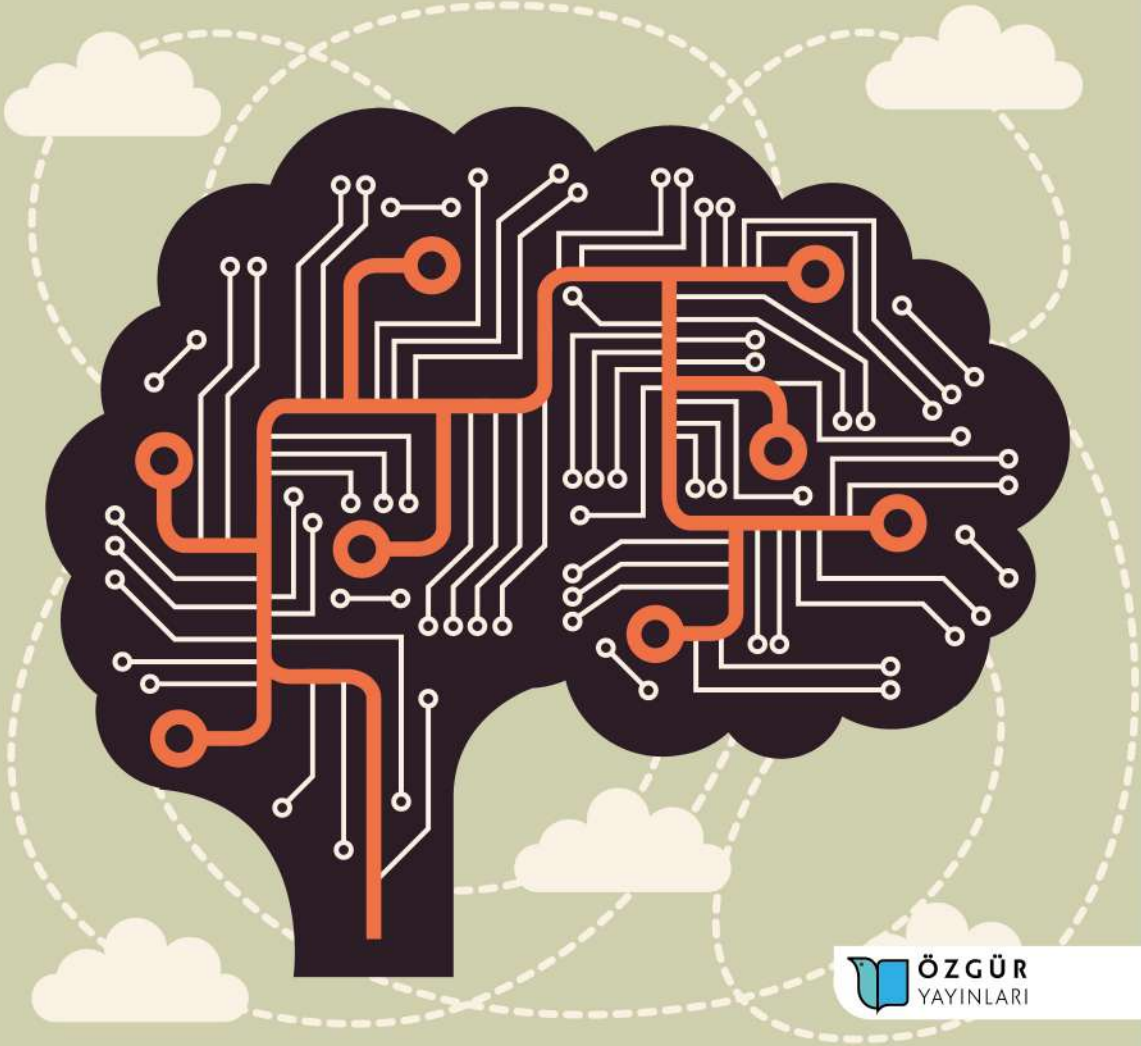


Nicel Karar Vermede Çok Kriterli Yaklaşımlar ve Makine Öğrenmesi Çalışmaları

Editör: Dr. Öğr. Üyesi Hasan Aykut KARABOĞA



ÖZGÜR
YAYINLARI

Nicel Karar Vermede Çok Kriterli Yaklaşımlar ve Makine Öğrenmesi Çalışmaları

Editör

Dr. Öğr. Üyesi Hasan Aykut KARABOĞA



Published by

Özgür Yayın-Dağıtım Co. Ltd.

Certificate Number: 45503

📍 15 Temmuz Mah. 148136. Sk. No: 9 Şhitkamil/Gaziantep

☎ +90.850 260 09 97

📞 +90.532 289 82 15

🌐 www.ozguryayinlari.com

✉ info@ozguryayinlari.com

Nicel Karar Vermede Çok Kriterli Yaklaşımlar ve Makine Öğrenmesi Çalışmaları

Editor: Dr. Öğr. Üyesi Hasan Aykut KARABOĞA

Language: Turkish

Publication Date: 2025

Cover design by Mehmet Çakır

Cover design and image licensed under CC BY-NC 4.0

Print and digital versions typeset by Çizgi Medya Co. Ltd.

ISBN (PDF): 978-625-5757-48-7

DOI: <https://doi.org/10.58830/ozgur.pub900>



This work is licensed under the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC 4.0). To view a copy of this license, visit <https://creativecommons.org/licenses/by-nc/4.0/>. This license allows for copying any part of the work for personal use, not commercial use, providing author attribution is clearly stated.

Suggested citation:

Karaboğa, H. A. (ed) (2025). *Nicel Karar Vermede Çok Kriterli Yaklaşımlar ve Makine Öğrenmesi Çalışmaları*.

Özgür Publications. DOI: <https://doi.org/10.58830/ozgur.pub900>. License: CC-BY-NC 4.0

The full text of this book has been peer-reviewed to ensure high academic standards. For full review policies, see <https://www.ozguryayinlari.com/>



Ön Söz

Günümüz dünyasında araştırmacılar ve sektör çalışanları benzeri görülmemiş bir veri bolluğu ve karmaşıklıkla karşı karşıyadır. Bireylerden kurumlara, endüstriden sosyal bilimlere kadar her alanda karar vericiler, çok sayıda alternatif ve birbiriyle çelişen kriterler arasından en doğru seçimi yapma zorluğuyla yüzleşmektedir. Bu dinamik ortamda, sezgisel yaklaşımların yerini alan kanıta dayalı, analitik ve nicel yöntemler, yalnızca bir rekabet avantajı değil, aynı zamanda bir zorunluluk haline gelmiştir.

Elinizdeki “Nicel Karar Vermede Çok Kriterli Yaklaşımlar ve Makine Öğrenmesi Çalışmaları” isimli eser bu ihtiyaca yanıt vermek üzere, alanında uzman akademisyenlerin ve araştırmacıların değerli katkılarıyla hazırlanmıştır. Kitap, karar verme biliminin iki güçlü ve modern sacayağını bir araya getirerek okuyucularına hem teorik bir temel hem de zengin bir uygulama yelpazesi sunmayı amaçlamaktadır.

Kitap, iki temel eksen etrafında şekillenmektedir. İlk ekseninde, en uygun alternatifi belirleme problemine odaklanan Çok Kriterli Karar Verme (ÇKKV) yöntemleri yer almaktadır. Bu bölümde, tedarikçi seçiminden elektrikli araçların sıralanmasına kadar uzanan güncel ve pratik problemler yenilikçi ve etkinliği kanıtlanmış ÇKKV metotlarıyla çözülmektedir. Bu çalışmalar, okuyuculara karmaşık karar problemlerini nasıl sistematik bir çerçeveye oturtabileceklerini göstermektedir. İkinci eksen ise, son yılların en dönüştürücü teknolojilerinden biri olan Makine Öğrenmesi (ML) uygulamalarını kapsamakta ve verinin içindeki gizli örüntüleri ortaya çıkararak geleceğe yönelik isabetli tahminler yapmayı hedefleyen algoritmaları ele almaktadır. Tedarik zinciri yönetiminde talep tahmininden, tüketici davranışlarının analizine; trafik kazalarında hayatta kalma olasılıklarının belirlenmesinden, üniversite öğrencilerindeki depresyon riskinin öngörülmesine kadar birbirinden farklı ve güncel konular güçlü makine öğrenmesi modelleriyle incelenmektedir.

Bu eser, akademisyenler, lisansüstü öğrenciler, veri bilimciler, analistler ve mühendisler için hem bir başvuru kaynağı hem de yeni araştırmalar için bir ilham kaynağı olmayı hedeflemektedir. Sunulan her bir bölüm, teorik altyapıyı somut ve gerçek dünya verileriyle birleştirerek, okuyucunun yöntemleri daha derinlemesine anlamasına ve kendi problemlerine uyarlamasına olanak tanımaktadır.

Sonu olarak, “Nicel Karar Vermede ok Kriterli Yaklařımlar ve Makine ğrenmesi alıřmaları”, okuyucuları modern dnyanın karmařık karar problemlerini özmek iin gerekli olan analitik aralarla donatmayı amalayan kapsamlı bir alıřmadır.

Bu kitabın hazırlanmasında titiz alıřmaları ile emek veren tüm yazarlara ve bu alıřmanın yayınlanmasında emeėi geenlere teřekkrlerimi sunarım.

Editr

Dr. ğr. yesi Hasan Aykut KARABOėA

İçindekiler

Ön Söz

iii

Bölüm 1

Yapay Zekâ Sohbet Robotlarının Çok Kriterli Karar Verme Süreçlerinde Kullanımı: Türkiye’deki Elektrikli SUV Modeller Üzerine Bir Uygulama 1

Serkan Akoğlu

Bölüm 2

ENTEGRE LOPCOW-APLOCO Yöntemleriyle Elektrikli Otomobil Seçimi Problemi 19

Mehmet Akif Kara

Bölüm 3

Makine Öğrenmesi ile Üniversite Öğrencilerinde Depresyonun Tahmini: Akademik, Psikolojik ve Demografik Faktörlerin Rolü 33

Hasan Aykut Karaboğa

Abdülkadir Keskin

Bölüm 4

Tedarik Zinciri Yönetiminde Makine Öğrenmesi Uygulamaları 53

Ersin Şener

Pakize Yiğit

Emrah Önder

Bölüm 5

Makine Öğrenmesi Yöntemleri ve Derin Sinir Ağlarıyla Trafik Kazalarında Hayatta Kalma Sonuçlarının Tahmin Edilmesi 71

Mustafa Bayram Gücen

Tüketicilerin Özel Günlerde Satın Alma Davranışlarının Rastgele Orman
Yöntemiyle İncelenmesi

103

Mert Ersen

Melik Masarifoğlu

Kemal Cem Söylemez

Yapay Zekâ Sohbet Robotlarının Çok Kriterli Karar Verme Süreçlerinde Kullanımı: Türkiye’deki Elektrikli SUV Modeller Üzerine Bir Uygulama

Serkan Akoğul¹

Özet

Bu çalışma, yapay zekâ tabanlı sohbet robotlarının (ChatGPT, Gemini, Copilot) çok kriterli karar verme (ÇKKV) süreçlerinde nasıl kullanılabileceğini incelemektedir. Günümüzde karar verme süreçleri giderek daha karmaşık hale gelirken, yapay zekâ destekli sistemler bu süreçleri hızlandırmak ve objektif bir bakış açısı sunmak açısından önemli bir potansiyele sahiptir. Bu bağlamda, çalışmada Türkiye otomotiv pazarında 2024 yılının ilk on ayında en çok satılan elektrikli SUV modelleri olan TOGG-T10X, Tesla Model Y, KG Mobility Torres, BMW XI, Mercedes EQB, Mini Countryman ve MG ele alınmıştır.

Söz konusu modeller, menzil, batarya kapasitesi, motor gücü, şarj süresi, hızlanma süresi, enerji tüketimi ve fiyat gibi çeşitli performans ve maliyet kriterleri çerçevesinde değerlendirilmiştir. Çalışmanın temel amacı, belirlenen kriterlerin ağırlıklarının tespitinde hem CRITIC gibi objektif bir yöntem hem de LBWA gibi sübjektif bir yöntemin yanı sıra, yapay zekâ sohbet robotları tarafından önerilen ağırlıkların karşılaştırmalı analizini yapmaktır.

Alternatiflerin sıralanmasında TOPSIS yöntemi kullanılarak, geleneksel yöntemlerle yapay zekâ destekli kriter ağırlıklandırma yaklaşımlarının sonuçları karşılaştırılmıştır. Bu sayede, yapay zekâ sohbet robotlarının ÇKKV süreçlerine katkısının ne ölçüde güvenilir ve tutarlı olduğu değerlendirilecektir. Elde edilen bulgular, yapay zekâ tabanlı sistemlerin ÇKKV süreçlerine entegrasyonu konusunda önemli çıkarımlar sunarak, bu tür sistemlerin karar destek mekanizmalarındaki rolü hakkında yeni bir perspektif sağlamaktadır. Ayrıca, geleneksel yöntemler ile yapay zekâ tabanlı yöntemlerin sunduğu avantajlar, dezavantajlar ve farklılıklar ortaya konmuştur.

1 Doç. Dr., Pamukkale Üniversitesi Fen Fakültesi İstatistik Bölümü, sakogul@pau.edu.tr, ORCID: 0000-0002-0346-4308

1. Giriş

Teknolojinin hızla gelişmesiyle birlikte yapay zekâ (YZ) tabanlı sistemler, karar destek mekanizmalarında önemli bir rol oynamaya başlamıştır. Özellikle üretken yapay zekâ sohbet robotları (ChatGPT, Gemini, Copilot), geniş veri kümelerini analiz ederek karar verme süreçlerinde yardımcı olabilecek araçlar haline gelmiştir. Çok Kriterli Karar Verme (ÇKKV) yöntemleri ise, karar verme süreçlerinde birden fazla kriterin dikkate alınmasını sağlayan analitik yöntemlerdir. Bu yöntemler, çeşitli alanlarda alternatifler arasından en uygun seçeneği belirlemek için kullanılmaktadır (Hwang ve Yoon, 1981). Son yıllarda, YZ tabanlı sohbet robotlarının ÇKKV süreçlerine nasıl entegre edilebileceği üzerine çeşitli araştırmalar yapılmıştır (Basmacı, 2022; Göktaş ve Güçlü, 2024). Bu bağlamda, sohbet robotlarının karar destek sistemlerinde bir araç olarak kullanılması hem akademik araştırmalar hem de uygulamalı çalışmalar açısından dikkat çekici bir konu haline gelmiştir.

Bu çalışma, elektrikli SUV modelleri üzerindeki bir uygulama ile Türkiye’de ÇKKV süreçlerinde yapay zekânın etkisini incelemektedir. Elektrikli araçlar, sürdürülebilir ulaşım çözümleri arasında giderek daha fazla önem kazanmaktadır ve tüketicilerin araç seçimi yaparken dikkate aldığı kriterler oldukça çeşitlidir. Araştırmada, 2024 yılı ilk on ayı itibarıyla Türkiye’de en çok satan elektrikli SUV modelleri olan TOGG-T10X, Tesla Model Y, KG Mobility Torres, BMW X1, Mercedes EQB, Mini Countryman ve MG markalarına ait araçlar değerlendirmeye alınmıştır. Bu modeller, menzil, batarya kapasitesi, motor gücü, şarj süresi, hızlanma, enerji tüketimi ve fiyat gibi kriterler çerçevesinde analiz edilmiştir.

Çalışmada, kriterlerin ağırlıklandırılmasında hem objektif hem de subjektif yöntemler kullanılmıştır. Objektif kriter ağırlıklandırma yöntemi olarak CRITIC (CRiteria Importance Through Intercriteria Correlation) yöntemi benimsenmiştir (Diakoulaki vd., 1995). Bu yöntem, kriterler arasındaki bağımlılıkları değerlendirerek nesnel ağırlıklar hesaplamaktadır. Subjektif kriter ağırlıklandırma yöntemi olarak ise LBWA (Level Based Weight Assessment) yönteminin (Žižović ve Pamucar, 2019) yanı sıra, üretken yapay zekâ sohbet robotları tarafından önerilen ağırlıklarda kullanılmıştır. Alternatiflerin yani elektrikli SUV modellerinin sıralanması için ise TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) yöntemi uygulanmıştır (Hwang ve Yoon, 1981).

Sonuç olarak, bu çalışmada yapay zekâ tabanlı sistemlerin ÇKKV süreçlerine entegrasyonu ve geleneksel yöntemlerle karşılaştırıldığında ortaya çıkan avantajları ve farklar değerlendirilmektedir. Elde edilen bulgular, yapay zekâ sohbet robotlarının karar verme süreçlerine katkı sağlayabileceğini ve

alternatif analizlerde etkili bir araç olarak kullanılabileceđini göstermektedir. YZ destekli yöntemlerin karar verme süreçlerine entegrasyonu, gelecekte çeşitli sektörlerde daha fazla uygulanabilirlik kazanarak karar destek sistemlerinde önemli bir bileşen haline gelebilir.

2. Literatür İncelemesi

Son yıllarda, yapay zekâ teknolojileri ve Çok Kriterli Karar Verme (ÇKKV) yöntemleri, farklı disiplinlerde karar destek sistemlerinin etkinliğini artırmak amacıyla yaygın olarak kullanılmaktadır. Özellikle ChatGPT, Gemini ve Copilot gibi yapay zekâ sohbet robotları, kullanıcıların karar alma süreçlerini yönlendirme potansiyeline sahip olup, bilgiye hızlı erişim ve analiz kapasitesi sunmaktadır. Öte yandan, elektrikli araçlar, sürdürülebilir ulaşım çözümleri kapsamında giderek artan bir ilgi görmektedir ve bu alanda karar verme süreçleri; menzil, batarya kapasitesi, şarj süresi gibi çok sayıda kriteri içermektedir. Bu bağlamda, literatürde ÇKKV yöntemleri ile elektrikli araç seçimleri üzerine yapılan çalışmaların yanı sıra, yapay zekâ sohbet robotlarının karar destek sistemlerindeki rolü de incelenmiştir. Bu bölümde, belirtilen konularda yapılan çalışmalar kronolojik sırayla ele alınmaktadır.

Gavcar ve Kara (2020) çalışmalarında, Türkiye’de satışa sunulan farklı model elektrikli otomobilleri batarya kapasitesi, beygir gücü, aerodinamik katsayısı, menzil ve satış fiyatı gibi kriterler baz alınarak karşılaştırmış ve sıralamıştır. Kriterlerin önem düzeyleri Entropi yöntemiyle belirlenmiş, otomobillerin performansları ise TOPSIS yöntemiyle analiz edilmiştir. Çalışma sonucunda, beygir gücü ve satış fiyatı kriterlerinin ön plana çıktığı, Tesla Model X LR ve Model 3 modellerinin en yüksek performansı elde ettiđi belirlenmiştir.

Karasu Asnaz ve Özdemir (2021) çalışmasında, elektrikli araçlar (EV) için şarj istasyonu altyapısının planlanması ve Balıkesir Üniversitesi Çağış Kampüsü’nde bir uygulama modeli geliştirilmesi amaçlanmıştır. Çalışmada, kampüs içinde optimum sayıda şarj ünitesi tespiti için kuyruk teorisi kullanılarak, şarj ünitelerinin yerleştirileceđi en uygun noktalar otoparklar olarak belirlenmiş ve bu dağıtım ÇKKV metotlarıyla analiz edilmiştir. Alternatiflerin öncelik sıralaması, Bulanık Analitik Hiyerarşi Prosesi (FAHP) ile gerçekleştirilmiş ve FAHP, PROMETHEE ve SMART yöntemleriyle karşılaştırılmıştır.

