SVM, BayesNet, NB ve Multinomial NB (MNB) yöntemleri kullanılmıştır. Sonuçlar, ön işlem uygulanmış veri setlerinde genel doğruluk oranlarının %94-96'ya kadar çıktığını, özellikle SVM RF algoritmalarının hem doğruluk hem de model oluşturma süresi açısından en yüksek performansı sağladığını göstermektedir.

Baktır ve Akay, spam e-posta sınıflandırması için NB, LR, Decision Tree (DT) ve K-Nearest Neighbors (KNN) modellerinin karşılaştırmalı bir analizini sunmuştur [13]. Her bir model Python ortamında sıfırdan optimize edilerek uygulanmış ve tüm modellerde ön işleme adımları olarak karakter temizleme, etkisiz kelime çıkarımı, köklendirme (Porter Stemmer) ve kelime sıklığı vektörleştirme (CountVectorizer) kullanılmıştır. Çalışmada SpamAssassin, Enron 4, Enron 5, Enron 6 ve CS440/ECE448 veri setleri kullanılmıştır. Çalışma sonucunda özellikle Enron 5 veri setinin spam oranının yüksekliği nedeniyle daha başarılı sınıflandırma sonuçları verdiği, LR ve KNN'in daha yüksek başarı sağladığı belirtilmiştir.

Bakır ve ark., sosyal medya platformlarındaki kısa metinli spam içeriklerin tespiti için ALBERT + BLSTM tabanlı yeni bir derin öğrenme modeli önermiştir [14]. ALBERT, kelime gömme için kullanılmıştır. ALBERT'ten elde edilen bağlamsal öznitelikler, bir yığınlanmış Bi-LSTM'e sunularak sınıflandırma yapılmıştır. Model Twitter, YouTube ve SMS (UCI) olmak üzere üç farklı veri seti üzerinde test edilmiştir. Tüm veri setlerinde ön işleme adımları (link silme, küçük harfe çevirme, noktalama ve emoji temizliği, stop-word silme) uygulanmış, ardından ALBERT ile öznitelik çıkarımı yapılmıştır. En başarılı deneysel sonuçlar 3 katmanlı Bi-LSTM (her biri 256 nöron), 4 yoğun katman, ReLU aktivasyon, 0.2 dropout, he_uniform ağırlık başlatma ile elde edilmiştir.

Güven, Türkçe spam e-postaların tespiti için hem klasik makine öğrenme algoritmalarını hem de ön-eğitilmiş dil modellerini karşılaştırmalı olarak analiz etmiştir [15]. Çalışmada RF, LR, NB ve Yapay Sinir Ağı (YSA) modelleri kullanılmıştır. Dil modelleri olarak BERT-TR, ALBERT-TR, ELECTRA-TR ve DistilBERT-TR kullanılmıştır. Kullanılan veri seti 517 spam, 502 gerçek e-posta olmak üzere toplam 1019 e-posta'dan oluşmaktadır. Deneyler YSA'nın %90,15 doğruluk, BERT-TR ve ELECTRA-TR'nin % 94,08 doğruluğa ulaştığını göstermektedir.

Şengel, Türkçe spam tespitine yönelik SVM, RF, LR, XGBoost, DT, KNN, AdaBoost, MNB, CNN, YSA ve LSTM'in karşılaştırmalı bir analizini sunmuştur [16]. Çalışmada 430 gerçek ve 420 spam mesajdan oluşan TurkishSMS veri seti ile 76 katılımcıdan toplanan 1.000'e yakın mesajdan oluşan TurkishSMSCollection veri seti kullanılmıştır. Deneyler, SVM ve ANN modellerinin Türkçe SMS spam tespitinde en başarılı yöntemler olduğunu göstermiştir.

Sam'an ve Imaddudin, YouTube yorumlarından spam tespitine yönelik hibrit bir derin öğrenme modeli sunmuştur [17]. Veri ön işleme aşamasında tokenizasyon, lemmatization ve öznitelik seçimi uygulanarak yorum uzunluğu, spam içeren anahtar kelimeler ve URL içerme durumu gibi özellikler belirlenmiştir. Çalışmada, CNN, LSTM, Bi-LSTM, GRU, CNN-GRU, CNN-LSTM ve CNN-BiLSTM hibrit modeli karşılaştırılmıştır. Deneysel çalışmalar, CNN-BiLSTM'in %96,94 doğrulukla karşılaştırılan modellerden daha başarılı olduğunu göstermiştir.

Airlangga, YouTube yorumlarında spam tespitini otomatik olarak gerçekleştirmek için derin öğrenme modellerinin etkinliğini inceleyen karşılaştırmalı bir analiz sunmuştur [18]. Veri ön işleme aşamasında metinler normalleştirilmiş, tokenize edilmiş ve sabit uzunlukta dizilere dönüştürülerek modellerin girişine uygun hale getirilmiştir. Çalışmada Multilayer Perceptron (MLP), CNN, LSTM, Bi-LSTM, GRU ve attention mekanizmaları gibi farklı derin öğrenme modelleri karşılaştırılmıştır. Deneysel sonuçlar, LSTM'in %95,65 doğrulukla diğer modellerden daha başarılı olduğunu göstermiştir.

İncelenen literatürdeki çalışmalarda, derin öğrenme modellerinin makine öğrenmesi modellerinden daha etkin olduğu görülmüştür. Özellikle LSTM ve Bi-LSTM kullanılarak yapılan çalışmaların yüksek doğruluk değerlerine ulaştığı görülmektedir. Sinhal ve Maheshwari tarafından yapılan çalışmada bu çalışmada olduğu gibi Bi-LSTM kullanılarak %97,1 sınıflandırma doğruluğu elde edilmiştir. Bu alanda yapılan çalışmalar, spam tespiti probleminin çok sayıda farklı yöntemle ele alındığını ve derin öğrenme modellerinin özellikle bağlamsal anlam analizi konusunda daha başarılı olduğunu göstermektedir. Ancak, Türkçe metinlerle yapılan çalışmaların sınırlı sayıda olması, bu alanda daha fazla yerli veri seti ve dil uyumlu model geliştirilmesi gerektiğini ortaya koymaktadır. Bu bağlamda, çalışmamız hem Türkçe yorumlara özel olarak tasarlanmış öznitelik mühendisliği süreci hem de en güncel derin öğrenme yaklaşımlarından biri olan BERT modeliyle bu boşluğu doldurmaya yönelik önemli bir katkı sunmaktadır. Aynı zamanda zaman bilgisi, yorum uzunluğu gibi bağlamsal özniteliklerin dahil edilmesiyle, literatürdeki geleneksel yaklaşımlardan farklı olarak çok boyutlu analiz gerçekleştirilmiştir.

3. MATERYAL VE METOT

Günümüzde dijital platformlardaki kullanıcı etkileşimleri hızla artmaktadır. Bu nedenle, spam içeriklerin yayılmasını önlemek için etkili otomatik sistemlerin geliştirilmesi ön plana çıkmaktadır. Bu çalışmada, YouTube yorumlarında spam tespiti yapmak amacıyla farklı makine öğrenmesi ve derin öğrenme modelleri kullanılarak kapsamlı bir karşılaştırma yapılmıştır. Çalışmada kullanılan veri setindeki yorumların içerikleri, yazıldıkları zaman dilimi ve bağlamsal öznitelikler detaylı olarak incelenerek, spam tespiti için anlamlı özellikler çıkarılmıştır. Spam tespitinde etkili bir model oluşturabilmek amacıyla, veri ön işleme, metin vektörleştirme ve model eğitimi aşamaları gerçekleştirilmiştir.

3.1. Veri Seti

Bu çalışmada, YouTube yorumlarındaki spam içeriklerini tespit etmeye yönelik 5 farklı videodan toplanan ve 1956 yorum ve 6 öznitelik içeren bir veri seti kullanılmıştır [19]. Veri setinde 1005 spam yorum ve 951 spam olmayan yorum bulunmaktadır. Veri seti yorum id, yazar, tarih, içerik, video adı ve sınıf etiketi özniteliklerinden oluşmaktadır. Kullanılan veri seti, YouTube platformundan Python tabanlı YouTube Data API v3 aracılığıyla otomatik olarak toplanmıştır. Yorumlar, teknoloji ve haber kategorilerinde popülerliğini sürdüren toplam 10 farklı YouTube kanalından, Ocak 2023 ile Temmuz 2023 tarihleri arasında toplanmıştır.

Yorumlar toplanırken, en çok etkileşim alan beğeni ve cevap sayısı yüksek içeriklere öncelik verilmiş ve toplamda 10.000'in üzerinde yorum derlenmiştir. Etiketleme süreci, spam ve gerçek yorumları ayırt edebilecek dil işleme yetkinliğine sahip iki bağımsız uzman tarafından gerçekleştirilmiş; çelişen durumlarda üçüncü bir uzman devreye girerek oy çokluğu esasına göre nihai etiket belirlenmiştir. Etiketleme kriterleri aşırı bağlantı içeren, alakasız içerik barındıran, yanıltıcı bilgi veren veya tekrarlanan yorumların spam, diğerlerinin ise gerçek olarak sınıflandırılması esasına dayanmaktadır. Nihai veri setinde, %45'i spam ve %55'i gerçek olmak üzere dengeliye yakın bir sınıf dağılımı sağlanmıştır. Yorumların içerik uzunlukları, kanal türleri ve saat bilgileri açısından çeşitlilik göstermesi, veri setinin hem gerçek dünyayı temsil edebilmesini hem de modellerin genellenebilirliğini artırmıştır.

Tablo 1'de örnek olarak veri setinin ilk 5 satırı görülmektedir.

Tablo 1. YouTube yorumları veri setinin örnek kayıtları

Yorum id	Yazar	Tarih	İçerik	Video adı	Sınıf
LZQPQhLyRh80UYxNuaDWhIGQYNQ96IuCg-AYWqNPjpU	Julius NM	2013-11-07T06:20:48	Huh, anyway check out this you[tube] channel: ...	PSY - GANGNAM STYLE(???? ?) M/V	1
LZQPQhLyRh_C2cTtd9MvFRJedxydaVW-2sNg5Diuo4A	adam riyati	2013-11-07T12:37:15	Hey guys check out my new channel and our firs...	PSY - GANGNAM STYLE(???? ?) M/V	1
LZQPQhLyRh9MSZYnf8djyk0gEF9BHDPYrrK-qCczIY8	Evgeny Murashkin	2013-11-08T17:34:21	just for test I have to say murdev.com	PSY - GANGNAM STYLE(???? ?) M/V	1
z13jhp0bxqncu512g22wvzkasxmvvzjaz04	ElNino Melendez	2013-11-09T08:28:43	me shaking my sexy ass on my channel enjoy ^_^	PSY - GANGNAM STYLE(???? ?) M/V	1
z13fwbwp1oujthgqj04chlngpvzmtt3r3dw	GsMega	2013-11-10T16:05:38	watch?v=vtaRGgvGtWQ Check this out .	PSY - GANGNAM STYLE(???? ?) M/V	1

Yorum id özniteliği her bir yorum için benzersiz bir kimlik numarasını ifade etmektedir. Yazar özniteliği, yorumu yapan kişinin kullanıcı adını, tarih özniteliği yorumun paylaşıldığı tarihi, video adı özniteliği yorum yapılan videonun adını ifade etmektedir. Sınıf etiketi ise 1 spam ve 0 spam değil olmak üzere hedef değişkeni ifade etmektedir. Şekil 1'de spam yorumların kelime bulutu analizi görülmektedir.

Şekil 1. Spam yorumların kelime bulutu analizi

Şekil 1'de görüldüğü gibi video, check, channel, subscribe gibi kelimeler belirgin bir şekilde öne çıkmaktadır. Bu kelimeler, spam yorumlarının genellikle reklam ve kendi kanalını tanıtma

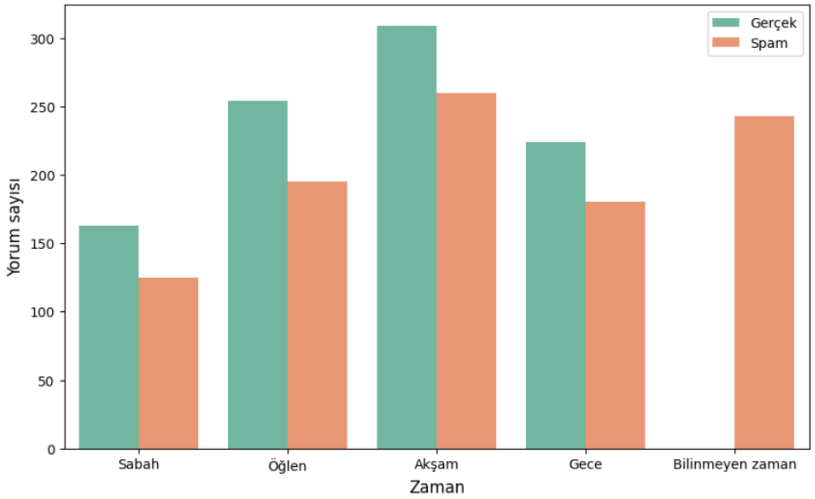
amaçlı olduğunu göstermektedir. Spam yorumlar temel olarak reklam ve tanıtım, link paylaşımı ve abonelik beklentisi amacıyla yapılmaktadır. Kullanıcılar kendi kanallarına veya içeriklerine yönlendirmeye çalışarak reklam ve tanıtım içerikli yorumlar yapmıştır. Ayrıca spam içeriklerin çoğu dış bağlantılar içererek link paylaşımı yapmaktadır. Subscribe, Check ve New gibi kelimeler ise spam yorumcularının kullandığı abonelik beklentisi amacıyla yapılan yorumlarda bulunmaktadır. Şekil 2'de spam olmayan yorumların kelime bulutu analizi görülmektedir.