Basmacı (2022) çalışmasında, sağlık sektöründe yapay zekâ (YZ) uygulamalarının adaptasyon sürecinde karşılaşılan zorlukları belirlemek ve bu zorlukları önceliklendirmek amacıyla Analitik Hiyerarşi Süreci (AHS) yöntemini kullanmıştır. Çalışmada, üçü tıp doktoru ve ikisi YZ uygulamaları

üzerine çalışan akademisyen olmak üzere beş uzmanın görüşlerine başvurulmuştur. Elde edilen bulgular, “Etik Problemler” başlığının en önemli zorluk olduğunu göstermiştir.

Büyükselçuk ve Tozan (2022) çalışmasında, Türkiye’de satışa sunulan e-SUV araçlarının performanslarını değerlendirmek için CRITIC ve EATWIOS yöntemleri kullanılmıştır. Sonuç olarak, enerji tüketimi, fiyat ve araç ağırlığının önemli faktörler olduğu belirlenmiş ve İngiltere menşeli e-SUV’un en yüksek performansa sahip olduğu ancak verimliliğin artırılması gerektiği vurgulanmıştır.

Pinochet vd. (2023) çalışmalarında, yükseköğretimde ChatGPT’nin kullanımına yönelik belirlenen değerlendirme kriterleri doğrultusunda, ChatGPT ile gerçekleştirilebilecek iş birliğine dayalı çalışma alternatiflerini analiz etmek için PROMETHEE ve SAPEVO-M1 yöntemlerini entegre etmişlerdir. Analiz sonuçlarına göre, “Özerk Öğrenmeye Destek” alternatifi en yüksek pozitif akış ve en düşük negatif akış değerlerine sahip olarak en iyi performansı sergilemiştir.

Ahangar ve Fietko (2023) çalışmasında, ChatGPT’nin finansal karar alma süreçlerinde nasıl kullanılabileceği incelenmiştir. Yapay zekânın, finansal analiz, yatırım kararları ve risk yönetimi gibi alanlarda veri analizi ve stratejik önerilerde bulunma potansiyeline sahip olduğu belirtilmiş, ancak doğrudan yatırım tavsiyesinde bulunmadığı vurgulanmıştır. Sonuç olarak, ChatGPT’nin finans uzmanları için yararlı bir kaynak olabileceği ancak sınırlı finansal uzmanlığa sahip bireyler için yeterli olmayabileceği ifade edilmiştir.

Çoşkun (2023) çalışmasında, elektrikli araç seçiminde etkili olan kriterler ele alınarak, çok kriterli karar verme teknikleriyle bu kriterlerin değerlendirilmesi ve sıralanması gerçekleştirilmiştir. Çalışmada, objektif ağırlıklandırma teknikleri olarak ENTROPI ve CRITIC, subjektif ağırlıklandırma teknikleri olarak ise AHP ve WINGS yöntemleri kullanılmıştır. Çalışma sonucunda, fiyat kriteri en yüksek öneme sahip kriter olarak belirlenmiş, diğer kriterlerin sıralamalarında ise kullanılan yöntemler arasında farklılıklar gözlemlenmiştir.

Göktaş ve Güçlü (2024) çalışmalarında, ÇKKV problemlerinde kriter ağırlıklarının belirsizliği durumunda uzman görüşü yerine yapay zekâ sohbet robotları ve geçmiş verilerden yararlanarak bir yaklaşım geliştirilmiştir. Borsa İstanbul’da işlem gören sekiz hisse senedinin portföy seçimi sürecinde ÇKKV yöntemleri kullanılarak karar verme süreçleri analiz edilmiştir. Sonuçlar, uzman bilgisi ile geçmiş verilerden elde edilen sonuçlar arasında genellikle pozitif ancak yüksek düzeyde olmayan bir ilişki olduğunu göstermiştir.

Güler (2024) alışmasında, Türkiye’de satışa sunulan elektrikli araçlar için önemli kriterler, BWM ve LOPCOW yöntemleri ile ağırlıklandırılmış ve sıralama yapılmıştır. Çalışma sonucunda, en önemli kriterin “Batarya Kapasitesi” olduğu, en az önemli kriterin ise “Ağırlık” olduğu tespit edilmiştir.

Kanmaz ve Ertuğrul (2024) çalışmalarında, Türkiye’de 2023 yılında en çok satan 10 farklı elektrikli aracı belirli kriterlere göre sıralamışlardır. Çalışmada, fiyat, beygir gücü, menzil, DC hızlı şarj süresi, batarya kapasitesi ve yakıt tüketimi gibi kriterler göz önünde bulundurulmuştur. Kriterlerin önem düzeyleri bulanık PIPRECIA yöntemiyle belirlenmiş ve CRADIS yöntemiyle analiz edilmiştir. Çalışma sonucunda, yakıt tüketimi ve fiyatın yüksek öneme sahip kriterler olduğu belirlenmiş ve Togg T10X V2 modeli en yüksek skoru elde etmiştir.

Arıkan Kargı (2024) çalışmasında, Türkiye’de satışta olan elektrikli SUV’lar arasından tüketiciler için en uygun e-SUV alternatifini belirlemek amacıyla CRITIC ve COPRAS yöntemleri kullanılmıştır. Analiz sonuçlarına göre, en önemli üç kriterin hızlı şarj süresi, enerji tüketimi ve fiyat olduğu, en yüksek sıralamaya sahip araçların ise Subaru Solterra, New MG ZS EV ve BMW iX3 eDrive20 olduğu belirlenmiştir.

Bu çalışmalar, yapay zekâ ve ÇKKV yöntemlerinin karar verme süreçlerinde nasıl entegre edilebileceğini ve özellikle elektrikli araç seçiminde nasıl kullanılabileceğini ortaya koymaktadır.

3. Yöntemler

Bu çalışmada, elektrikli SUV modellerinin ÇKKV sürecinde değerlendirilmesi için hem kriter ağırlıklandırma hem de alternatif sıralama yöntemleri kullanılmıştır. Kriter ağırlıklarını belirlemek amacıyla objektif bir yöntem olan CRITIC ve sübjektif bir yöntem olan LBWA’nın yanı sıra yapay zekâ sohbet robotları (ChatGPT, Gemini, Copilot) tarafından önerilen ağırlıklar kullanılmıştır. CRITIC yöntemi, kriterler arasındaki bağımlılıkları ve değışkenlikleri dikkate alarak ağırlıkları belirlerken, LBWA yöntemi ise uzman görüşlerine dayalı olarak kriter önem derecelerini tespit etmektedir. Yapay zekâ sohbet robotları tarafından belirlenen ağırlıklar ise çalışmaya uygun olarak verilen istemler (prompt) sonucunda elde edilmiştir. Alternatiflerin sıralanmasında ise TOPSIS yöntemi uygulanmıştır. TOPSIS, ideal ve negatif ideal çözüme olan uzaklıkları temel alarak en iyi alternatifin belirlenmesini sağlar. Bu yöntemler birlikte ele alınarak, yapay zekâ tabanlı sistemlerin ÇKKV süreçlerindeki etkinliği ve karar verme mekanizmalarına etkisi analiz edilmiştir.

3.1. CRITIC

CRITIC (Criteria Importance Through Intercriteria Correlation) yöntemi, ÇKKV problemlerinde kriter ağırlıklarını belirlemek için kullanılan objektif bir yöntemdir. Bu yöntem, kriterler arasındaki bağımlılıkları ve değişkenlikleri dikkate alarak her bir kriterin göreceli önemini hesaplamaktadır. CRITIC yöntemi, veri setindeki kriterlerin varyanslarını ve birbiriyle olan korelasyonlarını analiz ederek, yüksek bilgi içeriğine sahip ve diğer kriterlerden bağımsız olan kriterlere daha yüksek ağırlık atamaktadır (Diakoulaki vd., 1995).

CRITIC yönteminin uygulanması, öncelikle kriterlere ait karar matrisinin standartlaştırılmasıyla başlar. Standartlaştırmada kriterlerin fayda ve maliyet durumlarına göre hesaplamalar yapılır. Daha sonra, her bir kriterin varyansı hesaplanarak kriterlerin ayırma gücü belirlenir. Kriterler arasındaki korelasyon katsayıları kullanılarak, kriterlerin bağımlılık ilişkileri analiz edilir ve kriterlerin birbirinden ne kadar bağımsız olduğu ölçülür. Son aşamada, kriterlerin bilgi içeriği ve bağımsızlık düzeyi göz önünde bulundurularak ağırlık katsayıları hesaplanır. CRITIC yöntemi, karar verme süreçlerinde subjektif önyargıları ortadan kaldırarak, kriterler arasındaki ilişkileri istatistiksel olarak değerlendiren sağlam bir ağırlıklandırma yaklaşımı sunmaktadır (Büyükselçuk ve Tozan, 2022; Coşkun, 2023; Arıkan Kargı, 2024).

3.2. LBWA

LBWA (Level Based Weight Assessment) yöntemi, çok kriterli karar verme (ÇKKV) süreçlerinde kriter ağırlıklarını belirlemek için kullanılan subjektif bir yöntemdir. Bu yöntem, karar vericilerin bilgi ve deneyimlerinden yararlanarak kriterleri belirli önem seviyelerine göre gruplandırmalarına dayanır. LBWA'nın temel avantajı, karar vericilere esneklik sağlaması ve karmaşık matematiksel hesaplamalar gerektirmeden kriterlerin önem derecelerini belirleyebilmesidir. Yöntem, özellikle uzman görüşlerinin ön planda olduğu durumlarda etkin bir şekilde kullanılmakta ve karar vericilerin sezgisel değerlendirmelerini sistematik bir yapıya oturtmaktadır (Žižović ve Pamucar, 2019).

LBWA yönteminin uygulanmasında ilk adım, kriterlerin önem düzeylerine göre hiyerarşik bir sıralamaya tabi tutulmasıdır. Uzmanlar, kriterleri en yüksek öneme sahip olandan en düşük öneme sahip olana kadar seviyelendirerek sınıflandırır. Ardından, bu seviyelendirme doğrultusunda kriterler arasında doğrusal bir ilişki varsayılır ve her seviyedeki kriterlere göreceli ağırlık değerleri atanır. Son olarak, kriter ağırlıkları normalize edilerek nihai ağırlıklar elde

edilir. LBWA yöntemi, karar vericilerin subjektif değerdendirmelerini dikkate alan yapıyla, çok kriterli karar verme süreçlerinde esnek ve pratik bir ağırlıklandırma yöntemi sunmaktadır (Uluskan vd., 2022; Özekenci, 2024; Delice ve Akvıran,2024)

3.3. Yapay zekâ sohbet robotları (ChatGPT, Gemini, Copilot)

Yapay zekâ sohbet robotları (ChatGPT, Gemini, Copilot), kullanıcılardan gelen metin tabanlı girdileri (promptları) analiz ederek yanıt üretir. Prompt (istem), yapay zekâyâ verilen komut veya soru olarak tanımlanabilir ve modelin vereceğı cevabın kapsamını, ayrıntı düzeyini ve içeriğini doğrudan etkiler. Etkili bir şekilde formüle edilen promptlar, modelin daha doğru, tutarlı ve kapsamlı yanıtlar vermesine yardımcı olur. Özellikle ÇKKV süreçlerinde, sohbet robotlarına kriter ağırlıklandırma, alternatiflerin değerdendirilmesi ve sonuçların analizi gibi konularda yönlendirici sorular verilebilir. Bu sayede, modelin karar destek sistemlerine entegrasyonu mümkün hale gelir.

ChatGPT, OpenAI tarafından geliştirilen ve diyalog konusunda uzmanlaşmış bir yapay zekâ sohbet botudur. Bu chatbot, denetimli ve takviyeli öğrenme teknikleriyle ince ayar yapılmış büyük bir dil modeline dayanır. İlk versiyonu 2022’de yayınlanmış olup, GPT-3, GPT-3.5, GPT-4 ve GPT-4o sürümelerini çıkarmıştır. ChatGPT, çeşitli görevlerde yardımcı olmak için eğitilmiştir ve insan benzeri metinler oluşturma, farklı dillerde iletişim kurma, metin özetleme, soru yanıtlama, kod yazma ve hatta yaratıcı içerik oluşturma gibi yeteneklere sahiptir (<https://tr.wikipedia.org/wiki/ChatGPT>).

Gemini, Google AI tarafından geliştirilen, Google’ın LaMDA dil ailesini kullanan ve 2023’te yayınlanan bir sohbet robotudur. Google’ın Bard platformunun evrimleşmiş hali olarak tanıtılan Gemini, metin tabanlı görevlerin yanı sıra görselleri ve kodları anlama gibi çok modlu yeteneklere sahiptir. Google’ın arama motoruyla entegre çalışabilmesi, modelin güncel ve güvenilir bilgiler sağlamasına olanak tanımaktadır (<https://tr.wikipedia.org/wiki/Gemini>).

Copilot, Microsoft ve OpenAI iş birliğiyle geliştirilen, ilk olarak 2021’de GitHub Copilot olarak kod yazma süreçlerini desteklemek amacıyla ortaya çıkan, özellikle üretken yapay zekâ uygulamalarında kullanılan bir sohbet robotudur. Microsoft’un Office uygulamaları ve yazılım geliştirme platformlarıyla entegre çalışan Copilot, kullanıcıların kod yazma, içerik oluşturma, veri analizi, arama deneyimini geliştirme, doğal dil işleme, görsel oluşturma, kaynak belirtme gibi çeşitli özellikler sunarak kullanıcıların

verimliliğini artırmayı ve iş süreçlerini kolaylaştırmak için tasarlanmıştır. (https://tr.wikipedia.org/wiki/Microsoft_Copilot).

3.4. TOPSIS

TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) yöntemi, çok kriterli karar verme (ÇKKV) problemlerinde alternatiflerin sıralanması için yaygın olarak kullanılan bir tekniktir. Bu yöntem, her alternatifin ideal çözüme (en iyi değerlere) olan uzaklığını en aza indirirken, negatif ideal çözüme (en kötü değerlere) olan uzaklığını en üst düzeye çıkarmayı amaçlar. Böylece, en iyi alternatif, ideal çözüme en yakın ve negatif ideal çözüme en uzak olan olarak belirlenir. TOPSIS yöntemi, kriterlerin farklı birimlere sahip olduğu durumlarda ölçekleme yaparak karar verme sürecini objektif bir temele oturtur. Yöntem, özellikle performans değerlendirme, tedarikçi seçimi, yatırım kararları ve teknoloji seçimleri gibi alanlarda sıkça kullanılmaktadır (Hwang ve Yoon, 1981).

TOPSIS yönteminin uygulanmasında ilk adım, karar matrisinin normalize edilmesidir. Daha sonra, kriter ağırlıkları kullanılarak ağırlıklı normalize edilmiş karar matrisi oluşturulur. Üçüncü aşamada, ideal ve negatif ideal çözümler belirlenir. Ardından, her alternatifin ideal ve negatif ideal çözüme olan uzaklıkları hesaplanır. Son aşamada ise, alternatiflerin göreceli yakınlık katsayıları hesaplanarak alternatifler en iyiye en yakın olandan en uzak olana doğru sıralanır. TOPSIS yöntemi, kavramsal olarak basit olması ve karar vericilere net bir sıralama sunması nedeniyle ÇKKV süreçlerinde sıkça tercih edilen etkili bir yöntemdir (Akogul vd., 2020; Akogul ve Gögebakan, 2020; Güleriyüz ve Çokyavaşar, 2021; Abdulvahitoğlu vd., 2022)

4. Veri Seti ve Uygulama

Bu bölümde, çalışmada kullanılacak veri seti ve değerlendirme kriterlerine uygun elektrikli SUV modellerinin belirlenmesine yönelik uygulamalar sunulacaktır.

4.1. Veri Seti

Çalışmada kullanılan veri seti, 2024 yılının ilk on ayında Türkiye’de en çok satan ilk yedi elektrikli SUV modellerine ait teknik ve ekonomik özellikleri içermekte olup modellere ait satış rakamları Tablo 1 de verilmiştir. Veri setinde TOGG T10X, Tesla Model Y, KG Mobility Torres, BMW X1, Mercedes EQB, Mini Countryman ve MG modelleri yer almakta olup, değerlendirme kriterleri menzil (km), batarya kapasitesi (kWh), güç (hp), şarj süresi (dk-DC %80), hızlanma süresi (0-100 km/s), enerji tüketimi (kWh/100 km)

ve fiyat (TL) gibi parametrelerden oluşmaktadır. Veriler, üretici firmaların teknik dokümanları, otomotiv sektörüne yönelik piyasa analiz raporları ve güvenilir çevrimiçi kaynaklardan derlenmiştir (OMDM, 2024). Çalışmada kullanılan karar matrisi, SUV modellerin belirlenen kriterler doğrultusunda sayısal değerler ile temsil edildiği Tablo 2’de sunulmuştur. Çalışmanın amacı doğrultusunda, Menzil, Batarya ve Güç kriterleri fayda; Şarj Süresi, Hızlanma, Tüketim ve Fiyat kriterleri ise maliyet özelliği taşıyacak şekilde analizlere dahil edilmiştir.

Tablo 1. Ocak-Ekim 2024 Yılı Türkiye’de En Çok Satan Elektrikli SUV Modeller

Ocak-Ekim 2024	Satış Rakamları	Marka (Model)
1	20 bin 140	TOGG (T10X -V1 RWD)
2	7 bin 367	TESLA (MODEL Y-Standart Range RWD)
3	3 bin 982	KG Mobility (Torres-EVX)
4	2 bin 906	BMW (X1-eDrive20)
5	2 bin 446	Mercedes (EQB-250+)
6	2 bin 221	MINI (Countryman E)
7	2 bin 10	MG (ZS EV)

Tablo 2. Elektrikli SUV Modellerin Kriterlere Göre Oluşturulan Karar Matrisi

Modeller	Menzil	Batarya	Güç	Şarj Süresi	Hızlanma	Tüketim	Fiyat
TOGG-T10X	523	88,5	218	28	7,4	19,1	1.673.000
TESLA-Model Y	455	50	218	30	6,9	15,7	1.866.000
KG Mobility-Torres	635	73,4	207	37	8,1	18,7	1.880.000
BMW-X1	474	64,8	204	29	5,6	17,2	2.888.000
Mercedes-EQB	490	70,5	190	35	8,9	17,5	2.760.000
MINI-Countryman	462	64,6	204	29	8,6	17,4	1.922.000
MG -ZS EV	440	72,6	156	30	8,6	17,8	1.660.000

4.2. Uygulama

Bu bölümünde, çalışmada kullanılan kriterlerin ağırlıkları CRITIC, LBWA ve yapay zekâ sohbet robotların tarafından belirlenecektir. Elde edilen ağırlıklar karşılaştırılarak, yöntemler arasındaki farklılıklar analiz edilecektir.

Sonraki aşamada, alternatif SUV modelleri TOPSIS yöntemi ile sıralanarak, yapay zekâ tabanlı sistemlerin ÇKKV süreçlerine etkisi incelenecektir. Tablo 3’de kriterlere ait yöntemler tarafından hesaplanan ağırlıklar verilmiştir.

Tablo 3. Kriterlere Ait Yöntemler Tarafından Hesaplanan Ağırlıklar

Yöntemler		Menzil	Batarya	Güç	Şarj süresi	Hızlanma	Tüketim	Fiyat
CRITIC	<i>Değer</i>	0,15	0,13	0,11	0,14	0,12	0,17	0,18
	<i>Sıra</i>	3	5	7	4	6	2	1
LBWA	<i>Değer</i>	0,20	0,12	0,08	0,17	0,07	0,10	0,26
	<i>Sıra</i>	2	4	6	3	7	5	1
CHATGPT	<i>Değer</i>	0,25	0,20	0,15	0,10	0,05	0,10	0,15
	<i>Sıra</i>	1	2	3	5	7	5	3
GEMİNİ	<i>Değer</i>	0,25	0,20	0,15	0,15	0,10	0,05	0,10
	<i>Sıra</i>	1	2	3	3	5	7	5
COPILOT	<i>Değer</i>	0,25	0,15	0,10	0,10	0,10	0,10	0,20
	<i>Sıra</i>	1	3	4	4	4	4	2
YZ_ORT	<i>Değer</i>	0,25	0,18	0,13	0,12	0,08	0,08	0,15
	<i>Sıra</i>	1	2	4	5	6	6	3

YZ_ORT: Yapay zekâ sohbet robotlarının kriterlere verdikleri ağırlıkların ortalaması

Tablo 3’e göre yöntemlerin belirlediği kriter önem sıralaması şu şekildedir:

CRITIC: Fiyat > Tüketim > Menzil > Şarj süresi > Batarya > Hızlanma > Güç

LBWA: Fiyat > Menzil > Şarj süresi > Batarya > Tüketim > Güç > Hızlanma

CHATGPT: Menzil > Batarya > Fiyat = Güç > Şarj süresi = Tüketim > Hızlanma

GEMİNİ: Menzil > Batarya > Şarj süresi = Güç > Fiyat = Hızlanma > Tüketim

COPILOT: Menzil > Fiyat > Batarya > Şarj süresi = Güç = Tüketim = Hızlanma

YZ_ORT: Menzil > Batarya > Fiyat > Güç > Şarj süresi > Tüketim = Hızlanma

CRITIC yöntemi kullanılarak belirlenen ağırlıklar, karar matrisinin normalize edilmesinden sonra kriterlere ait varyansın ve kriterler arası korelasyonun kullanılması ile hesaplanmıştır.

LBWA ağırlıkları ise en önemli kriter Fiyat olmak üzere; birinci seviye: Fiyat, Menzil ve Şarj Süresi, ikinci seviye: Batarya ve Tüketim, üçüncü seviye: Güç ve Hızlanma olacak şekilde yazar tarafından seviyelendirilerek hesaplanmıştır.

Aşağıda verilen prompt kullanılarak, yapay zekâ sohbet robotları için kriterlerin ağırlıklandırılması yapılmıştır. Her bir sohbet robotunun çıktıları incelendiğinde, kriterlere verdikleri ağırlıklar ve nedenleri ile birlikte açıklayıcı raporlar görülebilmektedir.

PROMPT:

“Türkiye’de yaşayan biri olarak SUV model Elektrikli araba alacağım. Elektrikli arabaların benim için bazı önemli kriterlerini belirledim. Senden bu kriterleri toplamaları bir olacak şekilde ağırlıklandırmanı istiyorum. Böylelikle bu kriter ağırlıklarını kullanarak, seçmem için belirlediğim alternatif araçları çok kriterli karar verme yöntemleri kullanarak belirleyeceğim. Kriterleri kendini Türkiye’de yaşayan ve bu konuda uzman birisi olarak düşün ve ağırlıklıdır. Kriterler şunlar: Menzil, Batarya, Güç, Şarj Süresi, Enerji Tüketimi, Hızlanma ve Fiyat.”

CHATGPT:

“Elektrikli SUV Seçim Kriterleri ve Ağırlıkları

Elektrikli bir SUV seçerken Türkiye’deki kullanıcılar için önemli olan kriterlere uzman bakış açısıyla ağırlık vermek, karar sürecini oldukça güçlendirecektir. Ağırlıkları belirlerken Türkiye’deki elektrikli araç altyapısını, kullanıcıların ihtiyaçlarını ve maliyet hassasiyetlerini göz önünde bulundurarak aşağıdaki dağılımı öneriyorum:

- *Menzil (0,25): Türkiye’nin şarj altyapısı gelişme aşamasında olduğundan, menzil uzunluğu kritik bir kriterdir. Kullanıcılar uzun menzilli araçları tercih edebilir, böylece sık sık şarj etmek zorunda kalmazlar.*
- *Batarya Kapasitesi (0,20): Batarya kapasitesi menzil ile doğrudan bağlantılı olduğundan önemli bir diğer kriterdir. Geniş batarya kapasiteli araçlar, menzil kaygısını azaltır ve daha uzun ömürlüdür.*

- *Güç (0,15): Elektrikli SUV'lerde güç, performans ve çekiş gücü için önemlidir. Türkiye'nin çeşitli yol koşullarında güçlü motorlar kullanıcıya güven sağlar.*
- *Şarj Süresi (0,10): Şarj istasyonlarının sayısının sınırlı olması nedeniyle kısa şarj süreleri tercih edilir; fakat menzile kadar kritik değildir.*
- *Hızlanma (0,05): Hızlanma performansı kullanıcılar için cazip olsa da çoğu kullanıcı için çok kritik değildir; dolayısıyla daha düşük bir ağırlık verilebilir.*
- *Enerji Tüketimi (0,10): Enerji verimliliği, bataryanın uzun ömürlü olmasına ve maliyet tasarrufuna katkıda bulunur. Türkiye'de elektrik fiyatlarının da göz önünde bulundurulması, enerji tüketimini önemli bir faktör yapar.*
- *Fiyat (0,15): Elektrikli araçların fiyatları hala geleneksel araçlara göre yüksek olabilmekte, bu yüzden maliyet Türk kullanıcıları için önemli bir kriterdir."*

GEMİNİ:

"Türkiye'de Yaşayan Bir SUV Sever İçin Elektrikli Araç Kriterlerinin Ağırlıklandırılması

Türkiye'de elektrikli SUV arayışınızda size yardımcı olmaktan mutluluk duyarım. Belirttiğiniz kriterleri, Türkiye'deki elektrikli araç pazarı ve sürücü alışkanlıkları göz önünde bulundurularak ağırlıklandıralım.

- *Menzil (0,25): Türkiye'deki şarj altyapısının henüz tam olarak gelişmemiş olması nedeniyle, uzun yolculuklarda bile rahat edebilmek için yeterli menzile önemlidir. Ancak, günlük kullanımda şehir içindeki hareketler için çok uzun menzile ihtiyaç duyulmayabilir.*
- *Batarya Kapasitesi (0,20): Bataryanın ömrü, şarj tutma kapasitesi ve güvenilirliği, aracın uzun vadedeki kullanım maliyetini ve performansını doğrudan etkiler.*
- *Güç (0,15): SUV'ların güçlü olmaları beklenir. Elektrik motorlarının sunduğu yüksek tork, özellikle şehir içindeki trafiği ve arazi kullanımını kolaylaştırır.*
- *Şarj Süresi (0,15): Hızlı şarj altyapısının yaygınlaşmasıyla birlikte, şarj süresi önemli olsa da günlük kullanımda evde veya iş yerinde yavaş şarjın yeterli olabileceği durumlar da vardır.*

- *Hızlanma (0,10): Elektrikli araçların sunduğu yüksek hızlanma performansı, sürüş keyfini artırır ancak günlük kullanım için en önemli kriter olmayabilir.*
- *Enerji Tüketimi (0,05): Daha düşük enerji tüketimi, daha uzun menzil ve daha düşük kullanım maliyeti anlamına gelir.*
- *Fiyat (0,10): Bütçe, her alıcı için önemli bir karar faktörüdür. Ancak, sadece fiyata odaklanmak yerine, toplam sahip olma maliyeti (TCO) de göz önünde bulundurulmalıdır.*

COPILLOT:

“Türkiye’de yaşayan biri olarak SUV model elektrikli araba alırken dikkate almanız gereken kriterleri belirlediniz. Elektrikli araç seçiminde bu kriterlerin ağırlıklarını belirlemek önemli. İşte Türkiye’de yaşayan bir uzman olarak önerdiğim ağırlıklar:

- *Menzil: 0,25*
- *Batarya: 0,15*
- *Güç: 0,10*
- *Şarj Süresi: 0,10*
- *Hızlanma: 0,10*
- *Enerji Tüketimi: 0,10*
- *Fiyat: 0,20*

Bu ağırlıklar, her kriterin önemini Türkiye’deki sürüş koşulları, elektrikli araç altyapısı ve kullanıcı tercihlerini göz önünde bulundurarak belirlenmiştir. Bu ağırlıkları kullanarak alternatif araçları çok kriterli karar verme yöntemleri ile değerlendirebilirsiniz. Hangi modelleri düşündüğünüzü belirtirseniz, birlikte karşılaştırma yapabiliriz.”

Yöntemler tarafından belirlenen kriter ağırlıkları kullanılarak, alternatif elektrikli SUV modellerinin sıralaması TOPSIS yöntemi ile gerçekleştirilmiştir. TOPSIS yöntemi, alternatifleri en iyi ve en kötü çözüme olan mesafelerine göre sıraladığı için, yapay zekâ sohbet robotları ve geleneksel yöntemler tarafından belirlenen ağırlıkların karar verme sürecine etkisi karşılaştırılmıştır. Elde edilen TOPSIS puanları ve sıralamalar Tablo 4’de verilmiştir.

Tablo 4. Modellerin Yöntem Ağırlıkları Kullanılmış TOPSIS Sıralama(S) ve Puanları (P)

Modeller	TOPSIS-CRITIC		TOPSIS-LBWA		TOPSIS-CHATGPT		TOPSIS-GEMİNİ		TOPSIS-COPILOT		TOPSIS-YZ_ORT	
	S	P	S	P	S	P	S	P	S	P	S	P
TOGG-T10X	1	0,72	1	0,77	1	0,72	1	0,70	1	0,71	1	0,71
TESLA-Model Y	4	0,53	5	0,57	5	0,39	7	0,36	5	0,46	5	0,39
KG Mobility	2	0,62	2	0,69	2	0,71	2	0,65	2	0,71	2	0,69
BMW-X1	6	0,38	6	0,28	7	0,33	5	0,40	6	0,31	7	0,35
Mercedes-EQB	7	0,28	7	0,23	6	0,36	6	0,37	7	0,28	6	0,35
MINI-Count.	5	0,53	4	0,59	4	0,45	4	0,41	4	0,48	4	0,44
MG-ZS EV	3	0,56	3	0,63	3	0,48	3	0,43	3	0,52	3	0,48

Tablo 4'e göre tüm yöntemlerin kriter ağırlıkları kullanılarak hesaplanan TOPSIS sıralamasında birinci alternatif model TOGG-T10X olarak bulunmuştur. Bu sonuç TOGG markasının Türkiye satış sayılarında ilk sırada yer almasıyla da uyushmaktadır. Tüm sıralamaların benzerliklerinin karşılaştırılması için yöntemlerin arasındaki sıra korelasyonları hesaplanmış ve Tablo 5 de verilmiştir.

Tablo 5. Yöntem Ağırlıkları Kullanılmış TOPSIS Sıralamaları Arasındaki İlişkiler

	TOPSIS-CRITIC	TOPSIS-LBWA	TOPSIS-CHATGPT	TOPSIS-GEMİNİ	TOPSIS-COPILOT	TOPSIS-YZ_ORT
TOPSIS-CRITIC	1	0,964	0,929	0,786	0,964	0,929
TOPSIS-LBWA		1	0,964	0,893	1,000	0,964
TOPSIS-CHATGPT			1	0,857	0,964	1,000
TOPSIS-GEMİNİ				1	0,893	0,857
TOPSIS-COPILOT					1	0,964
TOPSIS-YZ_ORT						1

Tablo 5'e göre, geleneksel yöntemler (CRITIC ve LBWA) ile yapay zekâ sohbet robotları (ChatGPT, Gemini, Copilot) tarafından belirlenen ağırlıklar kullanılarak elde edilen TOPSIS sıralamalarının yüksek oranda benzerlik gösterdiği görülmektedir. Ayrıca LBWA yöntemi ile yapay zekâ sohbet robotlarından elde edilen ortalama ağırlıklar (YZ_ORT) arasındaki korelasyon da 0,964 seviyesindedir. Bu durum, yapay zekâ tabanlı yöntemlerin subjektif değerlendirme yöntemleriyle benzer sonuçlar verebildiğini göstermektedir.

5. Sonuç

Bu çalışmada, yapay zekâ sohbet robotlarının (ChatGPT, Gemini, Copilot) ÇKKV süreçlerindeki etkinliği incelenmiş ve elde edilen sonuçlar geleneksel kriter ağırlıklandırma yöntemleri (CRITIC ve LBWA) ile karşılaştırılmıştır. Türkiye'de 2024 yılının ilk on ayında en çok satan elektrikli SUV modelleri üzerinden gerçekleştirilen analizde, belirlenen kriterlere göre alternatifler sıralanmış ve farklı yöntemlerin sonuçlar üzerindeki etkisi değerlendirilmiştir. Yapay zekâ tabanlı yöntemlerin ÇKKV süreçlerine entegre edilmesi, karar destek sistemlerinde yeni bir perspektif sunarken, yöntemin avantajları ve sınırlılıkları da çalışmada tartışılmıştır.

Yapay zekâ sohbet robotlarının ÇKKV süreçlerinde kullanımı, bazı önemli avantajlar sağlamaktadır. Bu sistemler büyük veri setlerini işleyerek hızlı analiz yapabilmekte, karar vericilere objektif ve veri odaklı öneriler sunabilmektedir. Ayrıca, insan uzmanlardan bağımsız olarak 7/24 erişilebilir olmaları ve çok çeşitli kriterleri dikkate alarak alternatif çözümler üretebilmeleri, karar verme süreçlerinde esneklik sağlamaktadır. Ancak, yapay zekâ tabanlı sistemlerin belirli kısıtları da bulunmaktadır. Modelin eğitildiği veri kümesi ve algoritmik yapı nedeniyle bazı durumlarda önyargılı veya bağlamsal olarak yetersiz sonuçlar üretebilme ihtimali vardır. Ayrıca, şeffaflık eksikliği ve karar mekanizmalarının tam olarak anlaşılamaması, geleneksel yöntemlerle karşılaştırıldığında güvenilirlik açısından soru işaretleri doğurabilmektedir.

Çalışmada kullanılan CRITIC ve LBWA yöntemleriyle karşılaştırıldığında, yapay zekâ sohbet robotlarının belirlediği kriter ağırlıklarının belirli benzerlikler gösterdiği görülmüştür. Nitekim, farklı ağırlıklar kullanılarak yapılan TOPSIS sıralamalarında da klasik yöntemler ile yapay zekâ sohbet robotları arasında yüksek düzeyde benzerlik tespit edilmiştir.

Tüm yöntemlerle elde edilen alternatif sıralamalar karşılaştırıldığında, TOGG-T10X modelinin en iyi alternatif olarak öne çıktığı görülmüştür. Yapay zekâ sohbet robotları ve geleneksel ÇKKV yöntemleriyle yapılan analizlerde, TOGG-T10X'in menzil, batarya kapasitesi ve fiyat-performans

dengesi açısından diğer modellere göre daha avantajlı olduğu ortaya çıkmıştır. Bu bulgu, yapay zekâ tabanlı yöntemlerin karar destek süreçlerinde kullanılabilirliğini göstermektedir.

Sonuç olarak, yapay zekâ sohbet robotlarının ÇKKV süreçlerine entegrasyonu, hızlı ve dinamik analizler sağlayarak karar verme süreçlerini destekleyebilir. Ancak, modelin güvenilirliği ve şeffaflığı konularında bazı kısıtların olduğu unutulmamalıdır. Gelecekte, yapay zekâ tabanlı karar destek sistemlerinin geliştirilmesi ve daha açıklanabilir hale getirilmesi, bu teknolojinin çok kriterli karar verme süreçlerinde daha yaygın ve etkin bir şekilde kullanılmasını sağlayabilir.

Kaynakça

- Abdulahitoğlu, A., Abdulvahitoğlu, A., & Kılıç, M. (2022). Elektrikli araç bataryalarının bütünlük swara-topsis metodu ile değerlendirilmesi. Çukurova Üniversitesi Mühendislik Fakültesi Dergisi, 37(4), 1061-1076.
- Ahangar, R. G., & Fietko, A. (2023). Exploring the potential of ChatGPT in financial decision making. In Advancement in business analytics tools for higher financial performance (pp. 94-111). IGI Global.
- Akoğul, S., & Gögebakan, M. (2020). Modele Dayalı Kümeleme Analizinde Optimum Kümeleme İçin Yeni Bir Yaklaşım. Mühendislik Bilimleri Ve Tasarım Dergisi, 8(5), 218-229.
- Akoğul, S., Erisoğlu, M., & Erişoğlu, Ü. (2020). Çok Değişkenli Normal Dağılımların Karmasına Dayalı Kümelemede TOPSIS Yöntemiyle Küme Sayısının Belirlenmesi. Erciyes Üniversitesi Fen Bilimleri Enstitüsü Fen Bilimleri Dergisi, 36(3), 472-480.
- Arıkan Kargı, V. S. (2024). Multi-Criteria Decision Making In The Selection of Electric Sports Utility Vehicles: Integrated Critic-Copras Method. Abant Sosyal Bilimler Dergisi, 24(2), 474-485.
- Basmacı, G. (2022). Prioritizing Challenges in AI Applications in Healthcare: Multi-Criteria Decision Making Approach. Nevşehir Hacı Bektaş Veli Üniversitesi SBE Dergisi, 12(1), 194-206.
- Büyükselçuk, E. Ç., & Tozan, H. (2022). Elektrikli Araçların Performanslarının CRITIC-EATWIOS ile Değerlendirilmesi. Düzce Üniversitesi Bilim ve Teknoloji Dergisi, 10(4), 1670-1688.
- ChatGPT. Erişim Tarihi: 05.11.2024, Erişim adresi: <https://chatgpt.com/>
- Copilot. Erişim Tarihi: 05.11.2024, Erişim adresi: <https://copilot.microsoft.com/>
- Çoşkun, İ. T. (2022). Çok kriterli karar verme teknikleri ile elektrikli otomobil seçimi: SDMULTIMOORA yaklaşımı. Üçüncü Sektör Sosyal Ekonomi, 57 (1), 68-82.
- Çoşkun, İ. T. (2023). Subjektif ve Objektif Karar Verme Teknikleri ile Elektrikli Araç Seçiminde Etkili Olan Kriterlerin Değerlendirilmesi. Çukurova Üniversitesi İktisadi Ve İdari Bilimler Fakültesi Dergisi, 26(2), 173-190
- Delice, E. K., & Akvıran, E. (2024). A New Integrated Criteria Weighting Approach for Determining Production Priority in Demand Management. Journal of Turkish Operations Management, 8(2), 442-458.
- Diakoulaki, D., Mavrotas, G., & Papayannakis, L. (1995). Determining objective weights in multiple criteria problems: The critic method. Computers & Operations Research, 22(7), 763-770.
- Gavcar, E., & Kara, N. (2020). Elektrikli Otomobil Seçiminde Entropi ve TOPSIS Yöntemlerinin Uygulanması. İş Ve İnsan Dergisi, 7(2), 351-359.