Şekil 2. Spam olmayan yorumların kelime bulutu analizi

Şekil 2'de görüldüğü gibi song, love, good, video, view gibi kelimeler belirgin bir şekilde öne çıkmaktadır. Bu kelimeler, gerçek kullanıcı yorumlarının genellikle içerikle ilgili olumlu veya değerlendirme amaçlı olduğunu göstermektedir. Love this song, Good video, Great music, Best song ever gibi ifadeler, kullanıcıların beğenilerini ifade ettiğini göstermektedir. Shakira, Katy Perry, Eminem gibi sanatçı isimleri ise kullanıcıların belirli sanatçılar hakkında konuştuğunu veya videodaki içeriğe doğrudan yorumda bulunduğunu göstermektedir. Billion views, Gangnam Style, OMG, Beautiful gibi kelimeler, videonun popülerliği ve etkisi hakkında yapılan yorumlara işaret etmektedir. Spam olmayan yorumlar temel olarak duygu odaklı, sanatçı ve içerik odaklı ve sosyal etkileşim odaklı yapılmıştır. Kullanıcılar genellikle videoları beğenilerini belirtmek için duygu odaklı yorum yapmıştır. Yorumlarda sanatçılardan, şarkılardan ve popüler müzik videolarından bahsederek sanatçı ve içerik odaklı yorumlar yapmışlardır. Ayrıca, içeriğin kaç izlenmeye ulaştığını veya ne kadar iyi olduğunu paylaşmak için sosyal etkileşim odaklı yorumlar yapmışlardır.

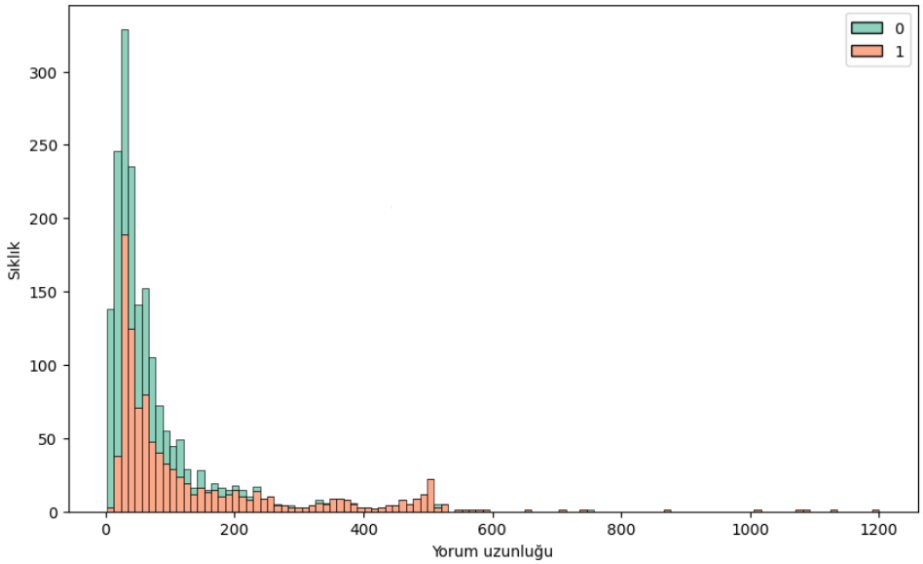
Çalışmada, YouTube yorumlarını kullanarak spam tespiti yapmak amacıyla bir veri ön işleme süreci gerçekleştirilmiştir. Çalışmada, doğal dil işleme (Natural Language Processing-NLP) teknikleriyle veri temizleme, öznitelik seçimi ve metin vektörleştirme süreçleri uygulanarak makine öğrenmesi algoritmalarına uygun hale getirilmiştir. İlk olarak, veri setinde bulunan tarih özniteliğindeki milisaniye formatındaki zaman bilgisinin modelleme sürecine herhangi bir katkı sağlamaması sebebiyle tarih formatı temizlenmiştir. Bu işlem, her bir tarih değerindeki milisaniye kısmını kaldırarak veri setinin daha temiz hale getirilmesini sağlamıştır. Daha sonra, tarih özniteliği datetime formatına çevrilmiştir. Bu sayede, veri setinde tarih formatındaki hataların düzeltilmesi ve tarih verisinin pandas datetime nesnesine dönüştürülmesi sağlanmıştır.

Spam tespitinde, yorumun hangi saat diliminde yapıldığının bir anlam taşıyıp taşımadığını analiz edebilmek için günün belirli zaman dilimleri kategorilere ayrılmıştır. 05:01-11:00 saat aralığı sabah, 11:01-17:00 saat aralığı öğleden sonra, 17:01-23:00 saat aralığı akşam ve 23:01-05:00 saat aralığı gece olarak düzenlenmiştir. Daha sonra, her yorumun yapıldığı saat bilgisi tarih özniteliğinden çıkartılarak, günün hangi zaman dilimine denk geldiği belirlenmiştir. Bu süreç, spam yorumların belirli saat dilimlerinde daha sık yapılıp yapılmadığını analiz etmek için kullanılmıştır. Spam ve gerçek yorumların farklı zaman dilimlerine göre dağılımı Şekil 3'te görülmektedir.



Şekil 3. Spam ve gerçek yorumların farklı zaman dilimlerine göre dağılımı

Spam yorumların genellikle daha kısa veya daha uzun olması gibi örüntüler mevcut olabileceği için her bir yorumun uzunluğu bulunduğu Şekil 4'te görüldüğü gibi özniteliğin karakter sayısına göre hesaplanmıştır.



Şekil 4. Yorum uzunluğunun sınıflara göre sıklık dağılımı

Spam tespitinde, yorumların içeriği tek başına yeterli olmaması nedeniyle, yorumun yapıldığı videonun adı, yazarın adı ve günün saat dilimi gibi ek bağlamsal bilgiler de dâhil edilmiştir. Yorumların, makine öğrenmesi modelleri tarafından işlenebilmesi için TF-IDF kullanılarak yorumlar sayısal vektörlere dönüştürülmüştür. TF-IDF kullanılarak her bir kelimenin önemini hesaplayarak spam ve gerçek yorumları daha iyi ayırt edebilecek bir vektör uzayı oluşturmuştur.

Bu çalışmada spam tespitine yönelik öznitelik seçiminde hem içerik tabanlı hem de bağlamsal özellikler dikkate alınmıştır. İçerik temelli öznitelikler arasında TF-IDF ile elde edilen kelime ağırlıkları yer alırken, bağlamsal öznitelikler olarak karakter sayısı ile ifade edilen yorum uzunluğu, gece, gündüz yada öğlen şeklinde ifade edilen yorumun gönderildiği saat dilimi, büyük harf oranı, noktalama işareti kullanımı ve link içerip içermeme durumu gibi göstergeler seçilmiştir. Bu öznitelikler, literatürde spam içeriklerin genellikle kısa, tekrar eden ve belirli zaman dilimlerinde yoğunlaştığı gözlemleriyle örtüşmektedir. Örneğin, spam yorumlar

çoğunlukla gece saatlerinde gönderilmekte ve link içermektedir; bu nedenle zaman ve bağlantı bilgisi ayrıştırıcı rol oynamaktadır. Modelleme sürecinde, özniteliklerin tek tek ve farklı kombinasyonlarıyla test edilmesi sonucunda, içerik ve bağlam özelliklerinin birlikte kullanıldığı yapıların en yüksek başarıyı sağladığı gözlemlenmiştir. Özellikle TF-IDF + yorum uzunluğu + saat dilimi birleşimi, hem kesinlik hem duyarlılık açısından en dengeli sonucu üretmiştir.

Son olarak, veri setinin %80'i eğitim ve %20'si test için ayrılmıştır. Çalışmada uygulanan modellerin en başarılı olabilecekleri hiper-parametreleri belirlemek amacıyla Grid Search kullanılmıştır. Grid Search, makine öğrenmesi ve derin öğrenme modelleri için en iyi hiper-parametreleri belirlemek amacıyla kullanılan sistematik bir arama yöntemidir. Grid Search, Belirlenen hiperparametrelerin farklı değerlerini içeren bir hiper-parametre uzayı oluşturur. Tüm olası hiper-parametre kombinasyonlarını tek tek dener ve her birini model üzerinde eğiterek değerlendirir ve en iyi sonucu veren hiper-parametre kombinasyonunu seçer.

Bu çalışmada kullanılan modellerin mimarileri ve optimizasyon süreçleri detaylı olarak yapılandırılmıştır. LR için belirlenen hiper-parametreler penalty: l2, C: 1, solver: liblinear ve max_iter: 200'dür. RF için belirlenen hiper-parametreler n_estimators: 200, max_depth: 20, min_samples_split: 5, min_samples_leaf: 3, max_features: sqrt ve bootstrap: True'dur. SVM için belirlenen hiper-parametreler C: 1, kernel: rbf, gamma: scale ve degree: 3'tür. XGBoost için belirlenen hiper-parametreler n_estimators: 300, max_depth: 6, learning_rate: 0.1, subsample: 0.8, colsample_bytree: 0.8, gamma: 0.1 ve lambda: 1'dir. Bi-LSTM için belirlenen hiper-parametreler embedding_dim: 128, lstm_units: 256, dropout: 0.5, batch_size: 32, epochs: 10, optimizer: adam ve loss: binary_crossentropy'dir.

BERT, 12 Transformer katmanı (layer), 768 gizli boyut (hidden size), 12 çoklu-başlıklı dikkat (multi-head attention) mekanizması ve toplamda yaklaşık 110 milyon öğrenilebilir parametre içeren bert-base-uncased sürümüyle uygulanmıştır. Bu mimari, her bir kelimeyi çift yönlü bağlamda temsil ederek, metin içerisindeki semantik ilişkileri derinlemesine analiz etme yeteneğine sahiptir. Modelin son katmanına bir sınıflandırıcı (fully connected dense layer) eklenmiş ve AdamW optimizasyon algoritması kullanılmıştır. Aşırı öğrenmeyi önlemek amacıyla weight decay=0.01 ve dropout=0.1 oranları uygulanmıştır. Öğrenme oranı (learning_rate=2e-5), eğitim stabilitesi sağlamak üzere 500 adımlık bir warmup süreciyle ayarlanmıştır.

Tüm modeller için hiperparametre optimizasyonu, sistematik bir arama stratejisi olan Grid Search yöntemiyle gerçekleştirilmiştir. Bu süreçte, her model için farklı kombinasyonlar denenmiş ve doğruluk ile F1-puanı gibi metrikler üzerinden en iyi sonuçları veren yapılandırmalar tercih edilmiştir. Örneğin, RF için n_estimators=200, max_depth=20; XGBoost için n_estimators=300, learning_rate=0.1 gibi değerler optimal performansa ulaşmada etkili olmuştur. Hiperparametrelerin bu şekilde seçilmesi, her modelin kendi yapısal avantajlarını en iyi şekilde ortaya koymasına imkân tanımıştır. Ayrıca, kullanılan modellerin avantaj ve sınırlılıkları da dikkate alınmıştır. LR hızlı ve yorumlanabilirken doğrusal sınırlara bağımlıdır. SVM yüksek doğruluk sağlar ancak büyük veri kümelerinde eğitim süresi uzundur. Bi-LSTM uzun vadeli bağımlılıkları öğrenmede etkilidir ancak yüksek hesaplama maliyeti gerektirir. BERT ise bağlamsal anlamı güçlü biçimde modelleyerek özellikle karmaşık metinler için üstün başarı sunar ancak GPU gibi güçlü donanım ihtiyacı doğurur.