- Gemini. Erişim Tarihi: 05.11.2024, Erişim adresi: <https://gemini.google.com/>
- Göktaş, F., & Güçlü, F. (2024). Belirsiz kriter ağırlıkları altında yeni bir ÇKKV yöntemi: Yapay zekâ sohbet robotlarına (ChatGPT4, Copilot, Gemini) dayalı portföy seçimi üzerine bir uygulama. *Afyon Kocatepe Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 26, 68-80.
- Güler, A. (2024). Türkiye’de Satışa Sunulan Elektrikli Araçların BWM ve LOP-COW Yöntemleriyle Ağırlıklandırılması ve Kriterlerin Sıralanması. *Ordu Üniversitesi Bilim Ve Teknoloji Dergisi*, 14(1), 106-120.
- Güleryüz, S. & Çokyaşar, A. (2021). Otomobil Seçimi İçin TOPSIS Temelli Bir Karar Verme Yaklaşımı. *European Journal of Science and Technology*, (31), 713-724.
- Hwang, C. L. & Yoon, K. (1981). *Multiple Attribute Decision Making: Methods and Applications*. Springer-Verlag, New York.
- Kanmaz, A. K., & Ertuğrul, İ. (2024). Fuzzy PIPRECIA and CRADIS integrated method in electric vehicle selection. *Pamukkale Journal of Eurasian Socioeconomic Studies*, 11(1), 19-36.
- Karasu Asnaz, M. S., & Özdemir, B. (2021). Elektrikli Araç Şarj İstasyonlarının Çok Kriterli Karar Verme Yöntemleri ile Optimal Konumlandırması. *Akıllı Ulaşım Sistemleri Ve Uygulamaları Dergisi*, 4(2), 175-187.
- ODMD (Otomotiv Distribütörleri ve Mobilite Derneği). (Erişim Tarihi: 05.11.2024, Erişim adresi: <https://www.odmd.org.tr/>)
- Özekenci, E. K. (2024). Personnel selection based on the LBWA, TOPSIS and GRA Methods: A case study on foreign trade company. *Fiscaeconomia*, 8(2), 646-665.
- Pinochet, L. H. C., Moreira, M. Â. L., Fávero, L. P., dos Santos, M., & Pardim, V. I. (2023). Collaborative work alternatives with ChatGPT based on evaluation criteria for its use in higher education: Application of the PROMETHEE-SAPEVO-M1 method. *Procedia Computer Science*, 221, 177-184.
- Uluskan, M., Topuz, D., & Çimen, C. (2022). AHP, BULANIK AHP, LBWA VE COPRAS Yöntemleri ile Tedarikçi Değerlendirme: Demiryolu Sektöründe Bir Uygulama. *Eskişehir Osmangazi Üniversitesi Mühendislik ve Mimarlık Fakültesi Dergisi*, 30(3), 412-430.
- Wikipedia. Erişim Tarihi: 05.11.2024, Erişim adresi: <https://tr.wikipedia.org/wiki/Gemini> ; https://tr.wikipedia.org/wiki/Microsoft_Copilot ; <https://tr.wikipedia.org/wiki/ChatGPT>
- Žižović, M., & Pamucar, D. (2019). New model for determining criteria weights: Level Based Weight Assessment (LBWA) model. *Decision making: applications in management and engineering*, 2(2), 126-137.

ENTEGRE LOPCOW-APLOCO Yöntemleriyle Elektrikli Otomobil Seçimi Problemi

Mehmet Akif Kara¹

Özet

Enerji verimliliğinin sağlanması, ekolojik krize karşı mücadele gibi bir çok nedenden dolayı elektrikli araçlar hızla yaygınlaşmaktadır. Farklı otomobil firmalarının ürettiği bu araçlar tüketiciler için bir seçim problemi ortaya çıkarmaktadır. Bu çalışmada literatürde ilk kez kullanılan LOPCOW-APLOCO entegre çok kriterli karar verme yöntemleriyle 6 alternatif elektrikli araç 8 farklı kriter kullanılarak sıralanmaktadır. Kriterlerin ağırlıklandırılmasında objektif bir yöntem olan LOPCOW yöntemi kullanılmıştır. Seçim problemi için ise APLOCO yöntemi kullanılmıştır. Çalışma sonucunda kriterler önem sırasına göre; fiyat, hızlanma, enerji tüketimi, ağırlık, azami hız, menzil, motor gücü ve tork olarak bulunmuştur. Seçim probleminin çözümü sonrasında ise elektrikli araçlar A2, A1, A3, A6, A5, A4 şeklinde sıralanmaktadır.

1. Giriş

Çok kriterli karar verme yöntemleri (ÇKKV), giderek karmaşıklaşan ve zorlaşan, stratejik öneme sahip karar problemlerinin çözümlemesinde kullanılmaktadır. Bu tür yöntemler, genel itibariyle karar vericilerin farklı ve çelişen hedefleri dikkate alması gereken durumlarda tercih edilmektedir. Birden fazla alternatif ve kriterin olduğu bu problemlerin sistematik ve mantıklı bir şekilde çözülmesi gereklidir. ÇKKV, yöneylem araştırması ve yönetim bilimi alanlarının bir alt dalı olarak, karar teorisi ve karar analizinin en yaygın kullanılan yöntemlerinden birisidir (Atan ve Yılmaz, 2020). Bu bağlamda; yatırım kararları, stratejik planlama, seçim problemleri, çevresel değerlendirmeler gibi farklı konularda, iktisat, işletme, sağlık, mühendislik gibi farklı bilim alanlarında kullanılmaktadır.

1 Dr.Öğr.Üyesi, Giresun Üniversitesi İİBF İşletme Bölümü Sayısal Yöntemler ABD, akifkara28@gmail.com, ORCID: 0000-0003-4308-9933.

Bu açıdan değerlendirdiğimizde özellikle teknolojinin gelişimi ile birlikte doğanın sürdürülebilirliğinin uyumunun sağlanması elzem bir ihtiyaç haline gelmiştir. Özellikle artan küresel nüfus ve fosil yakıtların hızla azalması nedeniyle karbondioksit emisyonları yoluyla sera gazı üretilmekte ve bu küresel ısınmaya neden olmaktadır. Sera gazı emisyonlarının en büyük yüzdesi de iklim değişikliğine bu minvalde katkıda bulunan ulaştırma sektöründen kaynaklanmaktadır (Vassileva ve Campillo, 2016). Dolayısı ile doğanın sürdürülebilirliğinin sağlanması için doğaya zarar vermeyen enerji türlerinin de kullanımının yaygınlaşması beklenmekte ve buna ilişkin çalışmalar yürütülmektedir. Elektrikli araç kullanımı da bu çabanın bir ürünü olarak değerlendirilebilir. Son 1 yılda farklı markaların elektrikli araç üretimlerinin arttığını göz önüne aldığımızda satış rakamlarının da bu minvalde arttığını görebilmekteyiz. Keza, elektrikli araçların 2030 yılına kadar uluslararası pazarda binek otomobil satışlarının %75'ini oluşturması beklenmektedir (Güler, 2024). Bu çalışmada da piyasada satışı yapılan elektrikli araçlara ilişkin çeşitli kriterler bağlamında bir seçim problemi oluşturulmakta ve literatüre görece yeni kazandırılan LOPCOW ve APLOCO yöntemleriyle çözümü gerçekleştirilmektedir. Çalışma da entegre LOPCOW-APLOCO yöntemleri literatürde ilk kez kullanılmaktadır.

Çalışmanın motivasyonunu görece yeni bir pazar olan elektrikli araç piyasasının durumunu ortaya koymak ve muhtemel tüketiciler için kolaylık sağlayacak bir sonuç elde etme fikrini oluşturmaktadır. Aynı zamanda LOPCOW-APLOCO yöntemlerinin entegre olarak ilk kez kullanılacak olması da ikincil motivasyon içerisinde yer almaktadır.

Elektrikli araç seçimine ilişkin literatürde çok az çalışma olmakla birlikte genel olarak araç seçimine ilişkin ÇKKV yöntemlerinin uygulandığı çok sayıda çalışma vardır. Gavcar ve Kara (2020) yaptıkları çalışmada 5 farklı kriter kullanmıştır. Çalışmalarında yöntem olarak ENTROPİ ve TOPSİS yöntemlerini kullanmış ve “batarya kapasitesi” en önemli kriter olarak bulunmuştur. Çalışmada 11 farklı araç modeli için seçim problemi gerçekleştirilmiştir. Sonar ve Kulkarni (2021) AHP ve MABAC yöntemlerini kullandıkları çalışmalarında altı elektrikli araç seçimi için altı kriteri ağırlıklandırmıştır. Çalışma sonucundan en önemli kriter “menzil” olarak tespit edilmiştir. Öztayşı vd. (2021) çalışmalarında elektrikli araç seçim problemi için KEMIRA yönteminin bir uzantısı olan Bulanık KEMIRA (F-KEMIRA) yöntemi ile 6 kriter ve 5 alternatif içeren bir karar modeli oluşturmuştur. Çalışmada kullanılan kriterlerden üçü maliyet üçü ise fayda yönündür. Çaloğlu Büyükselçuk ve Tozan (2022) yaptıkları çalışmalarında elektrikli araçların verimliliğini CRITIC ve EATWIOS yöntemlerini kullanarak değerlendirmişlerdir. Çalışma sonucunda enerji tüketimi, fiyat

ve araç ağırlığının önemli faktörler olduğu belirlenmiş ve İngiliz marka bir aracın en iyi tercih edilebilir elektrikli araç olduğu sonucuna ulaşılmıştır. Güler (2024) yaptığı araştırmasında elektrikli araçları BWM ve LOPCOW yöntemleriyle ağırlıklandırmakta ve kriterleri sıralamaktadır. Ardından Ortak Ağırlık Yöntemi ile nihai kriter ağırlıklarını elde etmiştir. Çalışmada kriterler arasında en önemlisi “batarya kapasitesi” olarak bulunmuştur. “Ağırlık” kriteri ise en önemsiz kriter olarak bulunmuştur.

2. Metodoloji

Çalışmanın bu bölümünde kullanılan yöntemlere ilişkin matematiksel adımlar, çalışmada kullanılan veri seti ve değişkenler hakkında bilgi verilecektir.

2.1. Yöntemler

Çalışmada kriterlerin ağırlıklandırılmasında LOPCOW yöntemi, elektrikli otomobil seçiminde ise APLOCO yöntemi kullanılmış ve yöntemlere ilişkin adımlara yer verilmiştir.

2.1.1. LOPCOW Yöntemi

LOPCOW yöntemi kriter ağırlıklandırma yöntemleri arasında yer almaktadır. 2022 yılında Ecer ve Pamucar (2022) tarafından geliştirilmiştir. LOPCOW yöntemi objektif yani karar vericinin görüşlerine ihtiyaç duymadan kriter ağırlığını belirleyen bir yöntemdir. Alternatiflerin negatif performans değerlerinin dikkate alınması, kriter ağırlıklarının belirlenmesinde önemli bir rol oynamakta ve çok sayıda kriter ile alternatifin etkin bir şekilde analiz edilmesini sağlamaktadır (Biswas vd., 2022). LOPCOW yöntemi, her bir kriterin standart sapmasını hesaplamak için, alternatif sayısına bağlı olarak logaritmik bir fonksiyon kullanarak yüzde değerlerini elde etmekte ve böylece kriterlerin daha önemli ve daha az önemli olanları arasındaki farkları daha uygun bir şekilde sunmaktadır.

Yöntem 3 adımdan oluşmaktadır. İlk olarak; karar probleminin belirlenmesi ve çözümlenmesi için m alternatif ve n kriterden oluşan bir içsel karar matrisi oluşturulmaktadır.

1. adım: Normalize karar matrisinin oluşturulması

$$r_{ij} = \frac{x_{ij} - x_j^{\min}}{x_j^{\max} - x_j^{\min}} \quad \text{fayda yönlü kriterler için} \quad (1)$$

$$r_{ij} = \frac{x_j^{max} - x_{ij}}{x_j^{max} - x_j^{min}} \text{ maliyet yönlü kriterler için} \quad (2)$$

2. adım: : Her Kriter İçin Yüzdelik Değerlerin PV_{ij} Matrisinin Oluşturulması

$$PV_{ij} = \left| \ln \left(\frac{\sqrt{\frac{\sum_{i=1}^m r_{ij}^2}{m}}}{\sigma} \right) \right| 100 \quad (3)$$

3. adım: Objektif Ağırlıkların (W_j) Hesaplanması

$$w_j = \frac{PV_{ij}}{\sum_{i=1}^n PV_{ij}} \quad (4)$$

2.1.2. APLOCO Yöntemi

APLOCO yöntemi Bulut tarafından 2018 yılında geliştirilmiştir. Literatürde kompleks ağ analizi alanında düğüm ağırlıklarının hesaplanmasında ve merkezi metriklerin belirlenmesi (Fındık ve Özkaynak, 2019; Özkaynak, 2020; Fındık ve Özkaynak, 2021), sosyal ağ analizi (Mishra vd., 2020) ve bulaşıcı olmayan hastalıkların risk alanlarına göre karşılaştırılması (Bulut, 2024) alanlarında kullanılmıştır. Yöntem 5 adımdan oluşmaktadır.

Yöntemin adımları şu şekildedir (Bulut, 2018);

Adım 1. Karar Matrisinin oluşturulması

$$x_{ij} = \begin{pmatrix} \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1r} \\ x_{21} & x_{22} & \cdots & x_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ x_{r1} & x_{r2} & \cdots & x_{cr} \end{pmatrix} \end{pmatrix} \quad (5)$$

Adım 2: Başlangıç noktası kriter (SPC) değerlerinin hesaplanması

$$P_{ij} = \begin{cases} \max_i(x_{ij}) - x_{ij}; P_{ij}, \text{maksimum başlangıç noktası değeriye} \\ x_{ij} - \min_i(x_{ij}); P_{ij}, \text{minimum başlangıç noktası değeriye} \end{cases} \quad (6)$$

$$P_{ij} = \begin{pmatrix} \max_i(x_{ij}) + x_{11} & \cdots & \max_i(x_{ij}) + x_{1r} \\ \vdots & \ddots & \vdots \\ \max_i(x_{ij}) + x_{c1} & \cdots & \max_i(x_{ij}) + x_{cr} \end{pmatrix} \quad (7)$$

$$P_{ij} = \begin{pmatrix} x_{11} - \min_i(x_{ij}) & \cdots & x_{1r} - \min_i(x_{ij}) \\ \vdots & \ddots & \vdots \\ x_{c1} - \min_i(x_{ij}) & \cdots & x_{cr} - \min_i(x_{ij}) \end{pmatrix} \quad (8)$$

Adım 3: Logaritmik dönüşüm (LC) matrisinin oluşturulması

$$\ln_x = \log e \text{ ve } L_{ij} = \frac{1}{\ln(p_{ij}) + 2} \quad (9)$$

$$L_{ij} = \begin{pmatrix} \frac{1}{\ln(P_{11} + 2)} & \cdots & \frac{1}{\ln(P_{1r} + 2)} \\ \vdots & \ddots & \vdots \\ \frac{1}{\ln(P_{c1} + 2)} & \cdots & \frac{1}{\ln(P_{cr} + 2)} \end{pmatrix} \quad (10)$$

Adım 4: Kriterlerin ağırlıkları (WC)'nın belirlenmesi ve ağırlıklı logaritmik dönüşüm (WLC) matrisinin hesaplanması

$$T_{ij} = \begin{pmatrix} w_1 x l_{11} & \cdots & w_1 x l_{1r} \\ \vdots & \ddots & \vdots \\ w_n x l_{c1} & \cdots & w_n x l_{cr} \end{pmatrix} \quad (11)$$

Adım 5: En iyi alternatif (BA)'ın belirlenmesi

$$B_j = \max(T_{ij}) \quad (12)$$

$$B_{sj} = \sum_{j=1}^n \{l_1, l_2, \dots, l_n\} \quad (13)$$

$$\alpha s_i = \sum_{j=1}^n \{t_1, t_2, \dots, t_n\} \quad (14)$$

$$\theta_i = \frac{\alpha s_i}{B_{sj}} \quad (15)$$

2.2. Veri Seti ve Değişkenler

Çalışmada kullanılan veri seti araç markalarının web sitelerinden elde edilmiştir. Araçlar seçilirken SUV/JEEP segmentindeki benzer araçlar kullanılmıştır. Bu minvalde altı araç seçim problemine dâhil edilmiştir. Çalışmada kriter olarak; başlangıç fiyatı, azami hız, motor gücü, hızlanma, enerji tüketimi, tork, ağırlık ve menzil değişkenleri kullanılmıştır. Çalışmada kullanılan kriterlerin tanımları aşağıdaki gibidir:

Başlangıç Fiyatı; Araç satın alımında ödenecek ilk miktardır.

Azami hız; Araçların ulaşabileceği en yüksek hız. Genellikle km/saat (km/h) cinsinden ifade edilir ve aracın performansını gösterir.

Motor gücü; Elektrik motorunun sağladığı güç. Genellikle kilowatt (kW) cinsinden ölçülür. Motor gücü, aracın hızlanma kapasitesini ve genel performansını etkiler.

Hızlanma; Aracın belirli bir süre içinde (genellikle 0'dan 100 km/saat) ulaştığı hız. Hızlanma süresi, aracın dinamik performansı hakkında bilgi verir.

Enerji tüketimi; Aracın belirli bir mesafe (genellikle 100 km) için harcadığı enerji miktarı. KWh/100 km cinsinden ifade edilir ve araç verimliliğini gösterir.

Tork; Motorun dönerken ürettiği kuvvet. Genellikle Nm (Newton metre) cinsinden ölçülür. Tork, aracın hızlanma yeteneğini ve yokuşlarda performansını etkiler.

Ağırlık; Aracın toplam ağırlığı. Genellikle kg (kilogram) cinsinden ifade edilir. Ağırlık, aracın enerji tüketimi ve performansı üzerinde doğrudan etkilidir.

Menzil; Araçla bir şarjla kat edilebilecek maksimum mesafe. Genellikle km (kilometre) cinsinden ölçülür ve aracın kullanım alanını belirler.

Tablo 1’de çalışmada kullanılan veri setine yer verilmektedir.

Tablo 1. Veri Seti

	Fiyat (maliyet yönlü)	Azami Hız (Fayda yönlü)	Motor Gücü (Fayda yönlü)	Hızlanma (Fayda yönlü)	Enerji Tüketimi (maliyet yönlü)	Tork (fayda yönlü)	Ağırlık (maliyet yönlü)	Menzil (fayda yönlü)
A1	1880000	175	207	8,1	18,6	339	1915	635
A2	4200000	200	487	3,7	21,2	860	2348	490
A3	1614000	162	156	8,8	14,6	255	1760	512
A4	1664000	167	204	7,8	16,2	255	2105	460
A5	1800000	185	218	7,8	16,7	350	2126	523
A6	1880000	150	204	9,6	15,6	320	1930	489

3. Bulgular

3.1. Kriter ağırlıklarının belirlenmesi

Çalışmada kullanılan kriterlerin ağırlıklandırılması için LOPCOW yöntemi kullanılmaktadır. Tablo 1’de yer alan veri seti karar matrisi olarak kullanılmıştır. Eşitlik 1 ve Eşitlik 2’de yer alan denklemler ile karar matrisi normalize edilmiştir. Tablo 2’de gösterilmektedir.