3.2. Sınıflandırma Modelleri

Bu çalışmada, YouTube spam yorumlarını tespit etmek amacıyla LR, RF, SVM, XGBoost ve BiLSTM gibi yapay zekâ modellerinin sınıflandırma performansları kapsamlı bir şekilde karşılaştırılmıştır.

İstatistiksel bir yöntem olan LR, doğrusal ayrılabilen veri setlerinde iyi performans gösteren bir doğrusal sınıflandırma algoritmasıdır [20]. LR, sigmoid fonksiyonu kullanarak giriş özelliklerini 0-1 aralığında bir olasılığa dönüştürür. LR, verinin iki sınıfa ayrılmasını sağlayan bir karar sınırı öğrenir [21]. Karar sınırı, doğrusal bir fonksiyon ile belirlenir. LR, öğrenme sürecinde, negatif log-likelihood kaybını minimize eden bir optimizasyon işlemi gerçekleştirir. LR, genellikle küçük ve orta ölçekli veri setlerinde iyi çalışır ve hesaplama açısından verimlidir [22]. Ancak, doğrusal olmayan sınırlar içeren veri setlerinde yalnızca doğrusal bir karar sınırı oluşturduğu için yeteri kadar başarılı değildir. Bu nedenle, daha karmaşık ve doğrusal olmayan örüntüleri yakalayabilen SVM, karar ağaçları veya sinir ağları gibi modeller genellikle daha iyi performans gösterir. L1 (Lasso) ve L2 (Ridge) regularizasyonu ile aşırı öğrenme (overfitting) önlenebilir [23].

RF, birden fazla karar ağacının birleşiminden oluşan güçlü bir ensemble öğrenme yöntemidir. Karar ağaçlarının bağımsız olarak eğitilmesi ve sonucun çoğunluk oylaması ile belirlenmesi, modelin aşırı öğrenme riskini önemli ölçüde azaltır [24]. Torbalama yöntemi kullanılarak her ağaç farklı rastgele alt kümeler üzerinde eğitilir. Bu sayede modele çeşitlilik kazandırılır ve tek bir karar ağacına kıyasla daha genelleştirilebilir hale gelir [25]. RF, özellikle çok boyutlu ve yüksek öznitelikli veri setlerinde etkili çalışır. Öznitelik seçimi ve önemi konusunda güçlüdür. RF, gürültülü verilere karşı dayanıklıdır ancak büyük veri setlerinde eğitim süresi uzayabilir ve fazla bellek tüketebilir [26].

SVM, en iyi ayrım yapan hiper-düzlemi belirleyerek sınıflandırma yapan güçlü bir makine öğrenmesi algoritmasıdır [27]. SVM, maksimum marj prensibini kullanarak sınıflar arasındaki en büyük mesafeyi bulmaya çalışır [28]. Destek vektörleri, karar sınırına en yakın olan veri noktalarıdır ve sınıflandırma sürecinde önemli bir rol oynar. Eğer veri doğrusal olarak ayrılabilir değilse, SVM kernel trick kullanarak veriyi daha yüksek boyutlu bir uzaya taşıyabilir ve burada ayrılabilir hale getirebilir. Yaygın kullanılan çekirdek fonksiyonları Lineer, Polinom, Radial Basis Function (RBF) ve Sigmoid'dir [29]. SVM, özellikle küçük ve orta ölçekli veri setlerinde iyi çalışır ve çok fazla öznitelik içeren veri setlerinde oldukça başarılıdır. Ancak, büyük veri setlerinde eğitim süresi uzun olabilir ve çekirdek fonksiyonu seçimi modelin başarısını önemli ölçüde etkileyebilir [30].

XGBoost, karar ağaçlarını ardışık olarak eğiten, yüksek doğruluk sağlayan güçlü bir ensemble öğrenme modelidir [31]. XGBoost, Gradient Boosting algoritmasını optimize ederek hesaplama hızını artırır ve bellek kullanımını azaltır. Hata oranını minimize eden ardışık ağaç öğrenme süreci, XGBoost'u diğer klasik karar ağaçları yöntemlerinden daha güçlü hale getirir [32]. Her yeni ağaç, bir önceki ağacın hatalarını düzeltecek şekilde eğitilir. Bu süreç, modelin karmaşık ilişkileri ve doğrusal olmayan desenleri öğrenmesini sağlar [33].

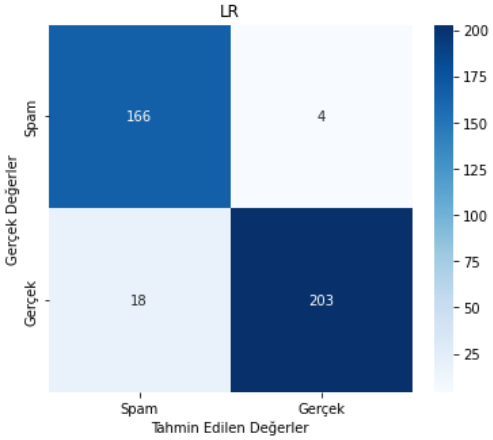
Bi-LSTM, kelime bağlamını hem ileri hem de geri yönde işleyerek klasik LSTM'den daha fazla bilgi yakalayabilen bir sinir ağı modelidir [34]. Geleneksel LSTM, uzun vadeli bağımlılıkları öğrenerek metin tabanlı verilerde başarılı sonuçlar veren bir modeldir [35]. Ancak, LSTM yalnızca geçmişten geleceğe doğru veri akışı sağlar, bu da gelecekteki kelimelerden gelen bağlamsal bilgiyi kaçırabilir. Bi-LSTM, girdileri ileri ve geri olmak üzere iki yönde işler. Bu

sayede, bir kelimenin anlamı hem önceki hem de sonraki kelimelerle birlikte değerlendirilir. Spam tespitinde, daha uzun bağlamları anlamlandırabilme yeteneği sayesinde doğruluk oranlarını artırabilir [36]. Ancak, Bi-LSTM eğitmek için büyük miktarda veri gereklidir ve hesaplama maliyeti yüksektir. GPU kullanımı olmadan, eğitim süresi oldukça uzun olabilir. Spam tespiti, duygu analizi ve metin tabanlı sıralı veri işlemede en güçlü modellerden biridir [37].

BERT, Dönüştürücü mimarisi sayesinde geleneksel tek yönlü modellerin aksine metin içerisindeki kelimelerin hem solundaki hem de sağındaki bağlamı aynı anda analiz edebilme özelliğine sahiptir [38]. Bu çift yönlü öğrenme yaklaşımı, özellikle bağlamsal anlam farklılıklarının önemli olduğu metin sınıflandırma görevlerinde büyük avantaj sağlamaktadır [39]. BERT, büyük çaplı dil verileri üzerinde önceden eğitilmiş olup, transfer öğrenme ile daha küçük veri kümeleri üzerinde kolayca yeniden uyarlanabilir. Bu yönüyle, veri miktarının sınırlı olduğu durumlarda bile yüksek başarı elde edilebilir [40]. Spam tespiti gibi anlam karmaşıklığı ve dilsel çeşitlilik barındıran görevlerde, BERT'in semantik açıdan zengin temsil gücü oldukça etkilidir. Özellikle ironi, mecaz ve dolaylı ifadelerin bulunduğu yorumların sınıflandırılmasında, kelime seviyesinde değil bağlam seviyesinde analiz yapabilmesi sayesinde klasik yöntemlerin ötesine geçmektedir.

4. DENEYSEL SONUÇLAR

Bu çalışmada, LR, RF, SVM, XGBoost ve Bi-LSTM modellerinin spam tespitindeki performansları test edilmiştir. Modellerin başarısı, doğruluk, kesinlik, duyarlılık ve F1-puanı metrikleri kullanılarak değerlendirilmiştir. Şekil 5 ve Tablo 2'de LR için karışıklık matrisi ve deneysel sonuçlar görülmektedir.



Şekil 5. LR için karışıklık matrisi

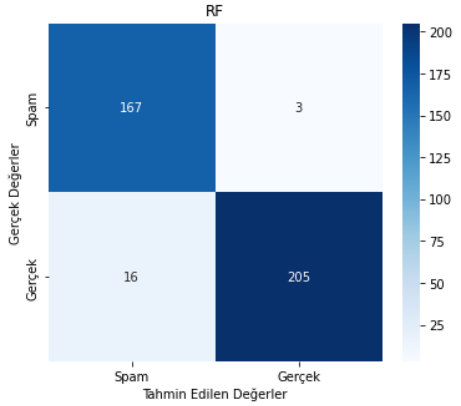
Şekil 5'te görüldüğü gibi LR, 170 adet spam yorumdan 166'sını ve 221 spam olmayan yorumdan 203'ünü doğru bir şekilde sınıflandırmıştır. LR, 391 yorumun 369'unu doğru, 22'sini ise yanlış sınıflandırmıştır.

Tablo 2. LR için deneysel sonuçlar

Sınıf	Doğruluk	Kesinlik	Duyarlılık	F1-puanı
Gerçek	%94,4	%98,1	%91,9	%94,9
Spam		%90,2	%97,6	%93,8
Makro ortalama		%94,1	%94,8	%94,4
Ağırlıklı ortalama		%94,2	%94,4	%94,3

Tablo 2'de görüldüğü gibi LR'nin sınıflandırma doğruluğu %94,4'tür. Spam sınıfı için %97,6 olan duyarlılık değeri, LR'nin spam olan yorumları büyük oranda doğru tespit ettiğini göstermektedir. Ancak, spam sınıfı için %90,2 olan kesinlik değeri, LR'nin bazı gerçek yorumları spam olarak sınıflandırdığını göstermektedir.

Şekil 6 ve Tablo 3'te RF için karışıklık matrisi ve deneysel sonuçlar görülmektedir.



Şekil 6. RF için karışıklık matrisi

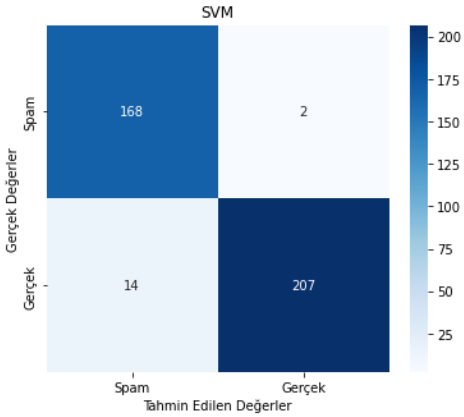
Şekil 6'da görüldüğü gibi RF, 170 adet spam yorumdan 167'sini ve 221 spam olmayan yorumdan 205'ini doğru bir şekilde sınıflandırmıştır. RF, 391 yorumun 372'sini doğru, 19'unu ise yanlış sınıflandırmıştır.

Tablo 3. RF için deneysel sonuçlar

Sınıf	Doğruluk	Kesinlik	Duyarlılık	F1-puanı
Gerçek	%95,3	%98,5	%92,8	%95,6
Spam		%91,2	%98,2	%94,6
Makro ortalama		%94,8	%95,5	%95,1
Ağırlıklı ortalama		%95,0	%95,3	%95,2

Tablo 3'te görüldüğü gibi RF'nin sınıflandırma doğruluğu %95,3'tür. Spam sınıfı için %98,2 duyarlılık değeri, RF'nin spam olan yorumların büyük çoğunluğunu doğru tespit ettiğini göstermektedir. Spam olmayan yorumlar için %92,8 kesinlik oranı LR'ye göre yüksektir ve RF'ın yanlış pozitifleri azalttığını göstermektedir.

Şekil 7 ve Tablo 4'te SVM için karışıklık matrisi ve deneysel sonuçlar görülmektedir.



Şekil 7. SVM için karışıklık matrisi

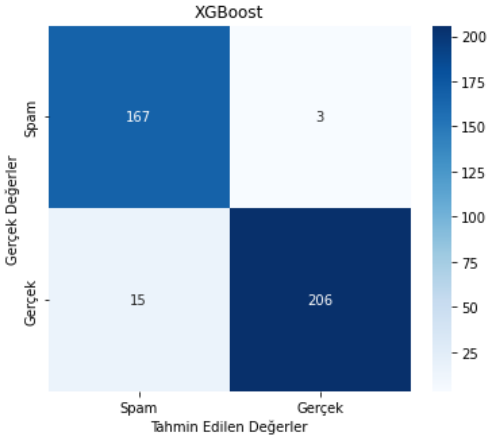
Şekil 7'de görüldüğü gibi SVM, 170 adet spam yorumdan 168'ini ve 221 spam olmayan yorumdan 207'ini doğru bir şekilde sınıflandırmıştır. SVM, 391 yorumun 375'ini doğru, 16'sını ise yanlış sınıflandırmıştır.

Tablo 4. SVM için deneysel sonuçlar

Sınıf	Doğruluk	Kesinlik	Duyarlılık	F1-puanı
Gerçek	%96,2	%99,0	%93,7	%96,3
Spam		%92,3	%98,8	%95,5
Makro ortalama		%95,6	%96,3	%95,9
Ağırlıklı ortalama		%95,8	%96,2	%96,0

Tablo 4'te görüldüğü gibi SVM'nin sınıflandırma doğruluğu %96,2'dir. Spam sınıfı için %98,8 duyarlılık değeri, SVM'nin spam olan yorumların büyük çoğunluğunu doğru tespit ettiğini göstermektedir. Spam olmayan yorumlar için %99,0 kesinlik değeri, SVM'nin yanlış pozitif oranının son derece düşük olduğunu göstermektedir. Makro ve ağırlıklı ortalama değerleri de %96 seviyelerindedir ve SVM tüm sınıflarda oldukça dengeli bir performans göstermiştir.

Şekil 8 ve Tablo 5'te XGBoost için karışıklık matrisi ve deneysel sonuçlar görülmektedir.



Şekil 8. XGBoost için karışıklık matrisi

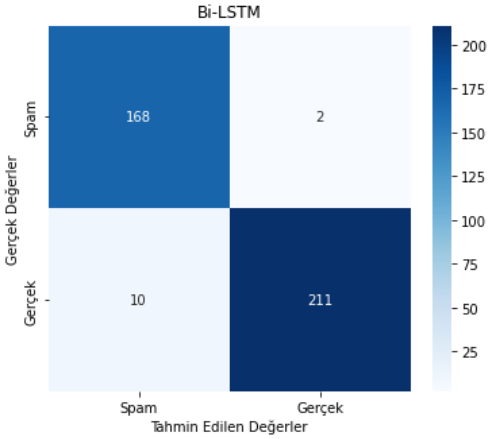
Şekil 8'de görüldüğü gibi XGBoost, 170 adet spam yorumdan 167'sini ve 221 spam olmayan yorumdan 206'sını doğru bir şekilde sınıflandırmıştır. XGBoost, 391 yorumun 373'ünü doğru, 18'ini ise yanlış sınıflandırmıştır.

Tablo 5. XGBoost için deneysel sonuçlar

Sınıf	Doğruluk	Kesinlik	Duyarlılık	F1-puanı
Gerçek	%95,6	%98,6	%93,2	%95,8
Spam		%91,7	%98,2	%94,8
Makro ortalama		%95,1	%95,7	%95,3
Ağırlıklı ortalama		%95,2	%95,6	%95,4

Tablo 5'te görüldüğü gibi XGBoost'un sınıflandırma doğruluğu %95,6'dır. Spam sınıfı için %98,2 olan duyarlılık değeri, XGBoost'un spam yorumları büyük ölçüde doğru bir şekilde tespit ettiğini göstermektedir. Spam olmayan yorumlar için %98,6 kesinlik değeri, yanlış pozitif oranının oldukça düşük olduğunu göstermektedir. Makro ve ağırlıklı ortalama metrikleri %95 seviyelerindedir ve XGBoost'un dengeli bir şekilde çalıştığını göstermektedir. XGBoost, SVM ile benzer sonuçlara sahiptir ancak XGBoost daha dengeli bir model sunmaktadır.

Şekil 9 ve Tablo 6'da Bi-LSTM için karışıklık matrisi ve deneysel sonuçlar görülmektedir.



Şekil 9. Bi-LSTM için karışıklık matrisi

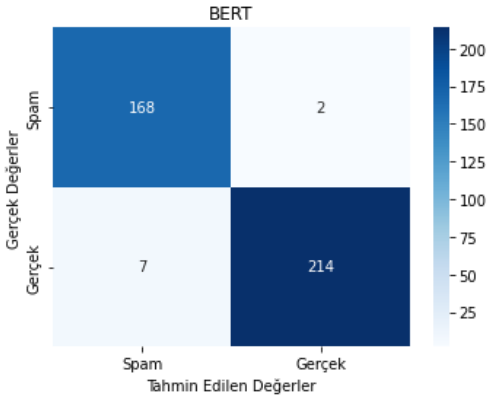
Şekil 9'da görüldüğü gibi Bi-LSTM, 170 adet spam yorumdan 168'ini ve 221 spam olmayan yorumdan 211'ini doğru bir şekilde sınıflandırmıştır. Bi-LSTM, 391 yorumun 379'unu doğru, 12'sini ise yanlış sınıflandırmıştır.

Tablo 6. Bi-LSTM için deneysel sonuçlar

Sınıf	Doğruluk	Kesinlik	Duyarlılık	F1-puanı
Gerçek	%97,1	%99,1	%95,5	%97,3
Spam		%94,4	%98,8	%96,6
Makro ortalama		%96,8	%97,2	%96,9
Ağırlıklı ortalama		%97,0	%97,1	%97,0

Tablo 6'da görüldüğü gibi Bi-LSTM'in sınıflandırma doğruluğu %97,1'dir. Bi-LSTM, karşılaştırılan modellerden daha yüksek doğruluk oranına ulaşarak en başarılı model olmuştur. Spam sınıfı için %98,8 duyarlılık değeri, Bi-LSTM'in spam olan yorumları neredeyse tamamen doğru tespit ettiğini göstermektedir. Spam olmayan yorumlar için %99,1 kesinlik değeri, Bi-LSTM'in hatalı bir şekilde spam olarak tespit edilen yorumları en aza indirdiğini göstermektedir. %97 seviyelerinde olan makro ve ağırlıklı ortalama metrikleri Bi-LSTM'in tüm sınıflarda daha dengeli bir performans sunduğunu göstermektedir.

Şekil 10 ve Tablo 7'de BERT için karışıklık matrisi ve deneysel sonuçlar görülmektedir.



Şekil 10. BERT için karışıklık matrisi

Şekil 10'da görüldüğü gibi BERT, 170 adet spam yorumdan 168'ini ve 221 spam olmayan yorumdan 214'ünü doğru bir şekilde sınıflandırmıştır. BERT, 391 yorumun 382'sini doğru, 9'unu ise yanlış sınıflandırmıştır.

Tablo 7. BERT için deneysel sonuçlar

Sınıf	Doğruluk	Kesinlik	Duyarlılık	F1-puanı
Gerçek	%97,7	%99,1	%96,8	%97,9
Spam		%96,0	%98,8	%97,4
Makro ortalama		%97,5	%97,8	%97,7
Ağırlıklı ortalama		%97,7	%97,7	%97,7

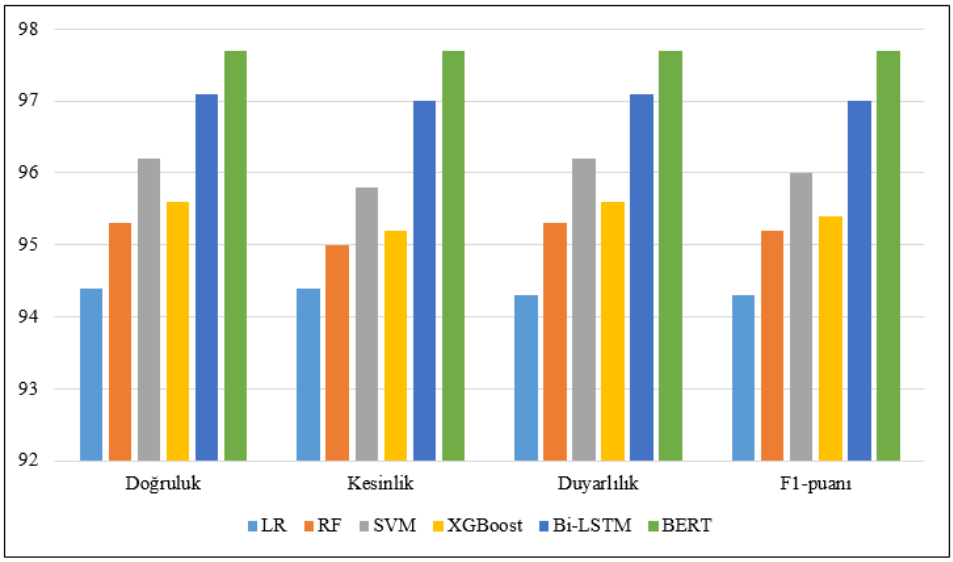
Tablo 7'de görüldüğü gibi BERT'in sınıflandırma doğruluğu %97,7'dir. BERT, hem spam tespiti hem de gerçek mesaj ayırt etme konusunda yüksek başarı göstermektedir. En güçlü yönü, %98,8 duyarlılık değeriyle spam mesajları neredeyse eksiksiz tespit edebilmesidir. Sınıf bazında incelendiğinde, Gerçek sınıfı için %99,1 kesinlik ve %96,8 duyarlılık, modelin yanlış pozitifleri minimum düzeyde tuttuğu ve bu sınıfı büyük oranda doğru tanıdığı görülmektedir. Spam sınıfı için %98,8 duyarlılık değeri, modelin spam içerikleri başarıyla ayırt edebildiğini göstermektedir. %96,0 kesinlik değeri ise bu tahminlerin çoğunun doğru olduğunu göstermektedir. Makro ve ağırlıklı ortalama metriklerinin birbirine çok yakın olması, modelin her iki sınıfa da dengeli şekilde öğrenme sağladığını ve veri setinde sınıf dengesizliğinden kaynaklı bir sapmanın oluşmadığını göstermektedir. F1-puanlarının her iki sınıf için de %97'nin üzerinde olması, kesinlik ve duyarlılık metrikleri arasında güçlü bir denge kurulduğunu göstermektedir.

Tablo 8 ve Şekil 11'de ağırlıklı ortalamaya göre karşılaştırmalı deneysel sonuçlar görülmektedir.

Tablo 8. Karşılaştırmalı deneysel sonuçlar

Model	Doğruluk	Kesinlik	Duyarlılık	F1-puanı
LR	%94,4	%94,4	%94,3	%94,3
RF	%95,3	%95,0	%95,3	%95,2
SVM	%96,2	%95,8	%96,2	%96,0
XGBoost	%95,6	%95,2	%95,6	%95,4
Bi-LSTM	%97,1	%97,0	%97,1	%97,0
BERT	%97,7	%97,7	%97,7	%97,7

Tablo 8 ve Şekil 11'de görüldüğü gibi BERT, tüm değerlendirme metrikleri açısından en başarılı performansa ulaşmıştır. BERT'in, sahip olduğu %97,7 doğruluk ile hem spam hem de gerçek yorumları ayırt etmede etkili olduğu görülmektedir.



Şekil 11. Karşılaştırmalı deneysel sonuçlar

LR, %94,4 doğruluk ile en düşük başarıyı gösteren model olmuştur. LR, doğrusal bir model olduğu için ve karmaşık yapıya sahip metin verilerinde yeterince iyi genelleme yapamamaktadır. RF, %95,3 doğrulukla LR'den daha iyi performans göstermiştir. RF, çok sayıda karar ağacı kullanarak karmaşık karar sınırlarını öğrenebildiği için metin sınıflandırmada LR'den daha başarılı olmuştur. SVM, %96,2 doğrulukla RF ve LR'ye göre daha başarılı olmuştur. Özellikle %96.2 duyarlılık değeriyle spam tespitinde oldukça başarılıdır. XGBoost, %95,6 doğruluk ile RF'den daha iyi ancak SVM'den daha düşük bir başarı göstermiştir. XGBoost, boosting mekanizması sayesinde RF'a kıyasla daha hassas karar sınırları belirleyebilse de, SVM ve Bi-LSTM kadar yüksek doğruluk sağlayamamıştır. En yüksek başarı, %97,1 doğruluk ile Bi-LSTM tarafından elde edilmiştir. Bi-LSTM, çift yönlü uzun kısa süreli bellek mekanizması sayesinde metinlerdeki kelime ilişkilerini daha iyi öğrenerek spam ve gerçek yorumları yüksek doğrulukla sınıflandırmıştır. En yüksek doğruluk oranı ise %97,7 ile BERT modeline aittir. BERT, önceden eğitilmiş bağlamsal dil temsil yeteneği sayesinde metinlerin anlam derinliğini öğrenmiş ve hem spam hem de gerçek yorumları yüksek hassasiyetle sınıflandırarak en başarılı model olmuştur.