Tablo 2. Normalize Edilmiş Karar Matrisi

	Fiyat (maliyet yönlü)	Azami Hız (Fayda yönlü)	Motor Gücü (Fayda yönlü)	Hızlanma (Fayda yönlü)	Enerji Tüketimi (maliyet yönlü)	Tork (fayda yönlü)	Ağırlık (maliyet yönlü)	Menzil (fayda yönlü)
A1	0,897	0,5	0,154	0,745	0,393	0,138	0,736	1
A2	0	1	1	0	0	1	0	0,171
A3	1	0,24	0	0,864	1	0	1	0,297
A4	0,980	0,34	0,145	0,694	0,757	0	0,413	0
A5	0,928	0,7	0,187	0,694	0,681	0,157	0,377	0,36
A6	0,897	0	0,145	1	0,848	0,107	0,710	0,165

Normalizasyon işleminin ardından yüzde değerleri hesaplanmaktadır. Tablo 3’de gösterilmektedir.

Tablo 3. Normalizasyon İşlemi Sonrası Yüzde Değerleri

Kriterler	Fiyat	Azami Hız	Motor Gücü	Hızlanma	Enerji Tüketimi	Tork	Ağırlık	Menzil)
Yüzde değerleri	259,03	216,18	159,78	255,74	241,15	141,52	234,36	185,93

Son olarak eşitlik 4 kullanılarak kriter ağırlıkları hesaplanmaktadır. Tablo 4’de gösterilmektedir.

Tablo 4. Kriter Ağırlıkları

Kriterler	Fiyat	Azami Hız	Motor Gücü	Hızlanma	Enerji Tüketimi	Tork	Ağırlık	Menzil)
Kriter ağırlıkları	0,152	0,127	0,094	0,150	0,142	0,083	0,138	0,109

Kriter ağırlıkları sırasıyla; fiyat, hızlanma, enerji tüketimi, ağırlık, azami hız, menzil, motor gücü ve tork olarak bulunmuştur.

3.2. Seçim probleminin çözülmesi

Tablo 1’de yer alan veri seti karar matrisi olarak kullanılmıştır. Eşitlik 6, eşitlik 7 ve eşitlik 8 kullanılarak başlangıç noktası kriter değerleri hesaplanmıştır. Tablo’da gösterilmektedir.

Tablo 5. Başlangıç Noktası Kriter Değerleri

	A1	A2	A3	A4	A5	A6
C1	2320000,00	0,00	2586000,00	2536000,00	2400000,00	2320000,00
C2	25,00	0,00	38,00	33,00	15,00	50,00
C3	51,00	331,00	0,00	48,00	62,00	48,00
C4	1,50	5,90	0,80	1,80	1,80	0,00
C5	2,60	0,00	6,60	5,00	4,50	5,60
C6	521,00	0,00	605,00	605,00	510,00	540,00
C7	433,00	0,00	588,00	243,00	222,00	418,00
C8	0,00	145,00	123,00	175,00	112,00	146,00

Eşitlik 9 ve eşitlik 10 kullanılarak logaritmik dönüşüm matrisi hesaplanmaktadır Tablo 6’da gösterilmektedir.

Tablo 6. Logaritmik dönüşüm matrisinin hesaplanması

	A1	A2	A3	A4	A5	A6
C1	0,07	1,44	0,07	0,07	0,07	0,07
C2	0,30	1,44	0,27	0,28	0,35	0,25
C3	0,25	0,17	1,44	0,26	0,24	0,26
C4	0,80	0,48	0,97	0,75	0,75	1,44
C5	0,66	1,44	0,46	0,51	0,53	0,49
C6	0,16	1,44	0,16	0,16	0,16	0,16
C7	0,16	1,44	0,16	0,18	0,18	0,17
C8	1,44	0,20	0,21	0,19	0,21	0,20

Ağırlıklı logaritmik dönüşüm matrisinin hesaplanması için eşitlik 11 kullanılmaktadır.

Tablo 7. Ağırlıklı logaritmik dönüşüm matrisi

	A1	A2	A3	A4	A5	A6
C1	0,010	0,221	0,010	0,010	0,010	0,010
C2	0,039	0,185	0,035	0,036	0,045	0,032
C3	0,024	0,016	0,136	0,024	0,023	0,024
C4	0,120	0,073	0,146	0,112	0,112	0,216
C5	0,093	0,205	0,066	0,073	0,076	0,070
C6	0,013	0,120	0,013	0,013	0,013	0,013
C7	0,023	0,199	0,022	0,025	0,026	0,023
C8	0,157	0,022	0,023	0,021	0,023	0,022

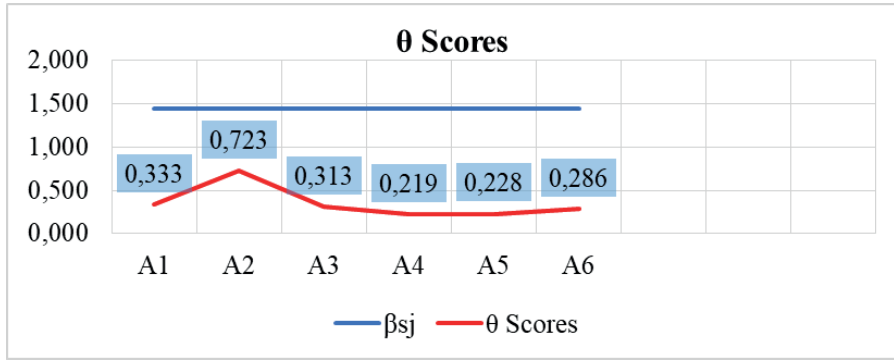
Daha sonra ise en iyi alternatif belirlenmektedir. Tablo 8, Tablo 9 ve Grafik 1’de gösterilmektedir.

Tablo 8. En iyi alternatifin belirlenmesi

	Bj	Alternatifler	asi
C1	0,221	A1	0,479
C2	0,185	A2	1,040
C3	0,136	A3	0,450
C4	0,216	A4	0,315
C5	0,205	A5	0,328
C6	0,120	A6	0,411
C7	0,199		
C8	0,157		
βsj	1,438		

Tablo 9. θ Skorlarının Hesaplanması Tablosu

Alternatif	θ Skorları	θ Skor Sıralaması
A1	1,443	0,333
A2	1,443	0,723
A3	1,443	0,313
A4	1,443	0,219
A5	1,443	0,228
A6	1,443	0,286

Grafik 1. θ Skorlarının Gösterilmesi

APLOCO yöntemi ile yapılan sıralama sonucunda A2, A1, A3, A6, A5, A4 şeklindedir.

4. Sonuç ve Tartışma

ÇKKV yöntemleri seçim, temin, sıralama vb. konularda yaşanan muhtemel sorunları en aza indirmek ve problemin çözümünü matematiksel adımlar kullanarak çözen yaklaşımlardır. Hem uygulayıcılar hem de paydaşlar açısından bu tür karar problemlerini kolaylaştıran yöntemlerdir.

Bu çalışmada ilk kez LOPCOW-APLOCO yöntemleri entegre olarak kullanılmıştır. Elektrikli araç seçim problemine uygulanan bu yöntemler optimum sonuca ulaşmaya çalışılmıştır. Her iki yönteminde farklı problemlerde kullanıldığı bilinmektedir.

Çalışmada kriter olarak fiyat, azami hız, motor gücü, hızlanma, enerji tüketimi, tork, ağırlık ve menzil olarak belirlenmiştir. Ayrıca, alternatif olarak 6 farklı elektrikli araç kullanılmıştır. LOPCOW yöntemi ile yapılan kriter ağırlıklandırma sonucunda; kriter ağırlıkları sırasıyla; fiyat, hızlanma, enerji tüketimi, ağırlık, azami hız, menzil, motor gücü ve tork olarak bulunmuştur.

Literatür ile karşılaştırıldığında ise kriterlerin ağırlık sıralamaları farklıdır. Gavcar ve Kara (2020) çalışmasında batarya kapasitesi, Sonar ve Kulkari (2021) çalışmasında menzil, Çaloğlu Büyükselçuk ve Tozan (2022) menzil, Güler (2024) ise batarya kapasitesi kriterlerini en önemli kriter olarak değerlendirmiştir. Ağırlıkların farklı çıkması veri seti, kullanılan yöntem gibi nedenlere dayanmaktadır. APLOCO yöntemi yapılan seçim işlemi sonrasında ise sırasıyla A2, A1, A3,A6, A5, A4 şeklindedir. Kriter ağırlıkları sıralama üzerinde doğrudan etkilidir. Fiyat kriterinin ağırlığı en yüksek kriter olması nedeniyle A2 alternatifinin ilk sırada olmasında etkilidir.

Kaynakça

- Atan, M. ve Yılmaz, E. (2020). Karar verme ve karar teorisi. *Örnek uygulamalarla çok kriterli karar verme yöntemleri* içinde (Edt. M.Atan ve Ş.Atan). 3-20, Gazi Kitabevi.
- Biswas, S., Bandyopadhyay, G., & Mukhopadhyaya, J. N. (2022b). A multi-criteria framework for comparing dividend pay capabilities: Evidence from Indian FMCG and consumer durable sector. *Decision Making: Applications in Management and Engineering*, 5(2), 140-175.
- Bulut, T. (2018). A New Multi Criteria Decision Making Method: Approach of Logarithmic Concept (APLOCO), *International Journal of Artificial Intelligence and Applications*, 9(1), 15-33.
- Bulut, T. (2024). Comparison of countries in European region according to risk factors of noncommunicable diseases by APLOCO method. *ADYÜ Sağlık Bilimleri Dergisi*, 10(3), 257-267.
- Çaloğlu Büyükselçuk, E. ve Tozan, H. (2022). Elektrikli araçların performanslarının CRITIC-EATWIOS ile değerlendirilmesi. *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, 10(4), 1670-1688.
- Ecer, F. ve Pamucar, D. (2022). A novel LOPCOW-DOBI multi-criteria sustainability performance assessment methodology: An application in developing country banking sector. *Omega*, 112, 102690.
- Fındık, O. ve Özkaynak, E. (2019). Node Weighting Method in Centrality Measure of Complex Networks. International Conference on Advanced Technologies, Computer Engineering and Science (ICATCES 2019), Apr 26-28, 2019 Alanya, Turkey, 342-348.
- Fındık, O. ve Özkaynak, E. (2021). Link Prediction Based on Node Weighting in Complex Networks. *Soft Comput.* 25, 2467-2482.
- Gavcar, E., ve Kara, N. (2020). Elektrikli otomobil seçiminde ENTROPI ve TOPSIS yöntemlerinin uygulanması. *Journal of Human and Work*, 7(2), 351-359.
- Güler, A. (2024). Türkiye’de satışa sunulan elektrikli araçların BWM ve LOPCOW yöntemleriyle ağırlıklandırılması ve kriterlerin sıralanması. *Ordu Üniversitesi Bilim Ve Teknoloji Dergisi*, 14(1), 106-120.
- Mishra, A. K., Joshi, N., Mathur, I. (2020). A Fuzzy Based Integrated Model For Identification Of Vital Node In Terrorist Network Using Logarithmic Concept, *Journal of Intelligent & Fuzzy Systems*, 39(3), 3617-3631.
- Özkaynak, E. (2020). Karmaşık Ağlarda Düğüm Ağırlıklı Bağlantı Yöntemlerinin Geliştirilmesi. Doktora Tezi, Karabük Üniversitesi Bilgisayar Mühendisliği.

- Öztayşı, B., Onar, S. Ç. ve Kahraman, C. (2021). Electric vehicle selection by using Fuzzy KEMIRA. *Multiple-Valued logic and soft computing*, 37(3-4), 437-461.
- Sonar, H. C. ve Kulkarni, S. D. (2021). An integrated AHP-MABAC approach for electric vehicle selection. *Research in Transportation Business & Management*, 41, 100665, 1-8.
- Vassileva, I., & Campillo, J. (2017). Adoption barriers for electric vehicles: Experiences from early adopters in Sweden. *Energy*, 120, 632-641.

Makine Öğrenmesi ile Üniversite Öğrencilerinde Depresyonun Tahmini: Akademik, Psikolojik ve Demografik Faktörlerin Rolü

Hasan Aykut Karaboğa¹

Abdülkadir Keskin²

Özet

Depresyon, üniversite öğrencileri arasında yaygın olarak görülen ve akademik başarıyı, sosyal ilişkileri ve genel yaşam kalitesini olumsuz etkileyen bir ruh sağlığı sorunudur. Bu çalışmada, farklı makine öğrenmesi algoritmalarını kullanarak üniversite öğrencilerinin depresyon durumunu tahmin etmek amaçlanmıştır. Araştırmada bağımlı değişken olarak depresyon durumu (var/yok) ele alınmış ve bağımsız değişkenler olarak yaş, cinsiyet, not ortalaması, uyku süresi, meslek, akademik baskı, çalışma memnuniyeti, iş memnuniyeti, beslenme alışkanlıkları gibi değişkenler kullanılmıştır.

Çalışmada, Destek Vektör Makineleri (SVM), Yapay Sinir Ağları (Neural Networks), Lojistik Regresyon, k-En Yakın Komşu (kNN) ve Gradient Boosting algoritmaları kullanılarak depresyon tahmini gerçekleştirilmiştir. Sonuçlar, Lojistik Regresyon ve SVM modellerinin en iyi genelleme yeteneğine sahip olduğunu göstermektedir. Öte yandan, Yapay Sinir Ağları ve Gradient Boosting algoritmalarının eğitim verisinde yüksek performans sergilediği ancak test verisinde aşırı öğrenme (overfitting) nedeniyle başarının düştüğü belirlenmiştir. Bu çalışma, makine öğrenmesi yöntemleriyle üniversite öğrencilerinde depresyon tahminine yönelik yeni bir bakış açısı sunmakta ve öğrenci ruh sağlığını iyileştirmeye yönelik veri odaklı yaklaşımlar geliştirilmesine katkıda bulunmaktadır.

- 1 Dr. Öğr. Üyesi, Amasya Üniversitesi, Eğitim Fakültesi, Eğitim Bilimleri Bölümü, h.aykut.karaboga@amasya.edu.tr, ORCID: 0000-0001-8877-3267
- 2 Arş. Gör. Dr., İstanbul Medeniyet Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, İstatistik Bölümü, abdulcadir.keskin@medeniyet.edu.tr, ORCID: 0000-0002-4795-1028

1. Giriş

Gençlerde depresyon hayat kalitesini düşüren ve gençlerin gelecek planlarını alt üst edebilen bir sorundur. Okul hayatları devam eden bireylerde akademik başarıları düşmekte, kaygı düzeyleri artmakta ve yaşam tarzları olan genel alışkanlıkları nedeniyle sosyal alışkanlıkları ve genel refah düzeyleri depresyon nedeniyle bozulmaktadır (Astafeva vd., 2024; Gillooly vd., 2024). Özellikle eğitimlerinde ilerlemeler kaydetmiş ve çalışma hayatına yakın olan üniversite öğrencileri arasında depresyon daha yaygın bir ruh sağlığı sorunudur. Üniversite döneminde genç yetişkinlik çağında bulunan öğrenciler hayatlarının en önemli döneminde akademik ve sosyal baskılarla karşı karşıya kalmaktadır. Bu da akademik olarak tatmin olmamalarına ve baskı hissetmelerine neden olmaktadır (Mirhosseini vd., 2021). Hissedilen baskılar, öğrencilerin psikolojik sağlıkları üzerinde olumsuz etkilere yol açabilmektedir. Bu durum depresyonun üniversite öğrencileri arasında daha yaygın bir sorun olarak ortaya çıkmasına neden olmaktadır (İbrahim vd., 2013; Trigueros vd., 2020). Problemi çözmek veya önerilerde bulunabilmek için depresyona katkıda bulunan faktörleri belirlemek ve erken müdahale stratejileri tasarlamak bu alandaki en değerli çalışmalardandır.

Depresyonu etkin biçimde tahmin ederek doğru zamanda müdahalelerde bulunmak, belirtilen bu problemlerin erkenden çözüme kavuşturulmasını sağlayabilecektir. Her ne kadar kişilerle birebir görüşülerek ya da çeşitli anketler yoluyla depresyonu tespit etmeye çalışan geleneksel yöntemler hala etkin olsa da erişilebilirlik sorunları ve yanıt yanlılığı gibi sorunlar nedeniyle sınırlı başarı gösterebilmektedir (İbrahim vd., 2013; Pedrelli vd., 2015). Yapay zekâ ve veri bilimindeki gelişmeler makine öğrenmesi yöntemlerinin depresyon riski taşıyan kişileri belirlemek veya depresyondaki kişileri tahmin etmek için güçlü altyapı sağlamaktadır. Ele alınan büyük ölçekli veri kümelerinden yararlanarak sosyolojik, akademik, psikolojik ve demografik faktörler yardımıyla makine öğrenmesi modelleri zihinsel sağlık sorunları yaşayan öğrenciler için erken teşhis ve kişiselleştirilmiş yardım önerileri sağlayabilir (Shatte vd., 2019; He, 2024).

Makine öğrenmesi, klasik yöntemlerle ulaşılması ve analizi zor olan sosyal medya verileri, yapılan anketlerin sonuçları ve kişilerin sanal âlemde bıraktıkları dijital izler gibi çeşitli kaynaklardan elde edilen verileri işleyerek bireylerin ruhsal durumunu ve varsa depresyon varlığını tahmin etme potansiyeline sahiptir. Dijital veriler, anketlerle desteklenerek depresyonun semptomlarının zaman içindeki seyrini izlemeye de olanak tanır. Bu bağlamda, makine öğrenmesi algoritmaları, bireylerin ruh halini etkileyen çok sayıda faktörü dikkate alarak daha doğru ve zamanında tahminler yapabilmektedir

(Liu vd., 2024). Makine öğrenmesi algoritmalarının kişilerin depresyon tahmininde klasik değerlendirme yöntemleri ile bütünleştirilmesi, özneliği en aza indirirken öngörü doğruluğunu artıran daha veri odaklı bir yaklaşıma olanak tanır. Böylece, depresyonun erken teşhis edilmesi ve bireylere uygun müdahalelerin uygulanması sürecinde makine öğrenmesi uygulamaları, ruh sağlığı profesyonellerine önemli derecede yardımcı olabilir (He, 2024).

Bu çalışmada yaş, cinsiyet, akademik performans, uyku süresi, iş ve akademik baskılar, çalışma memnuniyeti, beslenme alışkanlıkları ve coğrafi konum gibi bir dizi bağımsız değişkene dayalı olarak üniversite öğrencilerinde depresyon durumunu tahmin etmek için makine öğrenmesi yöntemlerinin kullanılabilirliğini araştırmaktadır. Çalışmanın öncelikli amacı öğrencileri depresyonda veya değil şeklinde ikili sınıflandırmada farklı makine öğrenmesi algoritmalarının etkinliğini değerlendirmektir. Bir diğer amaç ise model performans değerlendirmesinin yanında öğrenci depresyonunu arttıran faktörleri belirleyerek üniversitelerin ve politika yapıcılarının öğrencilere yönelik baskıları azaltıcı müdahaleler geliştirmesine yardımcı olabilmektir.