Deneysel sonuçlar, basit yöntemlerden karmaşık yöntemlere doğru gidildikçe spam tespit performansının arttığını göstermektedir. LR gibi basit modeller, hızlı ve yorumlanabilir olmalarına rağmen karmaşık dil örüntülerini öğrenmede sınırlı kalmaktadır. RF, SVM ve XGBoost gibi ileri seviye klasik yöntemler, daha güçlü karar sınırları öğrenebilmiş ve doğruluk oranlarını artırmıştır. En yüksek başarı, karmaşık dil bağlamlarını anlayabilen Bi-LSTM ve BERT modeline ait olup, özellikle bağlamsal ve yapısal olarak karmaşık spam yorumlarını ayırt etmede etkili olmuştur.

5. SONUÇLAR

Günümüzde dijital platformlarda spam içeriklerin yaygınlaşması, kullanıcı deneyimini olumsuz etkilemekte ve yanıltıcı bilgilere maruz kalma riskini artırmaktadır. Özellikle YouTube gibi büyük kullanıcı kitlesine sahip sosyal medya platformlarında, spam yorumların manuel olarak tespit edilmesi mümkün olmadığından, otomatik spam tespit sistemlerine duyulan ihtiyaç giderek artmaktadır. Bu çalışmada, YouTube yorumlarındaki spam içerikleri tespit etmek amacıyla farklı makine öğrenmesi ve derin öğrenme yöntemlerinin karşılaştırmalı bir analizi sunulmuştur. LR, RF, SVM, XGBoost, Bi-LSTM ve BERT modellerinin performansları

doğruluk, kesinlik, duyarlılık ve F1-puanı metrikleri kullanılarak değerlendirilmiştir. Deneysel çalışmalar, BERT modelinin %97,7 doğruluk oranı ile karşılaştırılan tüm modeller arasında en yüksek başarıyı sağladığını ortaya koymuştur. BERT'in, önceden büyük ölçekli metin verileri üzerinde eğitilmiş derin dil temsilleri sayesinde, bağlamsal ilişkileri güçlü bir şekilde modelleyebildiği ve bu nedenle hem spam hem de gerçek yorumları yüksek isabetle sınıflandırabildiği gözlemlenmiştir. Bi-LSTM modeli ise, kelimeler arasındaki ilişkileri çift yönlü olarak analiz edebilme yeteneği sayesinde spam tespitinde dikkat çeken bir başarı göstermiştir. SVM ve XGBoost modelleri yüksek doğruluk oranları elde etmiş olsalar da, metin tabanlı verilerdeki uzun vadeli bağımlılıkları öğrenme konusunda Bi-LSTM ve özellikle BERT kadar etkili olamamıştır.

Geleneksel kural tabanlı filtreleme sistemlerinin aksine yapay zekâ tabanlı modeller, spam içerikleri bağlamsal ve anlamsal özellikleri göz önünde bulundurarak daha doğru bir şekilde sınıflandırabilmektedir. Bu çalışmada kullanılan TF-IDF yöntemi ile metinlerin vektörleştirilmesi, makine öğrenmesi ve derin öğrenme modellerinin spam tespiti için uygun bir şekilde eğitilmesini sağlamış ve özellikle XGBoost ve SVM gibi modellerin performansını artırmıştır. Ancak, BERT ve Bi-LSTM dilin daha derin bağlamını anlayarak daha güçlü bir sınıflandırma performansına sahip olmuştur.

Deneysel sonuçlar, özellikle BERT ve Bi-LSTM modellerinin Türkçe spam tespitinde yüksek başarı oranlarıyla öne çıktığını göstermektedir. Bu sonuçlar, ALBERT+BLSTM gibi karma modellerle %95 üzeri doğruluk elde edilen güncel çalışmalarla benzerlik göstermekte, ancak çalışmada yalnızca yorum metni değil, saat dilimi ve yorum uzunluğu gibi bağlamsal özniteliklerin de modele dâhil edilmesi sayesinde daha güçlü sınıflandırma performansları elde edilmiştir. Ayrıca, önceki birçok çalışmada yalnızca İngilizce veya sınırlı boyuttaki veri setleri kullanılırken, bu çalışmada Türkçe içerikli YouTube yorumlarından oluşturulan özgün ve dengeli bir veri kümesi kullanılmıştır. Bununla birlikte, model eğitimi sırasında yüksek doğruluk elde edilse de, gerçek zamanlı uygulamalara yönelik testlerin yapılmaması çalışmanın bir sınırlılığı olarak değerlendirilebilir. Ayrıca, yorumlardaki ironi, kinaye gibi dilsel inceliklerin modellenememesi ve veri setinin yalnızca YouTube'a özgü olması genel geçerlilik açısından bir başka sınırlılıktır. Buna karşın, çok sayıda makine öğrenmesi ve derin öğrenme algoritmasının sistematik biçimde karşılaştırılması ve Türkçe metinler üzerinde uygulanması çalışmanın güçlü yönleri arasında yer almaktadır.

Elde edilen bulgular, sosyal medya platformlarında kullanılan mevcut spam filtreleme sistemlerinin, yapay zekâ temelli yöntemlerle önemli ölçüde geliştirilebileceğini göstermektedir. Özellikle BERT ve Bi-LSTM gibi bağlamsal öğrenme yeteneğine sahip modeller, sadece anahtar kelimelere değil, yorumun anlam bütünlüğüne ve bağlamına dayalı olarak spam tespiti gerçekleştirebildiğinden, daha az yanlış pozitif ve negatif sonuç üretmektedir. Bu da kullanıcıların yanlışlıkla engellenmesini ya da spam içeriklerin gözden kaçmasını önlemeye katkı sağlar. Geliştirilen modeller, YouTube gibi büyük ölçekli platformlara entegre edilerek, yorum denetleme süreçlerinin otomatikleştirilmesi ve gerçek zamanlı spam filtreleme altyapılarının oluşturulmasına olanak tanıyabilir. Bu sayede kullanıcı deneyiminin iyileştirilmesinin yanı sıra platformların bilgi güvenliğinin güçlendirilmesi sağlanabilir.

Gelecek çalışmalarda, spam tespit sistemlerinin performansını artırmak amacıyla RoBERTa, XLNet ve DeBERTa gibi farklı dil modelleri ve çok dilli veri setleri üzerinde incelemeler yapılabilir. Ayrıca, yorumların semantik yapısını daha derinlemesine analiz edebilecek attention tabanlı hibrit modeller geliştirilebilir. Spam tespitine yönelik modellerin

açıklanabilirliğini artırmak adına SHAP veya LIME gibi Açıklanabilir Yapay Zekâ (Explainable AI-XAI) yöntemlerinin entegre edilmesi de önemli katkılar sağlayacaktır. Gerçek zamanlı spam algılama senaryoları veya video içeriği ile yorumların ilişkilendirilmesi gibi bağlamsal analizler ise uygulama odaklı gelecek çalışmalara zemin hazırlayabilir.

KAYNAKLAR

[1] Susanto H, Fang Yie L, Mohiddin F, Rahman Setiawan A A, Haghi P K, Setiana D. Revealing social media phenomenon in time of COVID-19 pandemic for boosting start-up businesses through digital ecosystem. Applied system innovation. 2021;4(1).

[2] Humprecht E, Kessler S H. Unveiling misinformation on YouTube: examining the content of COVID-19 vaccination misinformation videos in Switzerland. Frontiers in Communication. 2024; 9.

[3] Lakshmi M S, Rani A S, Divya T S, Shravani J. Dynamic Spam Detection in Social Networks: Leveraging Convex Nonnegative Matrix Factorization for Enhanced Accuracy and Scalability. International Journal of Computer Engineering in Research Trends. 2024; 11(4), 1-11.

[4] Gongane V U, Munot M V, Anuse A D. Detection and moderation of detrimental content on social media platforms: current status and future directions. Social Network Analysis and Mining. 2022; 12(1).

[5] Wani M A, ElAffendi M, Shakil K A. AI-Generated Spam Review Detection Framework with Deep Learning Algorithms and Natural Language Processing. Computers. 2024; 13(10).

[6] Ahmed N, Amin R, Aldabbas H, Koundal D, Alouffi B, Shah T. Machine learning techniques for spam detection in email and IoT platforms: analysis and research challenges. Security and Communication Networks. 2022; 2022(1).

[7] Akinyelu A A. Advances in spam detection for email spam, web spam, social network spam, and review spam: ML-based and nature-inspired-based techniques. Journal of Computer Security. 2021; 29(5), 473-529.

[8] Al Saidat M R, Yerima S Y, Shaalan K. Advancements of SMS Spam Detection: A Comprehensive Survey of NLP and ML Techniques. Procedia Computer Science. 2024; 244, 248-259.

[9] Al-Adhaileh M H, Alsaade F W. Detecting and Analysing Fake Opinions Using Artificial Intelligence Algorithms. Intelligent Automation & Soft Computing. 2022; 32(1).

[10] Shinde S A, Pawar R R, Jagtap A A, Tambewagh P A, Rajput P U, Mali M K, Mulik S V. Deceptive opinion spam detection using bidirectional long short-term memory with capsule neural network. Multimedia Tools and Applications. 2024; 83(15), 45111-45140.

[11] Sinhal A, Maheshwari M. An Extensive Review on Contemporary Analysis of Comment Filtration of YouTube Videos Using Machine Learning Techniques. International Journal of Emerging Technology and Advanced Engineering. 2022; 12(9), 130-143.

[12] Shirzadova S, Uysal A K. Türkçe YouTube Yorumları Üzerinde Spam Filtreleme. Düzce Üniversitesi Bilim ve Teknoloji Dergisi. 2022; 10(4), 1793-1810.

[13] Baktır N, Atay Y. Makine Öğrenmesi Yaklaşımlarının Spam-Mail Sınıflandırma Probleminde Karşılaştırmalı Analizi. Bilişim Teknolojileri Dergisi. 2022; 15(3), 349-364.

[14] Bakır R, Erbay H, Bakır H. ALBERT4Spam: a novel approach for spam detection on social networks. Bilişim Teknolojileri Dergisi. 2024; 17(2), 81-94.

[15] Güven Z A. Türkçe e-postalarda spam tespiti için makine öğrenme yöntemlerinin ve dil modellerinin analizi. Avrupa Bilim ve Teknoloji Dergisi 2023; 47, 1-6.

[16] Şengel Ö. A comparative analysis of learning techniques in the context of Turkish spam detection. Batman Üniversitesi Yaşam Bilimleri Dergisi. 2024; 14(1), 43-56.

[17] Sam'an M, Imaddudin K. Hybrid deep learning model for YouTube spam comment detection. International Journal of Electrical and Computer Engineering (IJECE). 2024; 14(3), 3313-3319.

[18] Airlangga G. Spam Detection in YouTube Comments Using Deep Learning Models: A Comparative Study of MLP, CNN, LSTM, BiLSTM, GRU, and Attention Mechanisms. MALCOM: Indonesian Journal of Machine Learning and Computer Science. 2024; 4(4), 1533-1538.

[19] Waheed A. YouTube Yorumları Spam Veri Kümesi [İnternet]. Kaggle; [alıntılanma tarihi 6 Mart 2025]. Erişim adresi: https://www.kaggle.com/datasets/ahsenwaheed/youtube-comments-spam-dataset/data

[20] Bektaş J. EKSL: An effective novel dynamic ensemble model for unbalanced datasets based on LR and SVM hyperplane-distances. Information Sciences. 2022; 597, 182-192.

[21] Al-Najjar H A, Pradhan B, Kalantar B, Sameen M I, Santosh M, Alamri A. Landslide susceptibility modeling: an integrated novel method based on machine learning feature transformation. Remote Sensing. 2021; 13(16).

[22] Kim G, Yang S M, Kim D M, Choi J G, Lim S, Park H W. Developing a deep learning-based uncertainty-aware tool wear prediction method using smartphone sensors for the turning process of Ti-6Al-4V. Journal of Manufacturing Systems. 2024; 76, 133-157.

[23] Nwosu A, Aimufua G I O, Ajayi B A, Olalere M. The Impact of Regularization on Linear Regression Based Model. Journal of Artificial Intelligence and Computer Science. 2024; 1(1).

[24] Arabameri A, Chandra Pal S, Rezaie F, Chakrabortty R, Saha A, Blaschke T, Thi Ngo P T. Decision tree based ensemble machine learning approaches for landslide susceptibility mapping. Geocarto International. 2022; 37(16), 4594-4627.

[25] Sesa O, Haikal A Y, Elhosseini M A, Gad H H. Smart Bagged Tree-based Classifier optimized by Random Forests (SBT-RF) to Classify Brain-Machine Interface

Data. International journal of electrical and computer engineering systems. 2022; 13(10), 895-908.

[26] Jagannath A, Jagannath J, Kumar P S P V. A comprehensive survey on radio frequency (RF) fingerprinting: Traditional approaches, deep learning, and open challenges. Computer Networks. 2022; 219.

[27] Chandra M A, Bedi S S. Survey on SVM and their application in image classification. International Journal of Information Technology. 2021; 13(5), 1-11.

[28] Lai Z, Chen X, Zhang J, Kong H, Wen J. Maximal margin support vector machine for feature representation and classification. IEEE Transactions on Cybernetics. 2023; 53(10), 6700-6713.

[29] Negi H S, Dimri S C, Kumar B, Ram M. Support vector machine and classification, kernel trick for separating of data points. Mathematics in Engineering, Science & Aerospace (MESA). 2024; 15(2).

[30] Ding X, Liu J, Yang F, Cao J. Random radial basis function kernel-based support vector machine. Journal of the Franklin Institute. 2021; 358(18), 10121-10140.

[31] Natras R, Soja B, Schmidt M. Ensemble machine learning of random forest, AdaBoost and XGBoost for vertical total electron content forecasting. Remote Sensing. 2022; 14(15), 3547.

[32] Demir S, Sahin E K. An investigation of feature selection methods for soil liquefaction prediction based on tree-based ensemble algorithms using AdaBoost, gradient boosting, and XGBoost. Neural Computing and Applications. 2023; 35(4), 3173-3190.

[33] Ji S, Wang X, Lyu T, Liu X, Wang Y, Heinen E, Sun Z. Understanding cycling distance according to the prediction of the XGBoost and the interpretation of SHAP: A non-linear and interaction effect analysis. Journal of Transport Geography. 2022; 103.

[34] Shoubaki H, Abdallah S, Shaalan K. Deep Learning Techniques for Identifying Poets in Arabic Poetry: A Focus on LSTM and Bi-LSTM. Procedia Computer Science. 2024; 244, 461-470.

[35] Zhou Z G. Research on sentiment analysis model of short text based on deep learning. Scientific Programming. 2022; 2022(1), 2681533.

[36] Ahmed S, Saif A S, Hanif M I, Shakil M M N, Jaman M M, Haque M M U, Sabbir H M. Att-BiL-SL: Attention-based Bi-LSTM and sequential LSTM for describing video in the textual formation. Applied sciences. 2021; 12(1), 317.

[37] Odera D, Odiaga G. A comparative analysis of recurrent neural network and support vector machine for binary classification of spam short message service. World Journal of Advanced Engineering Technology and Sciences. 2023; 9(1), 127-152.

[38] Zhou C, Li Q, Li C, Yu J, Liu Y, Wang G, Sun L. A comprehensive survey on pretrained foundation models: A history from bert to chatgpt. International Journal of Machine Learning and Cybernetics, 2024, 1-65.

[39] Gao Z, Feng A, Song X, Wu X. Target-dependent sentiment classification with BERT. Ieee Access, 2019; 7, 154290-154299.
[40] Rosso M M, Marasco G, Aiello S, Aloisio A, Chiaia B, Marano G C. Convolutional networks and transformers for intelligent road tunnel investigations. Computers & Structures, 2023; 275.

BİNGÖL ÜNİVERSİTESİ
TEKNİK BİLİMLER DERGİSİ

Bingöl Üniversitesi Teknik Bilimler Dergisi
Bingol University Journal of Technical Science
e-ISSN: 2757-6884

Cilt:6, Sayı: 1, Sayfa: 51-63, 2025
Araştırma Makalesi

Volume:6, Number: 4, Page: 51-63, 2025
Research Article

Doi: 10.5281/zenodo.15719600

Weight Optimization of Weighted Naive Bayes Classifier for Efficient Classification

Gamzepelin AKSOY[1]*, Murat KARABATAK [2]

[1] Isparta University of Applied Sciences, Department of Computer Engineering, Isparta/Türkiye.

ORCID: 0000-0002-5328-2983, E-mail: gamzepelinaksoy@isparta.edu.tr

[2] Fırat University, Department of Software Engineering, Elazığ/Türkiye.

ORCID: 0000-0002-6719-7421, E-mail: mkarabatak@firat.edu.tr

(Alınış/Arrival: 02.06.2025, Kabul/Acceptance: 13.06.2025, Yayınlanma/Published: 25.06.2025)

Abstract

The Weighted Naive Bayes (WNB) classifier is an effective classification method based on the Naive Bayes algorithm; however, determining the appropriate weights within this algorithm remains a significant challenge. The number of optimization based WNB applications in the existing literature is quite limited. Grid Search methods used for weight determination often suffer from high computational costs and an inability to reach the global optimum. Although the Fast Weighted Naive Bayes classifier offers faster solutions, it remains limited in terms of classification performance. Therefore, optimizing weights is of great importance in achieving both computational efficiency and high classification accuracy. In this study, Genetic Algorithm (GA) and Particle Swarm Optimization (PSO) methods were employed to optimize the weights of the WNB algorithm. The proposed GAW-NB and PSOW-NB models were evaluated on five different datasets using 5-fold cross-validation. The experimental results demonstrated that both methods achieved significant improvements in terms of both execution time and classification performance.

Keywords: Weighted Naive Bayes, Classification, Genetic Algorithm, Particle Swarm Optimization.

Ağırlıklı Sade Bayes Sınıflandırıcısında Etkili Sınıflandırma İçin Ağırlıkların Optimizasyonu

Özet

Ağırlıklı Naive Bayes (WNB) sınıflandırıcısı, Naive Bayes algoritmasına dayanan etkili bir sınıflandırma yöntemidir ancak bu algoritmada ağırlıkların uygun biçimde belirlenmesi önemli bir problem olarak öne çıkmaktadır. Mevcut literatürde optimizasyon tabanlı WNB uygulamalarının sayısı oldukça sınırlıdır. Izgara arama yöntemi ile yapılan ağırlık belirleme süreçleri, yüksek hesaplama maliyeti ve küresel optimuma ulaşamama gibi dezavantajlara sahiptir. Hızlı Ağırlıklı Naive Bayes sınıflandırıcısı daha kısa sürede çözüm sunsa da sınıflandırma performansı açısından sınırlı kalmaktadır. Bu nedenle, ağırlıkların optimize edilmesi hem zaman verimliliği hem de sınıflandırma doğruluğu açısından büyük önem taşımaktadır. Bu çalışmada, WNB algoritmasındaki ağırlıkların optimizasyonu için Genetik

Algoritma (GA) ve Parçacık Sürü Optimizasyonu (PSO) yöntemleri kullanılmış; geliştirilen GAW-NB ve PSOW-NB modelleri beş farklı veri kümesi üzerinde 5 katlı çapraz doğrulama yöntemiyle test edilmiştir. Deneysel sonuçlar, her iki yöntemin de hem işlem süresi hem de sınıflandırma başarımı açısından anlamlı iyileşmeler sağladığını göstermektedir.

Anahtar Kelimeler: Ağırlıklı Sade Bayes, Sınıflandırma, Genetik Algoritma, Parçacık Sürü Optimizasyonu.

1. INTRODUCTION

Recent technological advancements have led to a significant increase in data generation across nearly all areas of the digital world. The broader reach of internet access, the integration of mobile devices as an indispensable part of daily life, and innovations such as the Internet of Things (IoT) have caused an exponential growth in data volume. This surge has made the accurate analysis of data and its transformation into meaningful insights inevitable, thereby increasing the importance of fields such as data analytics, Artificial İntelligence (AI), and big data.

Data mining constitutes a foundational methodology employed to extract meaningful patterns and knowledge from large-scale datasets. Core techniques within this domain include association rule mining, clustering, and classification. Among these, association rule mining is particularly instrumental in identifying and elucidating relationships among diverse attributes within databases, thereby revealing the underlying structure of such associations [1].

Clustering aims to generate meaningful data segments by grouping similar records within large-scale databases. While this technique forms groups based on similarities among data points, classification, on the other hand, seeks to assign new data instances to the correct categories using pre-labeled data. Classification is a supervised learning approach that analyzes the relationship between features and class labels in the training data during the model-building phase. Subsequently, this trained model is used to predict the class labels of test data, where only the features are known.

In the literature, classification problems have been extensively studied, and the algorithms developed for such tasks have found wide applicability across various domains. The Naive Bayes classification algorithm, as one of the statistical classification techniques, is widely favored by researchers due to its computational efficiency and speed. Ravinder et al. emphasized that web data mining plays a crucial role in extracting meaningful information from large-scale textual data, and demonstrated that the Naive Bayes algorithm provides an effective method for text classification and categorization in this context [2]. Jain et al. developed a model for fake news detection by employing natural language processing (NLP) techniques and machine learning methods. In their study, it was observed that the Naive Bayes and Random Forest algorithms demonstrated superior performance compared to other approaches [3]. Arrazyan et al. conducted a study to examine the effectiveness of the Naive Bayes algorithm in predicting diabetes. In their experiments, various training-to-testing data split ratios were assessed, and the highest classification accuracy 88.16% was recorded when 65% of the data was allocated for training and 35% for testing [4].

Research aimed at improving the performance of Bayes classifiers suggests that weighting techniques can offer an effective strategy. In particular, incorporating weighting schemes into

the Naive Bayes classifier can enhance the model's flexibility and lead to improved classification accuracy [5]. However, determining the optimal weight values constitutes a critical optimization problem that directly influences classification accuracy. In the literature, the grid search method is commonly employed for this purpose. Nevertheless, due to its reliance on nested loops, this method incurs high computational costs when applied to large datasets. This challenge underscores the need for more efficient approaches that can reduce processing time while maintaining classification performance. In this context, optimization algorithms offer a promising alternative for the effective determination of weight values [6]. In the study conducted by Lin et al., the Particle Swarm Optimization (PSO) algorithm was employed to optimize the weights in the Weighted Naive Bayes classifier, and its performance was evaluated against the standard Naive Bayes algorithm using multiple benchmark datasets [7].

To date, numerous optimization algorithms have been developed and applied to various types of problems. In particular, metaheuristic optimization techniques have emerged as powerful methods that provide effective solutions to nonlinear problems. Among these techniques, the most well-known and widely used algorithms are the Genetic Algorithm (GA) and Particle Swarm Optimization (PSO), both of which have found broad applicability across different domains [8]. Sun et al. proposed the automatic design of CNN architectures for image classification using GA. Their study reported that the GA-based approach outperformed both manual and other automated design methods in terms of classification accuracy, while also requiring lower computational cost [9]. Polap developed an adaptive technique that combines GA with convolutional classifiers for the analysis of microscopy images. The results demonstrated that the GA-based approach achieved 7.5% higher accuracy compared to traditional methods [10].

Particle Swarm Optimization (PSO) has been proposed for clustering problems and tested on two synthetic and five real-world datasets. The findings indicate that the algorithm successfully solves clustering tasks, demonstrating its effectiveness in this domain [11]. Huang developed a novel PSO-based hybrid cluster validity method to address complex clustering and classification problems. In his study, the performance of the proposed method was compared with semi-supervised classification, decision tree, and the Huang index method. The results demonstrated that the proposed approach outperformed existing methods in both clustering and classification tasks [12].

In this study, two methods are presented to improve the speed and efficiency of the classification process: the Genetic Algorithm-based Weighted Naive Bayes (GAW-NB) and the Particle Swarm Optimization-based Weighted Naive Bayes (PSOW-NB). The performance of the developed methods is evaluated by comparing them with the traditional Naive Bayes and Weighted Naive Bayes algorithms. In this context, various datasets from the literature are utilized to analyze accuracy rates and demonstrate the effectiveness of the methods.

The remainder of this paper is organized as follows. In Section 2, the fundamentals of the WNB classifier are introduced. Section 3 describes the optimization process and details the methods used to enhance the performance of the WNB classifier. Section 4 presents the experimental study, including dataset descriptions and implementation details. Section 5 reports and discusses the experimental results obtained through performance evaluations. Finally, Section 6 concludes the paper with a summary of the findings and suggestions for future research.

2. WEIGHTED NAIVE BAYES CLASSIFIER

Bayesian classifiers are among the statistical classification techniques and are preferred by researchers due to their computational efficiency and strong performance [13]. The Weighted Naive Bayes (W-NB) classifier is an enhanced version of the standard Naive Bayes classifier, developed by incorporating weight values into its structure.

The primary purpose of the weighting function is to modify or enhance the influence of each individual feature on the outcomes. A positive weight function defined on a set A, denoted as ω:A→R+ , is expressed as a positive function over A. If the weight function is defined as ω(a) = 1 for all a∈A, then all elements are assigned equal weight and the system can be regarded as unweighted.

$$\sum_{a\in A} f(a) \tag{1}$$

However, given a weight function ω:A → R + the weighted sum is expressed as follows:

$$\sum_{a\in A} f(a)\omega(a) \tag{2}$$

The weighting process within the structure of an artificial neural network is illustrated in Figure 1.