Çalışma 5 bölümden oluşmaktadır: ikinci bölümde öğrenciler arasındaki depresyon tahmininde kullanılan makine öğrenmesi uygulamalarıyla ilgili literatür incelemesi sunulmuştur. Üçüncü bölümde veri ön işleme, özellik seçimi ve model değerlendirmesi için kullanılan metodolojik yaklaşımlar ayrıntılı olarak verilmiştir. Dördüncü bölümde makine öğrenmesi modellerinin karşılaştırmalı analiz sonuçları gösterilmiştir. Son olarak, beşinci bölümde elde edilen temel bulguları, çalışmanın sınırlılıkları ve gelecekteki araştırma önerileri tartışılmıştır.

2. Literatür İncelemesi

Depresyon, dünya genelinde hemen hemen 280 milyon kişiyi etkileyen zorlu bir ruhsal hastalıktır (Dünya Sağlık Örgütü, 2021). Genel olarak her yaştan insanda görülebilse de üniversite öğrencileri arasında depresyon oranları finansal ve akademik baskılar ve sosyal beklentiler nedeniyle çok daha yüksektir (Hunt ve Eisenberg, 2010; Assari, 2018). Üniversite öğrencilerinin %34,5'i stres belirtileri gösterirken %23,6 ile %18,4 arası öğrenciler kaygı ve depresyon belirtileri göstermektedir. Bu durum genel olarak aile ilişkileri, üniversiteye uyum zorluğu ve bağımsız problem çözme becerisinin yeterliliği gibi faktörlerle ilişkili olduğu belirtilmiştir (Ramón-Arbués vd., 2020).

Çalışmalar, yaşın, cinsiyetin, uyku düzenlerinin, akademik stresin ve iş-yaşam dengesinin öğrencilerin ruh sağlığını önemli ölçüde etkilediğini göstermiştir (Beiter vd., 2015). Ayrıca, öğrencilerin geleceğe dair kaygıları, iş

yaşamına hazırlık süreçleri ve ailelerin onlara karşı tutumları gibi faktörler de depresyon belirtileri üzerinde etkili olmaktadır (Ozen vd., 2010). Ergenlikten genç yetişkinlik dönemine geçen üniversiteli bireyler bulundukları farklı ortamlardan ötürü sosyal, ekonomik, biyolojik ve psikolojik faktörlerin değişimi nedeniyle ruhsal hastalıklara karşı duyarlılık yaşarlar (Eisenberg vd., 2013). Bu durum üniversite öğrencilerinin genel nüfusa göre daha yüksek bir depresyon riski taşıdığını göstermektedir (İbrahim vd., 2013).

Ruhsal sorunlarda önemli sosyo-demografik faktörlerin yaş, cinsiyet ve ekonomik durum olduğu söylenebilir. Genç öğrenciler bulundukları ortamlara uyum sağlama konusunda yaşadıkları zorluklar nedeniyle daha büyük duygusal sıkıntılarla karşı karşıya kalırlar (Bayram ve Bilgel, 2008). Depresyon yaygınlığında kadınlar erkeklerden daha yüksek oranda depresif semptomlar göstermektedir (Piccinelli ve Wilkinson, 2000; Ceyhan ve Ceyhan, 2011). Biyolojik, sosyal ve kültürel faktörler bu farklılığa katkıda bulunur; çünkü kadınlar genellikle duygusal sıkıntıyı ifade etme ve yardım arama olasılıkları daha yüksekken, erkekler toplumsal beklentiler nedeniyle semptomları eksik bildirebilir (Hyde vd., 2008). Bu yönüyle erkek öğrenciler, kadın öğrencilere göre daha yüksek yalnızlık ve depresyon seviyelerine sahiptir. Çünkü depresyonun nedeniyle yalnızlaşan öğrencilerin sosyal etkileşimleri olumsuz etkilenmekte ve destek alabilecekleri sosyal destek ağları zayıflayacaktır (Ceyhan ve Ceyhan, 2011).

Öğrencinin yaşadığı bölgenin çevresel, sosyoekonomik ve kültürel faktörleri de ruh sağlığını olumsuz yönde etkileyebilir. Çalışmalar, kentsel öğrencilerin genellikle daha yüksek akademik stres ve rekabet yaşadığını, kırsal öğrencilerin ise sosyal izolasyon ve daha az ruh sağlığı kaynağıyla mücadele edebileceğini göstermektedir. Ayrıca üniversite öğrencilerinin karşılaştığı ekonomik zorlukların ruh sağlığı üzerindeki olumsuz etkileri ortaya konulmaktadır (Lund vd., 2010; Shao vd., 2020).

Akademik performans genellikle zihinsel refahla ilişkilendirilir ve hem yüksek hem de düşük başarı gösterenler depresyon riski altındadır. Düşük not ortalamalarına sahip öğrenciler öz saygı sorunları ve akademik baskı yaşayabilirken, yüksek başarı gösterenler mükemmeliyetçilik ve tükenmişlikle mücadele edebilir. Derse uyum ve başarı kaygıları ile mücadele öğrencilerin ruh sağlığı üzerinde önemli derecede olumsuz etki göstermektedir. (Ozen vd., 2009). Araştırmalar, akademik başarısızlık yaşayan öğrencilerin akranlarına kıyasla depresif semptomlar geliştirme olasılığının iki kat daha fazla olduğunu göstermektedir (Andrews ve Wilding, 2004).

İş ve akademik hayatı dengelemek, özellikle ders yüküyle birlikte yarı zamanlı işleri de idare eden üniversite öğrencileri için önemli bir stres kaynağı

olabilir (Arastaman vd., 2020). İş baskısı artan stres ve azalan akademik performansla ilişkilidir ve bu da bitkinliğe ve zihinsel sağlıkta düşüşe yol açar (Glozier, 2002). Benzer şekilde, genellikle sıkı teslim tarihleri, performans beklentileri ve başarısızlık korkusundan kaynaklanan akademik baskı, öğrenci depresyonunun önde gelen bir öngörücüsüdür (Arslan vd., 2009).

Bir öğrencinin çalışmalarından ve işinden duyduğu memnuniyet düzeyi, zihinsel sağlığını önemli ölçüde etkiler. Akademik yaşamdan duyulan memnuniyetsizlik strese ve depresif semptomlara yol açabilir (Hysaj vd., 2018). Aynı şekilde, çalışan öğrenciler arasında iş memnuniyetsizliği stresi artırabilir, genel yaşam memnuniyetini azaltabilir ve depresyon riskini artırabilir (Arastaman vd., 2020; Méndez vd., 2020; Satuf vd., 2018).

Uyku yoksunluğu, öğrenciler arasında ruh sağlığı sorunlarına en sık katkıda bulunan etkenlerden biridir. Araştırmalar, gecede altı saatten az uyuyan öğrencilerin depresyon ve anksiyete semptomları yaşama olasılığının daha yüksek olduğunu göstermektedir (Lund vd., 2010). Kötü uyku kalitesi, düzensiz uyku saatleri ve yatmadan önce aşırı ekran süresi de artan depresif semptomlarla bağlantılıdır (Baglioni vd., 2011; Musliner vd., 2016; Zhang vd., 2018). Beslenme koşulları da öğrencilerin ruh sağlığında kritik bir rol oynamaktadır. Bilinçsizce tüketilen işlenmiş gıdalar ve şeker açısından zengin diyetler daha yüksek depresyon oranlarıyla ilişkilendirilir (Jacka vd., 2010; Özenoğlu, 2018). Öğün atlama veya aşırı kafein ve fast food tüketme gibi sağlıksız beslenme alışkanlıkları olan öğrenciler, ruh sağlığı sorunları açısından daha büyük risk altındadır (O'Neil vd., 2014; Tokpunar, 2024).

Tüm bu değişkenler ve durumlar depresyon ve ruh sağlığı sorunları ile ilgili çalışmaları daha kritik ve daha karmaşık hale getirmektedir. Bu gibi durumlarda önleyici müdahaleler yapabilmek için tahmin yeteneği güçlü modellere ihtiyaç duyulmaktadır. Yapay zeka alanındaki gelişmeler, makine öğrenme teknikleri kullanılarak ruh sağlığı bozuklukları için güçlü tahmin modellerinin geliştirilmesini kolaylaştırmıştır (Thieme vd., 2020; Ahmed vd., 2020). Ruh sağlığı çalışmalarında makine öğrenmesi uygulamaları, metin analizi (Baminiwatta, 2022), sosyal medya aktivitesi (Villegas ve Martinez, 2022), giyilebilir cihazlar (Williamson vd., 2024) ve anket tabanlı veri kümeleri (Ferrari vd., 2018; Karam vd., 2014) gibi yaklaşımlarla kullanılabilir (D'Ambrosio vd., 2023).

Sınıflandırma görevleri için en basit ancak etkili yöntemlerden biri olan lojistik regresyon, yorumlanabilir olması ve verimliliği nedeniyle depresyon tahmini için kullanılabilir. Ayrıca SVM modelleri, yüksek boyutlu verilerdeki sağlamlıkları nedeniyle depresyon tespitinde kullanılabilir. Araştırmalar, anket ve klinik veriler üzerinde eğitilen SVM modellerinin yüksek doğruluk

oranı ile çalışabildiğini göstermektedir (Shatte vd., 2019). Çok katmanlı algılayıcılar (MLP) gibi modeller de yüksek doğruluk sunmasına rağmen kara kutu sayılan yöntemlerdir (Feng vd., 2019). Ayrıca k-En Yakın Komşu (kNN) ve Gradient Boosting algoritmaları da klasik makine öğrenmesi yöntemlerinin etkin yaklaşımlarındandır. Bu yöntemler ile depresyonda olan ve olmayan bireylerin tahmin başarısı irdelenerek ruh sağlığı araştırmalarında bu yöntemlerin kullanılabilirliği araştırılacaktır.

3. Materyal ve Metot

Bu çalışmada kullanılan veri seti, Kaggle platformundan alınmış olup 1500 bireyin demografik, akademik ve psikolojik özelliklerini içermektedir (Shamim, 2024). Veri setinin bağımlı değişkeni “Depresyon” olup, bireylerin depresyon durumlarını (1 = depresyon var, 0 = depresyon yok) olarak sınıflandırmaktadır. Depresyon değişkeninin dağılımına bakıldığında, bireylerin %58,6’sının depresyon belirtileri gösterdiği gözlemlenmiştir. Bu durum, veri setinde depresyonda olan bireylerin, olmayanlara kıyasla daha fazla olduğunu göstermektedir.

Bağımsız değişkenler bireylerin demografik, akademik, psikolojik ve sosyal özelliklerini kapsamaktadır. Yaş (Age) değişkeni sürekli bir değişken olup, veri setindeki bireylerin yaş ortalaması 25,74 olarak hesaplanmıştır. Bu değer, örneklemdaki bireylerin çoğunlukla genç yetişkinlerden oluştuğunu göstermektedir. Akademik baskı (Academic Pressure) değişkeninin ortalama değeri 3,16 olup, bu da bireylerin akademik sorumluluklarını orta düzeyde stres verici olarak algıladıklarını göstermektedir. Benzer şekilde, bireylerin finansal stres (Financial Stress) düzeyi ortalama 3,10 olup, birçok bireyin finansal sıkıntılar yaşadığını göstermektedir.

Akademik başarı açısından değerlendirildiğinde, bireylerin Genel Not Ortalaması (CGPA) 7,69 olup, en düşük not 5,06 ve en yüksek not 10’dur. Çalışma süresi açısından bakıldığında, bireylerin günlük ortalama çalışma süresi 7,14 saat olarak belirlenmiştir. Ancak bireyler arasında geniş bir dağılım bulunmaktadır; bazı bireyler hiç ders çalışmazken, bazıları günde 12 saate kadar çalışmaktadır. Çalışma memnuniyeti (Study Satisfaction) ortalama 2,91 olup, bireylerin akademik çalışmalarından tam anlamıyla tatmin olmadıklarını göstermektedir.

Psikolojik faktörler arasında bireylerin uyku düzenleri önemli bir yer tutmaktadır. Uyku süresi (Sleep Duration) kategorik bir değişken olup, bireylerin %29,4’ü günde 5 saatten az uyuduğunu belirtmiştir. Bu da akademik ve psikolojik baskıların uyku süresini kısaltabileceğini düşündürmektedir. Beslenme alışkanlıkları (Dietary Habits) açısından bakıldığında, bireylerin

%36,4'ü sağlıklı beslendiğini belirtmiştir. Ruh sağlığıyla ilişkili bir diğer önemli değişken olan intihar düşüncesi (Suicidal Thoughts) bireylerin %63,1'inde gözlemlenmiştir. Bu oldukça yüksek bir oran olup, bireylerin ruhsal iyi oluş düzeylerinin düşük olabileceğini göstermektedir.

Aile geçmişi de bireylerin ruhsal sağlığını etkileyebilecek önemli faktörlerden biridir. Ailede ruhsal hastalık geçmişi (Family History of Mental Illness) değişkeni incelendiğinde, bireylerin %51,1'inin ailesinde herhangi bir ruhsal hastalık geçmişi olmadığını belirtmesi dikkat çekicidir. Bu durum, depresyonun sadece genetik faktörlerden değil, çevresel ve akademik faktörlerden de etkilenebileceğini göstermektedir.

Tablo 1. Tanımlayıcı İstatistikler

Değişken	Türü	Ortalama	Std Sapma	Min	Medyan	Max	Mod	En Sık Kategori (%)
Yaş (Age)	Sayısal	25,740	4,880	18	25	42	-	-
Akademik Baskı (Academic Pressure)	Sayısal	3,160	1,410	1	3	5	-	-
Genel Not Ortalaması (CGPA)	Sayısal	7,690	1,460	5,060	7,800	10	-	-
Çalışma Memnuniyeti (Study Satisfaction)	Sayısal	2,910	1,350	1	3	5	-	-
İş Memnuniyeti (Job Satisfaction)	Sayısal	0,002	0,077	0	0	3	-	-
Çalışma/Saatleri (Work/Study Hours)	Sayısal	7,140	3,740	0	8	12	-	-
Finansal Stres (Financial Stress)	Sayısal	3,100	1,460	1	3	5	-	-
Depresyon (Bağımlı Değişken)	Kategorik	-	-	0	1	1	1	58,60%
Cinsiyet (Gender)	Kategorik	-	-	-	-	-	Erkek	54,20%
Uyku Süresi (Sleep Duration)	Kategorik	-	-	-	-	-	5 saatten az	29,40%
Beslenme Alışkanlıkları (Dietary Habits)	Kategorik	-	-	-	-	-	Sağlıksız	36,40%
İntihar Düşüncesi (Suicidal Thoughts)	Kategorik	-	-	-	-	-	Evet	63,10%
Ailede Ruhsal Hastalık Geçmişi (Family History of Mental Illness)	Kategorik	-	-	-	-	-	Hayır	51,10%

Bu çalışmada, öğrencilerin depresyon durumunu tahmin etmek amacıyla denetimli makine öğrenmesi algoritmaları uygulanmıştır. Veri seti, bağımlı değişken olarak depresyon durumunu (1 = depresyon var, 0 = depresyon yok) ve bağımsız değişkenler olarak bireylerin demografik, akademik,

Tablo 2. Kullanılan algoritmaların özellikleri

Algoritma	Varsayılan Özellikler/Parametreler	Notlar
SVM	- Kernel: RBF - C: 1,0 - Gamma: Otomatik (1/özellik sayısı)	Veri ölçeklendirmesi önemlidir.
Yapay Sinir Ağları (MLP)	- Gizli Katman Sayısı: 1 - Nöron Sayısı: Otomatik belirlenir - Öğrenme Hızı: 0,3 - Momentum: 0,2 - Epoch: 500	Overfitting durumunda düzenleme (dropout, erken durdurma vb.) gerekebilir.
Lojistik Regresyon	- Regularizasyon: L2 - C: 1,0 - Çözümleyici: LBFGS - Maksimum Iterasyon: 100	Basit ve yorumlanabilir; lineer sınıflandırmada etkilidir.
k-En Yakın Komşu	- k (Komşu Sayısı): 5 - Mesafe Ölçütü: Öklid - Ağırlıklandırma: Eşit ağırlık	Özellik ölçeklendirmesi özellikle önemlidir.
Gradient Boosting	- Estimator Sayısı: 100 - Öğrenme Hızı: 0,1 - Maksimum Derinlik: 3	Hiperparametre ayarları performansı büyük ölçüde etkileyebilir.

Modeli tahmin etmek için kullanılan beş algoritmaya ait genel özellikler Tablo 2’ de gösterilmektedir. Bu tabloda, her bir algoritmanın varsayılan ayarları ve temel parametreleri detaylandırılmıştır. Örneğin, SVM algoritması için kullanılan RBF çekirdek, C parametresi ve gamma değeri, verinin doğrusal olmayan yapısını yakalayabilmek adına özenle seçilmiştir. Ancak bu ayarların veri ölçeklendirmesi gibi ön işleme adımlarıyla birlikte optimize edilmesi gerekmektedir. Neural Network (MLP) modelinde ise tek katmanlı gizli yapı, öğrenme hızı ve momentum gibi parametreler modelin hızlı öğrenmesini sağlarken, eğitim sürecinde aşırı öğrenme (overfitting) riskine karşı dikkatle izlenmelidir. Lojistik Regresyon, L2 regularizasyonu ve LBFGS çözümleyicisiyle basit fakat güçlü bir yaklaşım sunarak, özellikle lineer sınıflandırmada net ve yorumlanabilir sonuçlar elde edilmesine olanak tanır. k-En Yakın Komşu algoritması, belirlenen komşu sayısı ve kullanılan mesafe ölçütü ile örnekler arasındaki benzerlikleri değerlendirirken, özelliklerin uygun şekilde ölçeklendirilmesinin önemine vurgu yapar. Son olarak, Gradient Boosting algoritması, sınırlı derinlik ve belirli bir öğrenme hızı ile çoklu zayıf tahminicileri birleştirerek güçlü bir sınıflandırma performansı hedefler; fakat hiperparametre ayarlarının modelin genelleme kabiliyeti üzerinde doğrudan etkisi olduğu unutulmamalıdır. Bu özelliklerin her biri, modelin tahmin performansını, öğrenme hızı, genelleme yeteneğini ve aşırı

öğrenme riskini belirlemede önemli rol oynar. Dolayısıyla, her algoritmanın özellikleri ve varsayılan ayarları, uygulanan veri setine ve probleme göre dikkatlice değerlendirilmelidir.

3.1. Makine Öğrenmesi Algoritmalarının Performansı

Bu çalışmada, model performansını değerlendirmek için beş temel ölçüt kullanılmıştır: sınıflandırma doğruluğu (CA), F1 skoru, duyarlılık (recall), kesinlik (precision) ve eğri altındaki alan (AUC). AUC, ROC eğrisi ile ilişkili olup, modelin duyarlılık ve özgüllük dengesini ne kadar iyi sağladığını gösterir. Değerin 1'e yaklaşması, modelin yüksek performans sergilediğini gösterir. Sınıflandırma doğruluğu (CA), doğru tahmin edilen örneklerin toplam örnek sayısına oranını verir. Ancak, veri setinde dengesizlik olması durumunda tek başına yeterli bir değerlendirme ölçütü olmayabilir. F1 skoru, duyarlılık ve kesinlik arasındaki dengeyi sağlayarak modelin genel başarısını ölçer. Duyarlılık (recall), modelin pozitif sınıfları doğru şekilde tespit etme oranını ifade ederken, kesinlik (precision), modelin pozitif olarak tahmin ettiği örneklerin ne kadarının gerçekten pozitif olduğunu gösterir.