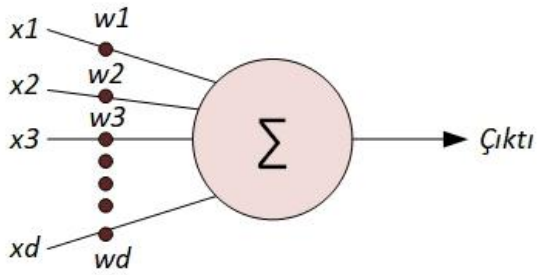


Figure 2. 1 The weighting process for the summation function

This weighting approach is also applicable to the Naive Bayes classifier. Nevertheless, since the primary operation in the Bayes classifier is multiplication, weights are integrated by adding them to the input values rather than multiplying. The principle behind incorporating these weights into the input values is detailed in Equation 3 and illustrated in Figure 2 [6].

$$\prod_{a\in A} (f(a) + \omega(a)) \tag{3}$$

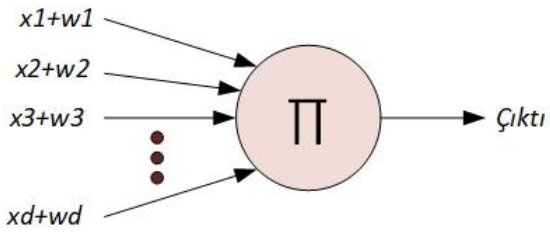


Figure 2. 2 The weighting process for the multiplication function

3. OPTIMIZATION PROCESS

Optimization is the process of determining the values that decision variables should take in order to maximize the value of an objective function defined under certain constraints [14]. In the literature, various optimization algorithms have been applied to different types of optimization problems. Optimization techniques are generally categorized into three main groups: analytical methods, heuristic methods, and metaheuristic methods.
Genetic Algorithm and Particle Swarm Optimization are among the most well-known and widely used algorithms within the family of metaheuristic methods. In this study, Genetic Algorithm and Particle Swarm Optimization were employed to optimize the weights of the Weighted Naive Bayes (W-NB) classifier.

3.1. Genetic Algorithm

Genetic Algorithm (GA) is one of the earliest and most widely used optimization algorithms in the field of evolutionary computation. Inspired by natural evolutionary processes, this algorithm was developed to solve search and optimization problems. Introduced by John Henry Holland in 1975, GA is based on the principle of natural selection. In this algorithm, solutions are obtained through a randomized search within a defined solution space.
Genetic Algorithms simulate natural selection and biological reproduction processes to enable the evolution of the fittest individual over successive generations. Operating without reliance on strict mathematical formulations, this method derives optimal solutions from the most successful individuals of previous generations, and is considered a nonlinear and stochastic process [15]. The ability of genetic algorithms to generate effective solutions in a relatively short amount of time is considered one of their major advantages.
The fundamental steps of the Genetic Algorithm can be summarized as follows [16]:

1. All possible solutions are encoded as binary strings.
2. The initial population is generated from randomly selected sets of candidate solutions.
3. The fitness value of each individual is calculated using a predefined fitness function.
4. Reproduction is performed based on the fitness values of the individuals.
5. The fitness values of the newly generated individuals are evaluated; crossover and mutation operations are applied to produce a new population.
6. All processes are repeated until a predefined number of generations is reached or a stopping criterion is satisfied.
7. Once the number of iterations reaches the stopping criterion, the best individuals are selected according to the fitness function, and the process is terminated.

3.2. Particle Swarm Optimization Algorithm

Particle Swarm Optimization (PSO) is a widely recognized optimization technique categorized under swarm intelligence algorithms, which draws inspiration from the collective behavior observed in natural phenomena, such as bird flocking or fish schooling. Originally proposed by Kennedy and Eberhart in 1995 [17], PSO operates similarly to Genetic Algorithms by starting with a randomly initialized set of candidate solutions. Unlike Genetic Algorithms, however, PSO does not use crossover or mutation mechanisms. Instead, it guides the search for optimal

solutions by continuously updating particle positions based on each particle's historical best (personal best) and the overall best position (global best) within the population.

The steps of the Particle Swarm Optimization (PSO) algorithm are outlined as follows [18]:

1. The initial positions and velocities of the particles are randomly generated in a d-dimensional search space.
2. The optimization fitness function, defined over d variables, is evaluated for each particle.
3. The fitness value of each particle is compared with its personal best value. If the current fitness is better than the personal best, the personal best is updated and the current position is recorded as the new personal best.
4. The fitness value is also compared with the global best value in the population. If the current fitness is better than the global best, the global best is updated accordingly, along with the index and value of the corresponding particle.
5. The velocities and positions of particles are updated according to the standard PSO update equations.These updates allow the particles to move through the search space toward better solutions. Velocity and position updates are performed using Equation (4) and Equation (5), respectively.

$$V_{id} = V_{id} + c_1 * rand( ) * (P_{id} - X_{id}) + c_2 * rand( ) * (G_{id} - X_{id}) \quad (4)$$

$$X_{ld} = X_{id} + V_{id} \quad (5)$$

6. Iterations continue from Step 2 until convergence criteria are met or the maximum number of generations is completed.

4. EXPERIMANTAL STUDY

In this study, the GA and PSO methods were applied to optimize the weights of the Weighted Naive Bayes (W-NB) classifier. Two variants, namely the Genetic Algorithm-based Weighted Naive Bayes Classifier (GAW-NB) and the Particle Swarm Optimization-based Weighted Naive Bayes Classifier (PSOW-NB), were tested on five distinct datasets. The performance of the proposed methods was evaluated by comparing them with each other. All datasets were obtained from the UCI Machine Learning and Intelligent Systems Repository [19]. The characteristics of the datasets used in the experimental study are presented below.

4.1. Tic-Tac-Toe Dataset

Tic-tac-toe is one of the oldest and most popular logic-based games. The game is played on a 3x3 board using the symbols "x" and "o". The Tic-tac-toe dataset encodes all possible board configurations under the assumption that the "x" player always makes the first move. The dataset contains 9 features and a total of 960 instances. The attributes of the Tic-tac-toe dataset are presented in Table 1. In the dataset, "x" represents the x-player, "o" represents the o-player, and "b" denotes a blank space [20].

Tablo 1. Attribute information of the Tic-Tac-Toe dataset

Attributes	Values

Top-left	o, b, x
Top-middle	o, b, x
Top-right	o, b, x
Middle-left	o, b, x
Middle-middle	o, b, x
Middle-right	o, b, x
Bottom-left	o, b, x
Bottom-middle	o, b, x
Bottom-right	o, b, x
Class	Positive, Negative

4.2. Post-Operative Patient Dataset

This dataset is designed to classify patients' postoperative conditions and assist in determining the appropriate hospital unit for referral following surgery. The postoperative period, which begins after a major surgical intervention, is often marked by complications such as hypothermia and is considered a critical phase in patient recovery. The dataset comprises 8 features and includes 90 instances. The detailed attributes of the dataset are provided in Table 2 [21].

Tablo 2. Feature details of the dataset for postoperative patient classification

Attributes	Values
Internal Temperature (°C)	High (>37), Medium (36–37), Low (<36)
Surface Temperature (°C)	High (>36.5), Medium (35–36.5), Low (<35)
Oxygen Saturation (%)	Excellent (≥98), Good (90–98), Medium (80–90), Poor (<80)
Last Blood Pressure Measurement	High (>130/90), Medium (90/70–130/90), Low (<90/70)
Surface Temperature Condition	Stable, Moderately Stable, Unstable
Internal Temperature Condition	Stable, Moderately Stable, Unstable
Blood Pressure Stability	Stable, Moderately Stable, Unstable
Perceived Comfort Level	0–20 (numerical value)
Class	I (Patient referred to Intensive Care Unit), S (Patient ready for discharge), A (Patient referred to general hospital floor)

4.3. The Qualitative Bankruptcy Dataset

The Qualitative Bankruptcy dataset is developed to estimate the likelihood of corporate bankruptcy based on a range of qualitative risk factors. The dataset includes 7 features and 250 instances. Detailed attribute information is provided in Table 3. The features are categorized into the following risk dimensions [22]:

- Industrial Risk**:** Evaluates factors such as government regulations, international treaties, market rivalry intensity, supply chain stability, macroeconomic fluctuations, competitiveness in domestic and global markets, and product lifecycle stages.
- Management Risk**:** Assesses managerial competency and stability, the relationship between management and ownership, effectiveness of human resource strategies, business growth strategies, operational outcomes, feasibility of strategic plans, and likelihood of success.
- Financial Flexibility**:** Measures the firm’s capacity to access and utilize direct, indirect, and alternative financial resources.
- Credibility**:** Reflects the firm’s credit history, reliability of financial disclosures, and strength of its relationships with financial institutions.
- Competitiveness**:** Examines market positioning, depth of core competencies, and uniqueness of business strategies adopted.

- Operational Risk**:** Covers aspects such as procurement stability and diversity, production efficiency, operational continuity, market demand, product and sales diversification, pricing strategies, and distribution network efficiency.

Tablo 3. Attribute information of the Qualitative Bankruptcy dataset

Attribute	Values
Industrial Risk	Positive, Average, Negative
Management Risk	Positive, Average, Negative
Financial Flexibility	Positive, Average, Negative
Credibility	Positive, Average, Negative
Competitiveness	Positive, Average, Negative
Operational Risk	Positive, Average, Negative
Decision	Bankrupt, Non-bankrupt

4.4. The Mammographic Mass Dataset

The Mammographic Mass dataset is designed to differentiate between benign and malignant breast lesions. It may serve as a valuable tool for classifying the malignancy of mammographic masses using BI-RADS attributes and patient age information. After removing instances with missing values, the dataset comprises a total of 825 samples, including 400 benign and 425 malignant cases. Detailed information regarding the six attributes included in the dataset is provided in Table 4 [23].

Tablo 4. Attribute information of the Mammographic Mass dataset

Attribute	Values
BI-RADS Assessment	Ordinal values from 1 to 5
Age	Patient's age (integer)
Shape	Mass shape: 1 = Round, 2 = Oval, 3 = Lobular, 4 = Irregular (nominal)
Margin	Mass margin: 1 = Circumscribed, 2 = Microlobulated, 3 = Obscured, 4 = Ill-defined, 5 = Spiculated (nominal)
Density	Mass density: 1 = High, 2 = Iso-dense, 3 = Low, 4 = Fat-containing
Class	Malignant = 1, Benign = 0

4.5. The Breast Cancer Dataset

This dataset was developed to support decision-making regarding the administration of radiation therapy in breast cancer patients, based on specific clinical and diagnostic features. It comprises 680 instances and includes 9 attributes. Detailed information about the features of the breast cancer dataset is provided in Table 5.

Tablo 5. Attribute information of the Breast Cancer dataset

Attribute	Values
Age	10–19, 20–29, 30–39, 40–49, 50–59, 60–69, 70–79, 80–89, 90–99
Menopause	Below 40, 40 and above
Tumor Size	0–4, 5–9, 10–14, 15–19, 20–24, 25–29, 30–34, 35–39, 40–44, 45–49, 50–54, 55–59
Involved Nodes	0–2, 3–5, 6–8, 9–11, 12–14, 15–17, 18–20, 21–23, 24–26, 27–29, 30–32, 33–35, 36–39
Node-Caps	No, Yes

Degree of Malignancy	1,2,3
Breast	Left, Right
Breast Quadrant	Left-upper, Left-lower, Right-upper, Right-lower, Central
Radiation	No, Yes
Class	no-recurrence-events, recurrence-events

4.6. Weighted Naive Bayes Process

In the Weighted Naïve Bayes (W-NB) classifier, the number of weights corresponds to the number of attributes in the dataset. In the conventional Global W-NB approach, these weights are typically determined using the grid search method. In this study, the grid search technique was utilized to identify the optimal weight values, with all weights constrained within the range [0, 1]. However, as the number of attributes increases, the computational complexity of W-NB also grows due to the nested loop structure inherent in grid search. To mitigate this computational burden, a step size of 0.2 was adopted for weight intervals. Nevertheless, this method remains computationally expensive, especially with larger datasets. Therefore, the use of optimization algorithms is proposed as an effective alternative to enhance efficiency in determining weight values.

4.7. Genetic Algorithm Process

In this study, each weight in the Weighted Naïve Bayes (W-NB) classifier was defined as a decision variable within the Genetic Algorithm (GA) framework. These decision variables were constrained to lie within the range [0, 1]. The classification accuracy served as the objective function of the algorithm, with the primary aim being to maximize this accuracy. To initiate the optimization, the initial population was randomly generated. Each individual in the population was encoded using 8-bit binary strings. These binary strings were used to represent real-valued feature weights by scaling their integer equivalents to the [0, 1] range. This encoding scheme allowed binary representations to be effectively used in a continuous optimization space. The fitness value of each individual was calculated based on its classification accuracy, which reflects the proportion of correctly classified samples over the total number of samples, including both positive and negative predictions.