AUC (Area Under the Curve): AUC, ROC (Receiver Operating Characteristic) eğrisinin altında kalan alandır ve modelin sınıflandırma performansını ölçer. AUC değeri 0 ile 1 arasında değişir. 1'e yakın değerler, iyi bir sınıflandırma performansını gösterir. AUC, ROC eğrisinin altındaki alanı hesaplar. Matematiksel olarak belirli bir formülü yoktur, ancak ROC eğrisinin hesaplanmasıyla elde edilir.

CA (Correct Accuracy) / Doğru Doğruluk: Doğru doğruluk, doğru sınıflandırılan örneklerin toplam örnek sayısına oranıdır. Modelin genel başarısını gösterir ve aşağıda gösterildiği şekilde hesaplanır:

$$CA = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

F1-Score: F1 skoru, modelin hem doğruluk (precision) hem de duyarlılık (recall) dengesi üzerinde yoğunlaşan bir metriktir. F1, precision ve recall'un harmonik ortalamasıdır. Özellikle sınıflar arasındaki dengesizlik durumunda kullanılır. Aşağıdaki şekilde ifade edilmektedir:

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (2)$$

Precision (Kesinlik): Kesinlik, modelin pozitif tahminlerinden kaç tanesinin doğru olduğunu gösterir. Kesinlik, modelin pozitif olarak tahmin ettiği örneklerin pozitif olanlara oranıdır ve aşağıdaki şekilde hesaplanır:

$$\text{Precision} = \frac{FP}{FP+TN} \quad (3)$$

Recall (Duyarlılık): Duyarlılık, gerçek pozitiflerin kaç tanesinin doğru bir şekilde tespit edildiğini gösterir. Hassasiyet, gerçek pozitif örneklerin toplam pozitif örneklere oranıdır ve aşağıdaki şekilde hesaplanır:

$$\text{Recall} = \frac{TP}{TP+FN} \quad (4)$$

Bu formüller, sınıflandırma modelinin performansını ölçmek için kullanılır ve her metrik modelin farklı yönlerini değerlendirmeye yardımcı olur. **AUC**, ROC eğrisinin altında kalan alanı ölçerken, **CA** sınıflandırma doğruluğunu gösterir. **F1 skoru**, kesinlik (precision) ve duyarlılık (recall) arasındaki dengeyi sağlar. Kesinlik ve duyarlılık, modelin pozitif ve negatif tahminlerindeki doğruluğunu değerlendirir.

4. Bulgular

Elde edilen sonuçlar değerlendirildiğinde, Yapay Sinir Ağları ve Gradient Boosting modellerinin eğitim verisi üzerinde oldukça yüksek başarı gösterdiği görülmektedir. Yapay Sinir Ağlarının eğitim setindeki AUC değeri 0,974, doğruluk oranı 0,919 ve F1 skoru 0,932 iken Gradient Boosting için bu değerler sırasıyla 0,973, 0,913 ve 0,926 olarak hesaplanmıştır. Ancak test verisi üzerinde bu başarıların belirgin şekilde düştüğü gözlemlenmiştir. Yapay Sinir Ağlarının test AUC değeri 0,866'ya, doğruluk oranı 0,787'ye gerilemiştir. Benzer şekilde, Gradient Boosting'in test setindeki AUC değeri 0,871 olarak hesaplanmıştır. Bu durum, bu modellerin eğitim verisine fazla uyum sağladığını ve yeni verilere genelleme yapmada zorluk yaşadığını (overfitting) göstermektedir. Elde edilen sonuçlar Tablo 3'te gösterilmektedir.

Tablo 3. Algoritmaların Eğitim ve Test Aşamalarının Performans Ölçümleri

	Train					Test				
	AUC	CA	F1	Precision	Recall	AUC	CA	F1	Precision	Recall
SVM	0,927	0,843	0,867	0,856	0,877	0,885	0,796	0,836	0,799	0,877
Neural Network	0,974	0,919	0,932	0,917	0,946	0,866	0,787	0,826	0,803	0,851
Logistic Regression	0,927	0,846	0,869	0,862	0,876	0,890	0,793	0,832	0,805	0,862
k-NN	0,935	0,870	0,890	0,870	0,912	0,823	0,762	0,810	0,773	0,851
Gradient Boosting	0,973	0,913	0,926	0,925	0,926	0,871	0,784	0,823	0,806	0,840

Öte yandan, SVM ve lojistik regresyon modelleri hem eğitim hem de test verisi üzerinde daha dengeli sonuçlar vermiştir. Lojistik regresyonun test setindeki AUC değeri 0,890, doğruluk oranı 0,793 ve F1 skoru 0,832 olarak hesaplanmıştır. SVM modeli de test setinde AUC 0,885, doğruluk oranı 0,796 ve F1 skoru 0,836 ile istikrarlı bir performans sergilemiştir. Bu sonuçlar, lojistik regresyon ve SVM'nin aşırı öğrenmeye daha az eğilimli olduğunu ve daha iyi genelleme sağladığını göstermektedir.

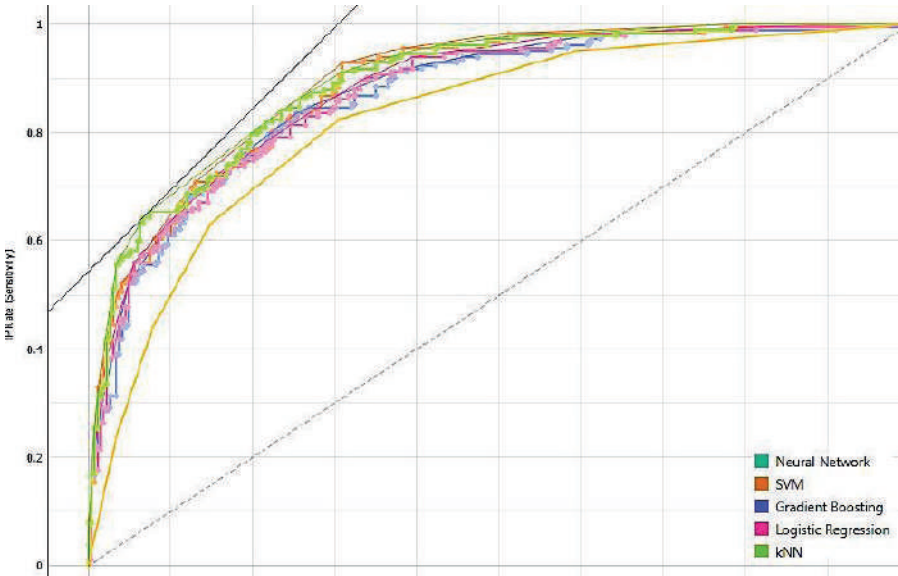
Yapay sinir ağları modeli, eğitim setinde AUC 0,974, doğruluk oranı 0,919 ve F1 skoru 0,932 gibi oldukça yüksek başarı göstermiştir. Kesinlik (0,917) ve duyarlılık (0,946) değerleri de oldukça güçlüdür. Ancak test setindeki başarısı belirgin şekilde düşmüştür (AUC 0,866, doğruluk 0,787, F1 skoru 0,826). Bu düşüş, modelin eğitim verisine fazla uyum sağladığını ve genelleme gücünün zayıfladığını göstermektedir. SVM modeli, eğitim verisi üzerinde AUC 0,927, doğruluk 0,843 ve F1 skoru 0,867 ile dengeli bir performans sergilemiştir. Test setinde de bu başarı büyük oranda korunmuştur (AUC 0,885, doğruluk 0,796 ve F1 skoru 0,836). Eğitim ve test skorları arasındaki farkın diğer modellere kıyasla daha az olması, SVM'nin genelleme yeteneğinin güçlü olduğunu göstermektedir. Lojistik regresyon modeli, eğitim setinde AUC 0,927, doğruluk 0,846 ve F1 skoru 0,869 ile tutarlı bir başarı sergilemiştir. Test setindeki sonuçları da benzer şekilde AUC 0,890, doğruluk 0,793 ve F1 skoru 0,832 olarak hesaplanmıştır. Modelin eğitim ve test skorları arasındaki farkın minimal olması, aşırı öğrenme problemine karşı dayanıklı olduğunu ve istikrarlı bir model sunduğunu göstermektedir.

kNN modeli eğitim setinde AUC 0,935, doğruluk 0,8695 ve F1 skoru 0,890 ile güçlü bir performans göstermiştir. Ancak test setinde başarısı belirgin şekilde düşmüştür (AUC 0,823, doğruluk 0,762 ve F1 skoru 0,810). Bu fark, kNN'nin aşırı öğrenme eğiliminde olduğunu ve genelleme yeteneğinin düşük olduğunu göstermektedir. Gradient Boosting, eğitim verisi üzerinde AUC 0,973, doğruluk 0,913 ve F1 skoru 0,926 gibi oldukça yüksek başarılar sergilemiştir. Ancak test setindeki değerlerin düşmesi (AUC 0,871, doğruluk 0,784 ve F1 skoru 0,823), modelin aşırı öğrenme eğiliminde olduğunu ortaya koymaktadır.

Model performansları genel olarak, lojistik regresyon ve SVM'nin en dengeli sonuçları verdiği görülmektedir. Lojistik regresyon modeli, test setinde en yüksek AUC değerine (0,890) sahip olup, doğruluk ve F1 skorları açısından da güçlüdür. SVM de benzer şekilde başarılı bir genelleme göstermektedir. Öte yandan, Yapay Sinir Ağları ve Gradient Boosting gibi karmaşık modeller eğitim setinde çok yüksek doğruluk sergilemelerine rağmen, test setinde bu başarıyı sürdürememiştir. Bu durum, bu modellerin

eğitim verisine aşırı uyum sağladığını ve yeni veriler üzerinde performans kaybı yaşadığını göstermektedir. kNN modeli ise test setinde en düşük doğruluk oranına sahip olup, genelleme gücü açısından diğer modellere kıyasla daha zayıf bir performans göstermiştir.

Aşağıdaki şekilde tüm değişkenlere ait ROC eğrisini göstermektedir. ROC eğrileri incelendiğinde, modellerin pozitif sınıflandırmayı ne kadar iyi ayırt edebildiği görsel olarak ortaya konmaktadır. Eğrilerin altında kalan alan (AUC) değerleri incelendiğinde, özellikle SVM ve Lojistik Regresyon modellerinin test verisinde 0,885 ve 0,890 seviyelerinde AUC elde etmesi, bu iki modelin diğerlerine göre daha iyi genelleme yaptığını göstermektedir. Diğer taraftan, Yapay Sinir Ağı ve Gradient Boosting eğitim verisinde oldukça yüksek AUC değerlerine ulaşmış olsa da test aşamasında daha düşük değerlere gerilemesi, potansiyel bir aşırı öğrenme belirtisidir. Dolayısıyla, ROC eğrilerinin genel şekli ve AUC değerleri birlikte değerlendirildiğinde hem doğru sınıflandırma oranının hem de yanlış pozitif oranının dengeli bir şekilde optimize edildiğini görmek, model seçimi ve hiperparametre iyileştirmesi açısından önemli ipuçları sunmaktadır.



Şekil 2. Algoritmaların ROC eğrileri

Sonuç olarak, en iyi genelleme yeteneğini gösteren modeller lojistik regresyon ve SVM olmuştur. Eğer aşırı öğrenme azaltılabilirse Yapay Sinir Ağları ve Gradient Boosting gibi güçlü algoritmalar daha başarılı sonuçlar

üretebilecektir. Ancak mevcut durumda Lojistik regresyon ve SVM, en istikrarlı ve güvenilir sonuçları sağlamaktadır.

5. Tartışma ve Sonuç

Bu çalışmada ele alınan çeşitli demografik, akademik, sosyolojik ve psikolojik faktörleri göz önünde bulundurarak makine öğrenmesi yöntemleri yardımıyla üniversite öğrencilerinin depresyonda olup olmadığının tahmini amaçlanmıştır. Kullanılan yöntemlerden Lojistik Regresyon ve Destek Vektör Makineleri (SVM) en iyi genellemeyi sağlamıştır. Yapay Sinir Ağları ve Gradient Boosting yöntemleri ise eğitim verilerinde yüksek doğruluk sergilemesine rağmen test verilerinde daha düşük performansla aşırı uyum eğilimleri göstermiştir. Bu bulgular, makine öğreniminin zihinsel sağlık bozuklukları için erken tespit ve müdahale stratejilerini geliştirebileceğini gösteren önceki araştırmaları desteklemektedir.

Elde edilen bulgular incelendiğinde akademik stres, uyku süresi, iş tatmini ve beslenme alışkanlıkları gibi değişkenlerin üniversite öğrencilerinde depresyon tahmininde kullanılabilecek önemli değişkenler olduğunu göstermektedir. Bu bulgular, akademik baskının ve yaşam tarzı ile ilgili faktörlerin öğrenci ruh sağlığı üzerindeki etkisini vurgulayan önceki araştırmalarla uyum göstermektedir (Andrews ve Wilding, 2004; Bayram ve Bilgel, 2008; Eisenberg vd., 2007). Özellikle akademik stres üniversite öğrencileri açısından depresyonu tetikleyen önemli bir risk faktörüdür. Yüksek sosyal beklentiler, başarısızlık korkusu ve iş yükü baskıları depresyonun tetikleyicilerindendir (Hunt ve Eisenberg, 2010; Arslan vd., 2009). Daha yüksek akademik baskıya maruz kalan öğrencilerin depresif semptomlar yaşama olasılığı daha yüksektir (Mirhosseini vd., 2021). Uykusuzluk da bilişsel işlevi, duygusal düzenlemeyi ve fiziksel sağlığı olumsuz etkileyerek öğrencileri depresyona daha yatkın hale getirir (Lund vd., 2010; Baglioni vd., 2011; Astafeva vd., 2024). Araştırmalar öğrenim görürken yarı zamanlı çalışan ve düşük iş tatmini olan öğrencilerin depresyon yaşama olasılığının daha yüksek olduğunu göstermektedir (Glozier, 1998; Assari, 2018; Satuf vd., 2018). Ayrıca beslenmenin ruh sağlığındaki üzerindeki rolü giderek daha fazla ilgi görmektedir. Bu da beslenme şekillerinin depresyon üzerinde etkili olabileceğini göstermektedir (Jacka vd., 2010; O'Neil vd., 2014; Tokpunar, 2024).

Bu çalışmanın sonuçları depresyon tahmininde makine öğrenmesi modellerinin kullanılabilme potansiyelini ve son otuz yılda ruh sağlığı değerlendirmesinde artan ilginin desteklendiğini göstermiştir (Baminiwatta, 2022). Bazı çalışmalar davranışsal, fizyolojik ve demografik verileri

birleştirirken, depresyonu ve diğer psikolojik bozuklukları yüksek doğrulukla tespit etmek için makine öğrenimi tekniklerini kullanmıştır (Karam vd., 2014; Liu vd., 2024).

Depresyon ile ilişkili biyolojik, psikolojik ve sosyal faktörlerin birbiri ile karmaşık etkileşiminden tahmin sonuçları etkilenmektedir (Hyde vd., 2008; Musliner vd., 2016). Bu karmaşık yapının bir temsilcisi olarak alınan bu çalışmada elde edilen bulgular tüm üniversite öğrencilerini tam olarak temsil etmeyebilir. Daha çeşitli veri kümeleri ile model performansını ve sonuçların genelleştirilebilirliği test edilebilir. Kesitsel bir veri yapısının kullanıldığı bu çalışmanın sonuçlarının iyileştirilmesi için panel veriler ile depresyon risk faktörlerinin zaman içinde nasıl değiştiği anlaşılabilir. Değişkenler açık erişimli belirli özellikleri içeren bir veri setidir. Bu veri setinde yer almayan ancak depresyon üzerinde etkin olabilecek, karmaşıklığı arttırabilecek çeşitli değişkenler yer alabilir. Bu değişkenlerin de araştırılarak ilerleyen çalışmalarda daha kapsamlı değişken seçimi yapılabilir.

Yapay Sinir Ağları ve Gradient Boosting gibi bazı modeller verilerle aşırı uyum sergileyerek düzenleme ve hiperparametre ayarlaması gibi daha fazla optimizasyon tekniğine ihtiyaç olduğunu göstermiştir. Depresyon tahmininde makine öğrenmesi kullanımı, veri gizliliği, modellerdeki olası önyargılar ve yanlış sınıflandırma riskiyle ilgili etik endişeler de mevcuttur. Gelecekteki araştırmalarda karar alma sürecinde şeffaflık ve hesap verebilirlik konularında iyileştirmeler yapıp farklı yapay zeka tekniklerini değerlendirme süreçlerine entegre ederek bu zorluklar aşılabılır.

Sonuçlar, makine öğrenmesinden yararlanarak ruh sağlığı profesyonelleri ve eğitimcilerin öğrencilerin refahını desteklemek için veri odaklı stratejiler geliştirebileceğini göstermiştir. Makine öğrenmesi teknikleri yapay zekâ alanındaki gelişmelerle birlikte gelişmeye devam ettikçe kişiselleştirilmiş müdahaleler ve ruhsal sorunların erken tespit edilmesi için büyük potansiyel içermektedir. Böylece depresyon gibi diğer psikolojik rahatsızlıkların da erken tanı ve tespitinin yanında şeffaf şekilde yönetilebileceği de anlaşılmıştır.

Kaynakça

- Ahmed, Z., Mohamed, K., Zeeshan, S., & Dong, X. (2020). Artificial intelligence with multi-functional machine learning platform development for better healthcare and precision medicine. *Database*, 2020, baaa010.
- Andrews, B., & Wilding, J. M. (2004). The relation of depression and anxiety to life-stress and achievement in students. *British journal of psychology*, 95(4), 509-521.
- Arastaman, G., Çetin, O. U., Arslan, S. Y., & Kerimoğlu, P. N. G. (2020). Eğitim yönetimi alanında doktora öğrencisi olmak: Lisansüstü öğrencilerin bakış açısından fenomenolojik bir inceleme. *Eğitimde Nitel Araştırmalar Dergisi*, 8(4), 1323-1346.
- Arslan, G., Ayranci, U., Unsal, A., & Arslantas, D. (2009). Prevalence of depression, its correlates among students, and its effect on health-related quality of life in a Turkish university. *Upsala journal of medical sciences*, 114(3), 170-177.
- Assari, S. (2018). Multiplicative effects of social and psychological risk factors on college students' suicidal behaviors. *Brain Sciences*, 8(5), 91.
- Astafeva, D., Konstantinov, D., Syunyakov, T., Kolsanov, A., Davydkin, I., Pavlichenko, A., ... & Smirnova, D. (2024). Depression and Anxiety Among COVID-19 Patients Admitted to a "Red Zone" Intensive Care Unit: Do the Factors of Self-Awareness of Somatic Health, Anxiety, and Depression Contribute to the Course of Sars-Cov-2 Infection?. *Psychiatria Danubina*, 36(suppl 2), 103-114.
- Baglioni, C., Battagliese, G., Feige, B., Spiegelhalder, K., Nissen, C., Voderholzer, U., ... & Riemann, D. (2011). Insomnia as a predictor of depression: a meta-analytic evaluation of longitudinal epidemiological studies. *Journal of affective disorders*, 135(1-3), 10-19.
- Baminiwatta, A. (2022). Global trends of machine learning applications in psychiatric research over 30 years: A bibliometric analysis. *Asian Journal of Psychiatry*, 69, 102986.
- Bayram, N., & Bilgel, N. (2008). The prevalence and socio-demographic correlations of depression, anxiety and stress among a group of university students. *Social psychiatry and psychiatric epidemiology*, 43, 667-672.
- Beiter, R., Nash, R., McCrady, M., Rhoades, D., Linscomb, M., Clarahan, M., & Sammut, S. (2015). The prevalence and correlates of depression, anxiety, and stress in a sample of college students. *Journal of affective disorders*, 173, 90-96.
- Ceyhan, E., & Ceyhan, A. (2011). Loneliness and Depression Levels of Students Using a University Counseling Center. *Eğitim ve Bilim-Education and Science*, 36(160).