Subsequent generations were formed using the Roulette Wheel selection method, which probabilistically favors individuals with higher fitness values, thereby increasing their likelihood of contributing to the next generation. In the crossover phase, 60% of the population underwent recombination. For this purpose, randomly selected 8-bit individuals were paired, and a single-point crossover was applied by exchanging 4-bit segments between them. This process aimed to produce offspring with improved fitness characteristics inherited from their parents. Following crossover, a mutation operation was performed by flipping a single bit in selected individuals to introduce genetic diversity and prevent premature convergence. The stopping criterion was defined as no improvement in the fitness value over 2000 consecutive iterations. The control parameters used in this GA process were selected based on recommendations by De Jong and are presented in Table 6 [24].

Tablo 6. Genetic Algorithm parameters based on De Jong's recommendations

Control Parameters	Value
Population Size	100
Crossover Rate	0.60

Mutation Rate	0.001

4.9. Particle Swarm Optimization Process

In the PSO-based optimization, the weights of the WNB are treated as particles within the swarm. Each particle represents a complete weight vector, with each dimension corresponding to a feature weight constrained in the range [0, 1]. The objective of employing PSO is to maximize the classification accuracy. Initially, 100 particles were randomly generated for each weight. While the inertia (approach) velocity can be initialized randomly in general, all initial velocities were set to zero in this study. The PSO parameters c1 and c2 were both set to 2, whereas the random coefficients r1 and r2 were drawn from a uniform distribution in the range [0, 1]. Fitness was measured as the ratio of correctly classified instances to the total number of instances. During the optimization process, particles updated their positions by considering both their personal best solutions and the global best solution found by the swarm. The algorithm was iterated for 1000 generations to determine the optimal weight values. This entire optimization process was repeated 50 times, and the average of the classification accuracies obtained across all runs was computed.

5. EXPERIMENTAL RESULTS

In this section, the classification performances of the standard Naive Bayes, Weighted Naive Bayes, and the Weighted Naive Bayes models optimized via Genetic Algorithm (GAW-NB) and Particle Swarm Optimization (PSOW-NB) are compared across five benchmark datasets. All experiments were conducted using MATLAB. The performance metrics were calculated by averaging the classification results obtained from each fold of the 5-fold cross-validation. This validation strategy was employed to enhance the robustness and generalizability of the evaluation. The detailed fold-wise classification results for each dataset are presented in Table 7.

Tablo 7. Comparison Of NB, WNB, GA-WNB and PSO-WNB Results

Database Name	NB	W-NB	GAW-NB	PSOW-NB
Tic-Tac-Toe Dataset				
1. Fold	69.79	77.60	79.16	78.48
2. Fold	75.00	80.20	81.77	81.44
3. Fold	72.39	77.60	77.60	79.11
4. Fold	68.75	75.00	75.52	75.55
5. Fold	67.70	73.95	74.47	75.59
Accuracy	**70.72**	**76.87**	**77.70**	**78.03**
Post-Operative Patient Dataset				
1. Fold	44.44	83.33	83.33	83.33
2. Fold	55.55	55.55	55.55	67.22
3. Fold	66.66	77.77	72.22	78.70
4. Fold	61.11	73.61	72.22	72.22
5. Fold	77.77	77.77	77.77	82.03
Accuracy	**61.11**	**73.61**	**72.21**	**76.70**
Qualitative Bankruptcy Dataset				
1. Fold	98.00	100.00	100.00	100.00
2. Fold	100.00	100.00	100.00	100.00
3. Fold	100.00	100.00	100.00	100.00
4. Fold	100.00	100.00	100.00	100.00
5. Fold	100.00	100.00	100.00	100.00
Accuracy	**99.60**	**100.00**	**100.00**	**100.00**

Mammographic Mass Dataset				
1. Fold	87.34	86.06	86.66	87.27
2. Fold	80.72	86.06	86.06	86.66
3. Fold	84.33	82.42	83.63	83.03
4. Fold	83.13	89.09	89.09	89.69
5. Fold	80.72	88.48	88.48	88.48
Accuracy	**83.25**	**86.42**	**86.78**	**86.69**
Breast Cancer Dataset				
1. Fold	97.05	99.26	99.26	99.26
2. Fold	93.38	98.52	98.52	98.52
3. Fold	94.11	96.32	97.05	97.79
4. Fold	97.05	98.52	98.52	98.52
5. Fold	97.79	98.52	98.52	98.52
Accuracy	**95.88**	**98.23**	**98.37**	**98.52**

As presented in Table 7, the GAW-NB and PSOW-NB classifiers achieved higher classification accuracy than the standard NB and W-NB models across all five datasets. On the Tic-Tac-Toe dataset, the PSOW-NB model achieved the highest accuracy at 78.03%, which is 7.31% higher than NB and 1.16% higher than W-NB. In the Post-Operative Patient dataset, PSOW-NB reached 76.70%, outperforming NB by 15.59%. In the Mammographic Mass dataset, the highest accuracy was obtained by the GAW-NB model with 86.78%, which is 3.53% higher than NB and slightly better than PSOW-NB. For the Breast Cancer dataset, PSOW-NB achieved 98.52% accuracy, improving upon NB by 2.64%. In the Qualitative Bankruptcy dataset, all models except NB reached 100% accuracy.

Despite the differences in dataset sizes, structures, and class distributions, both optimization-based approaches consistently produced higher accuracy values across all experiments. These findings demonstrate that integrating metaheuristic algorithms into the W-NB framework not only improves classification accuracy but also enhances the generalizability of the models and their ability to deliver consistent results across different datasets.

6. CONCLUSSION

The Weighted Naive Bayes (W-NB) classifier is an enhanced version of the standard Naive Bayes algorithm, offering improved classification performance through the incorporation of feature weights. However, the reliance on the grid search method in W-NB leads to significantly increased computational time, particularly when dealing with high-dimensional datasets. In such cases, determining the optimal set of weights may become computationally infeasible.

In this study, two metaheuristic optimization algorithms Genetic Algorithm and Particle Swarm Optimization were employed to optimize the weights of the W-NB classifier. The proposed GAW-NB and PSOW-NB classifiers aimed to achieve higher classification accuracy than the standard Naive Bayes classifier while reducing the computational cost associated with the traditional W-NB approach.

All experiments were carried out on five benchmark datasets using 5-fold cross-validation. The results demonstrated that both GAW-NB and PSOW-NB consistently outperformed the NB and W-NB classifiers in terms of classification accuracy. For example, PSOW-NB achieved the highest overall accuracy of 98.52% on the Breast Cancer dataset, while GAW-NB yielded the best performance of 86.78% on the Mammographic Mass dataset. On the Qualitative Bankruptcy dataset, all models except NB reached 100% accuracy. These findings provide

strong evidence that integrating metaheuristic optimization techniques into the W-NB framework can lead to the development of fast and highly accurate classifiers.

However, some limitations should be acknowledged. The proposed methods were tested on relatively small and clean datasets. The effectiveness and scalability of the approaches for large-scale or noisy datasets remain to be examined. Furthermore, since metaheuristic algorithms rely on stochastic processes, results may vary across runs, and care should be taken to avoid premature convergence or overfitting.

Future studies may explore alternative optimization algorithms or hybrid approaches to further enhance the efficiency, robustness, and scalability of the W-NB classifier, especially in applications involving high-dimensional and complex data environments.

7. ACKNOWLEDGEMENTS

This article was written as a part of master's thesis titled "Optimization of the weights of Weighted Naive Bayesian Classifier" at Firat University. Thesis no: 527473.

REFERENCES

[1] Aggarwal CC, Yu PS. Data Mining Techniques for Associations, Clustering and Classification. In: Zhong N, Zhou L, editors. Methodol. Knowl. Discov. Data Min., Berlin, Heidelberg: Springer; 1999, p. 13–23. https://doi.org/10.1007/3-540-48912-6_4.

[2] Ravinder B, Seeni SK, Prabhu VS, Asha P, Maniraj SP, Srinivasan C. Web Data Mining with Organized Contents Using Naive Bayes Algorithm. 2024 2nd Int. Conf. Comput. Commun. Control IC4, 2024, p. 1–6. https://doi.org/10.1109/IC457434.2024.10486403.

[3] Jain J, Upadhyay SK, Nayak SK. Analyzing the Effectiveness of Machine Learning Algorithms in detecting Fake News. Comput. Commun. Intell., CRC Press; 2025.

[4] Arrayyan AZ, Setiawan H, Putra KT. Naive Bayes for Diabetes Prediction: Developing a Classification Model for Risk Identification in Specific Populations. Semesta Tek 2024;27:28–36. https://doi.org/10.18196/st.v27i1.21008.

[5] Karabatak M. A new classifier for breast cancer detection based on Naïve Bayesian. Measurement 2015;72:32–6. https://doi.org/10.1016/j.measurement.2015.04.028.

[6] Aksoy G, Karabatak M. Performance Comparison of New Fast Weighted Naïve Bayes Classifier with Other Bayes Classifiers. 2019 7th Int. Symp. Digit. Forensics Secur. ISDFS, 2019, p. 1–5. https://doi.org/10.1109/ISDFS.2019.8757558.

[7] Lin J, Yu J. Weighted Naive Bayes classification algorithm based on particle swarm optimization. 2011 IEEE 3rd Int. Conf. Commun. Softw. Netw., 2011, p. 444–7. https://doi.org/10.1109/ICCSN.2011.6014307.

[8] Tian Z, Fong S, Tian Z, Fong S. Survey of Meta-Heuristic Algorithms for Deep Learning Training. Optim. Algorithms - Methods Appl., IntechOpen; 2016. https://doi.org/10.5772/63785.

[9] Sun Y, Xue B, Zhang M, Yen GG, Lv J. Automatically Designing CNN Architectures Using the Genetic Algorithm for Image Classification. IEEE Trans Cybern 2020;50:3840–54. https://doi.org/10.1109/TCYB.2020.2983860.

[10] Połap D. An adaptive genetic algorithm as a supporting mechanism for microscopy image analysis in a cascade of convolution neural networks. Appl Soft Comput 2020;97:106824. https://doi.org/10.1016/j.asoc.2020.106824.

[11] Cura T. A particle swarm optimization approach to clustering. Expert Syst Appl 2012;39:1582–8. https://doi.org/10.1016/j.eswa.2011.07.123.

[12] Huang KY. A hybrid particle swarm optimization approach for clustering and classification of datasets. Knowl-Based Syst 2011;24:420–6. https://doi.org/10.1016/j.knosys.2010.12.003.

[13] Kotsiantis SB. Supervised Machine Learning: A Review of Classification Techniques 2007.

[14] Dagdia, Zaineb Chelly, Mirchev, Miroslav. Optimization Problem - an overview 2020.

[15] Alhijawi B, Awajan A. Genetic algorithms: theory, genetic operators, solutions, and applications. Evol Intell 2024;17:1245–56. https://doi.org/10.1007/s12065-023-00822-6.

[16] Gen, Mitsuo, Cheng, Runwei. Foundations of Genetic Algorithms. Genet. Algorithms Eng. Optim., John Wiley & Sons, Ltd; 1999, p. 1–52. https://doi.org/10.1002/9780470172261.ch1.

[17] Jain M, Saihjpal V, Singh N, Singh SB. An Overview of Variants and Advancements of PSO Algorithm. Appl Sci 2022;12:8392. https://doi.org/10.3390/app12178392.

[18] Eberhart, Shi Y. Particle swarm optimization: developments, applications and resources. Proc. 2001 Congr. Evol. Comput. IEEE Cat No01TH8546, vol. 1, 2001, p. 81–6 vol. 1. https://doi.org/10.1109/CEC.2001.934374.

[19] Home - UCI Machine Learning Repository n.d. https://archive.ics.uci.edu/ (accessed January 30, 2025).

[20] Aha D. Tic-Tac-Toe Endgame 1991. https://doi.org/10.24432/C5688J.

[21] Sharon Summers LW. Post-Operative Patient 1991. https://doi.org/10.24432/C5DG6Q.

[22] Kim M-J, Han I. The discovery of experts' decision rules from qualitative bankruptcy data using genetic algorithms. Expert Syst Appl 2003;25:637–46. https://doi.org/10.1016/S0957-4174(03)00102-7.

[23] Elter M. Mammographic Mass 2007. https://doi.org/10.24432/C53K6Z.

[24] De Jong K. Learning with genetic algorithms: An overview. Mach Learn 1988;3:121–38. https://doi.org/10.1007/BF00113894.