- D'Ambrosio, E., Abrol, A., & Pigoni, A. (2023). Machine learning and psychosis: Diagnosis, prognosis and treatment. *Frontiers in Psychiatry*, 14, 1133072.
- Eisenberg, D., Gollust, S. E., Golberstein, E., & Hefner, J. L. (2007). Prevalence and correlates of depression, anxiety, and suicidality among university students. *American journal of orthopsychiatry*, 77(4), 534-542.
- Feng, L., Xu, S., Wang, F., Liu, S., & Qiao, H. (2019). Rough extreme learning machine: A new classification method based on uncertainty measure. *Neurocomputing*, 325, 269-282.
- Ferrari, C., Macis, A., Rossi, R., & Cameletti, M. (2018). Multivariate statistical techniques to manage multiple data in psychology. *Open Access Journal of Behavioural Science & Psychology*, 1(2).
- Gillooly, A., Dagnan, D., Hastings, R., Hatton, C., McMeekin, N., Baines, S., ... & Jahoda, A. (2024). Behavioural activation for depressive symptoms in adults with severe to profound intellectual disabilities: Modelling and initial feasibility study. *Journal of Applied Research in Intellectual Disabilities*, 37(2), e13197.
- Glozier, N. (1998). Workplace effects of the stigmatization of depression. *Journal of occupational and environmental medicine*, 40(9), 793-800.
- He, Y. M. (2024). Online Assessment of Mental Health Micromedia for College Students Incorporating Bayesian Network Algorithm. *International Journal of Maritime Engineering*, 1(1), 83-96.
- Hunt, J., & Eisenberg, D. (2010). Mental health problems and help-seeking behavior among college students. *Journal of adolescent health*, 46(1), 3-10.
- Hyde, J. S., Mezulis, A. H., & Abramson, L. Y. (2008). The ABCs of depression: integrating affective, biological, and cognitive models to explain the emergence of the gender difference in depression. *Psychological review*, 115(2), 291.
- Hysaj, A., Elkhoully, A., Qureshi, A. W., & Abdulaziz, N. (2018, December). Analysis of engineering students' academic satisfaction in a culturally diverse university. In *2018 IEEE international conference on teaching, assessment, and learning for engineering (TALE)* (pp. 755-760). IEEE.
- Ibrahim, A. K., Kelly, S. J., Adams, C. E., & Glazebrook, C. (2013). A systematic review of studies of depression prevalence in university students. *Journal of psychiatric research*, 47(3), 391-400.
- Jacka, F. N., Kremer, P. J., Leslie, E. R., Berk, M., Patton, G. C., Toumbourou, J. W., & Williams, J. W. (2010). Associations between diet quality and depressed mood in adolescents: results from the Australian Healthy Neighbourhoods Study. *Australian & New Zealand journal of psychiatry*, 44(5), 435-442.

- Karam, Z. N., Provost, E. M., Singh, S., Montgomery, J., Archer, C., Harrington, G., & Mcinnis, M. G. (2014, May). Ecologically valid long-term mood monitoring of individuals with bipolar disorder using speech. In *2014 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 4858-4862). IEEE.
- Liu, H., Li, C., Zhang, X., Zhang, F., Wang, W., Ma, F., ... & Zhang, X. (2024, March). Depression detection via capsule networks with contrastive learning. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 38, No. 20, pp. 22231-22239).
- Lund, H. G., Reider, B. D., Whiting, A. B., & Prichard, J. R. (2010). Sleep patterns and predictors of disturbed sleep in a large population of college students. *Journal of adolescent health, 46*(2), 124-132.
- Méndez, I., Martínez-Ramón, J. P., Ruiz-Esteban, C., & García-Fernández, J. M. (2020). Latent profiles of burnout, self-esteem and depressive symptomatology among teachers. *International Journal of Environmental Research and Public Health, 17*(18), 6760.
- Mirhosseini, S., Bazghaleh, M., Basirinezhad, M. H., Abbasi, A., & Ebrahimi, H. (2021). The relationship between depression and academic satisfaction in medical science students. *The Journal of Mental Health Training, Education and Practice, 16*(2), 99-111.
- Musliner, K. L., Munk-Olsen, T., Eaton, W. W., & Zandi, P. P. (2016). Heterogeneity in long-term trajectories of depressive symptoms: Patterns, predictors and outcomes. *Journal of affective disorders, 192*, 199-211.
- O'neil, A., Quirk, S. E., Housden, S., Brennan, S. L., Williams, L. J., Pasco, J. A., ... & Jacka, F. N. (2014). Relationship between diet and mental health in children and adolescents: a systematic review. *American journal of public health, 104*(10), e31-e42.
- Özenoğlu, A. (2018). Duygu durumu, besin ve beslenme ilişkisi. *Acıbadem Üniversitesi Sağlık Bilimleri Dergisi, 9*(4), 357-365.
- Pedrelli, P., Nyer, M., Yeung, A., Zulauf, C., & Wilens, T. (2015). College students: mental health problems and treatment considerations. *Academic psychiatry, 39*(5), 503-511.
- Piccinelli, M., & Wilkinson, G. (2000). Gender differences in depression: Critical review. *The British Journal of Psychiatry, 177*(6), 486-492.
- Ramón-Arbués, E., Gea-Caballero, V., Granada-López, J. M., Juárez-Vela, R., Pellicer-García, B., & Antón-Solanas, I. (2020). The prevalence of depression, anxiety and stress and their associated factors in college students. *International journal of environmental research and public health, 17*(19), 7001.

- Sakin Ozen, N., Ercan, I., Irgil, E., & Sigirli, D. (2010). Anxiety prevalence and affecting factors among university students. *Asia Pacific Journal of Public Health*, 22(1), 127-133.
- Satuf, C., Monteiro, S., Pereira, H., Esgalhado, G., Marina Afonso, R., & Loureiro, M. (2018). The protective effect of job satisfaction in health, happiness, well-being and self-esteem. *International journal of occupational safety and ergonomics*, 24(2), 181-189.
- Shamim, A. (2024, July 19). Student Depression Dataset. Kaggle. <https://www.kaggle.com/datasets/adilshamim8/student-depression-dataset>
- Shao, R., He, P., Ling, B., Tan, L., Xu, L., Hou, Y., ... & Yang, Y. (2020). Prevalence of depression and anxiety and correlations between depression, anxiety, family functioning, social support and coping styles among Chinese medical students. *BMC psychology*, 8, 1-19.
- Shatte, A. B., Hutchinson, D. M., & Teague, S. J. (2019). Machine learning in mental health: a scoping review of methods and applications. *Psychological medicine*, 49(9), 1426-1448.
- Thieme, A., Belgrave, D., & Doherty, G. (2020). Machine learning in mental health: A systematic review of the HCI literature to support the development of effective and implementable ML systems. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 27(5), 1-53.
- Tokpunar, M. (2024). Sürdürülebilir Diyet Modellerinin Depresyon ile İlişkisi: Derleme. *Genel Sağlık Bilimleri Dergisi*, 6(2), 346-363.
- Trigueros, R., Padilla, A., Aguilar-Parra, J. M., Lirola, M. J., García-Luengo, A. V., Rocamora-Pérez, P., & López-Liria, R. (2020). The influence of teachers on motivation and academic stress and their effect on the learning strategies of university students. *International journal of environmental research and public health*, 17(23), 9089.
- Villegas, C. A., & Martinez, M. J. (2022). Lessons from Harvey: Improving traditional damage estimates with social media sourced damage estimates. *Cities*, 121, 103500.
- Williamson, G., Trompeter, N., Murphy, D., Saba, S., Pedersen, E. R., Davis, J. P., & Leightley, D. (2024). Using passive and active data to predict posttraumatic stress disorder symptoms and cannabis use in recently discharged UK veterans: A protocol for the MAVERICK feasibility study. *Mental Health Science*, 2(3), e75.
- World Health Organization (WHO). (2021). Depression and other common mental disorders: Global health estimates. Retrieved from https://www.who.int/mental_health/management/depression/en/
- Zhang, Y., Peters, A., & Chen, G. (2018). Perceived stress mediates the associations between sleep quality and symptoms of anxiety and depression

among college nursing students. *International journal of nursing education scholarship*, 15(1), 20170020.

Tedarik Zinciri Yönetiminde Makine Öğrenmesi Uygulamaları

Ersin Şener¹

Pakize Yiğit²

Emrah Önder³

Özet

Tedarik zinciri verilerinden mevcut kalıpları keşfetmek her işletmede bir dönüm noktası olabilir. Yönetici müdahalesi veya analize rehberlik edecek taksonomiler olmadan Makine Öğrenimi (ML) algoritmaları ile tedarik zinciri verilerindeki yeni modeller günlük olarak bulunabilir. Algoritmalar kısıt tabanlı modelleme kullanarak verileri yinelemeli olarak sorgular ve en yüksek tahmin doğruluğuna sahip çözümü önerir. Özünde, Yapay Zeka (AI) ve makine öğrenimi algoritmaları, işletmelere envanter, tedarikçi kalitesi, potansiyel talep, tedarikten ödemeye, siparişten nakde, üretim planlaması, lojistik yönetimi ve daha fazlasını etkileyen Tedarik Zinciri Yönetimi (TZY) faktörlerini yönetme konusunda içgörüler sağlamaktadır. Üretilen yeni bilgi ve içgörüler, TZY’de verimli karar alma süreçleri ve işletmeler için karlı yatırım önerileri sunar.

Tüm bu faydalar işletmeleri hızla geliştirerek isabetli kararlar alınmasını, öngörülebilirliği ve yüksek yatırım ve kârlılık getirisini beraberinde getirebilir. Ticaretin doğası, yatırımın kârlılıkla geri dönüşüne bağlıdır. Bu açıdan bakıldığında bu bölüm, TZY’deki güncel ML uygulamaları hakkında bilgi sahibi olmak ve mevcut literatürü takip etmek isteyen okuyucular için bir çözüm sunmaktadır.

- 1 Dr. Öğr. Üyesi, Kırklareli Üniversitesi Fen Edebiyat Fakültesi Matematik Bölümü Uygulamalı Matematik Anabilim Dalı, ersinsener@klu.edu.tr, ORCID: 0000-0002-5934-3652
- 2 Doç. Dr., İstanbul Medipol Üniversitesi Tıp Fakültesi Temel Tıp Bilimleri Bölümü Biyoistatistik ve Tıp Bilişimi Anabilim Dalı, pyigit@medipol.edu.tr, ORCID: 0000-0002-5919-1986
- 3 Prof. Dr., İstanbul Üniversitesi İşletme Fakültesi İşletme Bölümü Sayısal Yöntemler Anabilim Dalı, emrah@istanbul.edu.tr, ORCID: 0000-0002-0554-1290

1. Giriş

Ürünlerin sadece bir tık uzakta olduğu bir dünyada, Tedarik Zinciri Yönetimi (TZY) yöneticilere en son teknolojileri kullanarak verimli kararlar alma yetkisi verir. Bu küreselleşmiş pazar, doğal olarak alıcıları daha uygun fiyatlı seçenekler aramaya, daha hızlı tedarik ve hizmet ile güvenilir servis sonrası ağları talep etmeye yönlendirmektedir. Şirketler küresel pazarda rekabet edebilmek için doğru kararları hızlı ve verimli bir şekilde almalıdırlar. Tedarik zinciri yönetimlerini otomatikleştirmek, şirketlerin küresel oyuncularla rekabet edebilmelerini sağlar. TZY, günümüz iş operasyonlarının önemli bir bileşenidir ve ana tedarikçilerden son kullanıcılara kadar çeşitli tedarik zinciri faaliyetlerini sorunsuz bir şekilde entegre etmeyi sağlar (Lambert vd., 1998). TZY'nin temel amacı, müşteri değerini en üst düzeye çıkaran ve sürdürülebilir rekabet avantajı sağlayan yüksek kaliteli hizmetler, ürünler ve bilgiler sağlamaktır. Etkili TZY stratejileri, kuruluşlara operasyonel verimliliğini optimize etmeyi, maliyetlerini düşürmeyi ve müşteri memnuniyetini artırma gibi olanaklar sunmaktadır. Bu nedenle, TZY, giderek daha karmaşık ve dinamik hale gelen bir pazar ortamında rekabetçi kalmaya çalışan çeşitli sektörlerdeki işletmeler için kilit bir odak alanı olarak ortaya çıkmıştır (Deveci, 2023; Tirkolace vd., 2021; Wenzel vd., 2019; Younis vd., 2021).

TZY beş temel alanı kapsar: hammadde seçimi ve tedarigi, üretim, dağıtım (son kullanıcıya), satış ve hizmet. Müşterinin geri bildirimi, tedarik zinciri sürecinin sürekli olarak iyileştirilmesi için çok önemlidir. Müşteri geri bildirimleri, mevcut ürünlerin yeni versiyonlarının tasarlanmasına ve hizmet ağı üzerindeki iş yükünün potansiyel olarak en aza indirilmesine yol açabilir. Bu açıdan üretim, satış, geri bildirim ve yeniden üretim süreci bir zinciri temsil eder ve dolayısıyla TZY kavramını oluşturur. TZY, her bir süreçte kaydedilen verilerin analiz edilmesiyle mümkündür. Bununla birlikte, son yıllarda üretilen verilerin hacminin büyüklüğü, işletmelerin anlamlı içgörüler elde etmesini giderek zorlaştırmaktadır. TZY'de büyük veriyi yönetmek ve analiz etmek önemli zorluklar doğurmaktadır. Bu sorunu çözmek için birçok şirket, geleneksel matematiksel ve istatistiksel yöntemler kullanmak yerine yazılım endüstrisinde kendine yer bulan Makine öğrenimini (ML), tercih etmektedir.

ML kullanımı, büyük miktarda veriyi işleme kapasitesi nedeniyle TZY'ndeki araştırmacıların ilgisini çekmiştir. Geleneksel yaklaşımlar, büyük ölçekte yapılandırılmamış verileri ve doğrusal olmayan problemleri etkili bir şekilde analiz edememektedir. Ancak ML teknikleri bu zorlukları çözebilmekte ve tedarik zinciri performansını etkileyen faktörleri tahmin

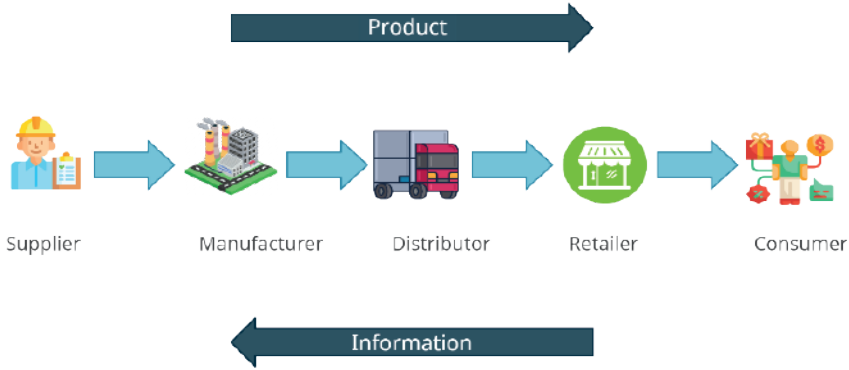
edebilmektedir. ML, TZY’de büyük verileri analiz etmek isteyen işletmeler için paha biçilmez bir araçtır (Tirkolace vd., 2021).

Etkili Tedarik Zinciri Yönetimi, modern iş dünyasında rekabet avantajı elde etmek için çok önemlidir. En son teknoloji ve gelişmiş veri analitiği sayesinde işletmeler, iş süreçlerini kolaylaştırmak ve karar verme yeteneklerini geliştirmek için yeni fırsatlar yakalayabilir. Bu bölüm, şirketler için TZY’de ML yöntemlerinin potansiyel uygulama alanlarını ve gelecekteki bilimsel keşifler için potansiyel araştırma alanlarını belirlemeye yardımcı olmayı amaçlamaktadır. Aşağıdaki bölümde, TZY’deki çeşitli makine öğrenmesi uygulamalarına ilişkin bir literatür taraması sunulmuştur. İş operasyonlarının optimize edilmesindeki önemi de tartışılmaktadır. İkinci bölüm, veri analitiği ve makine öğreniminin büyüleyici dünyasını incelemektedir. Bu bölümde özellikle TZY’de en çok kullanılan ML algoritmalarından bazıları özetlenmiştir. Üçüncü bölümde, okuyuculara bu algoritmaların endüstriyel elektrik tüketimine ilişkin bir veri setine aydınlatıcı bir uygulaması sunulmaktadır. Bu veri seti, Endüstri 4.0’da teknolojinin artan rolü için ilgi çekici bir vaka çalışması sunmaktadır. Bölüm, çeşitli ML algoritmalarının performansının anlaşılır bir şekilde karşılaştırılması ve TZY’ye katkılarının yorumlanmasıyla sona ermektedir. Son olarak, bu heyecan verici alanda makine öğrenmesinin gelecekteki uygulamaları için öneriler sunulmuştur.

1.1. Literatür Taraması

Günümüz iş dünyasında şirketler, hammadde tedarikinden nihai ürünlerin müşterilere dağıtımına kadar tedarik zincirleri boyunca büyük miktarda veri üretmektedir. İşletmelerin müşteri memnuniyeti ve yüksek kâr elde edebilmeleri için bu verilerden değerli içgörüler elde etmeleri çok önemlidir. Endüstri 4.0’da işletmeler için temel zorluk veri toplamak değil, bu bilgilerden değerli içgörüler elde etmek ve bunları etkili bir şekilde uygulamaktır. Endüstri 4.0’ın sunduğu teknolojik gelişmeler, TZY ortamını hızla dönüştürmektedir (Şahin, 2023).

TZY süreçlerinin görsel bir haritası Şekil 1.1’de sunulmuştur. Şekilde üretim süreçleri soldan sağa doğru gösterilmiş ve gelen bilgilerle üretimde kaliteyi artırmak için sağdan sola doğru bir zincir oluşturulmuştur.



Şekil 1.1 Tedarik Zinciri Yönetim Süreci (Supply Chain Diagram | Supply Chain Flowcharts | Creately, 2021)

Geçtiğimiz birkaç yıl içinde makine öğrenmesi, teknolojik gelişmeler ve bilgi işlem gücü sayesinde giderek daha popüler hale geldi. Tedarik zinciri yönetiminde ML ve AI kullanan birçok çalışma bulunmaktadır. Bunlardan ML kullanan bazıları aşağıda ele alınmıştır.

Talep tahmini yönetimi her işletmenin önemli bir unsurudur. Gelecekteki talebi doğru bir şekilde tahmin etmek, kuruluşların envanterlerini optimize etmelerine, tedarik zincirlerini yönetmelerine ve üretim programlarını planlamalarına yardımcı olabilir. Etkili talep yönetimi, işletmelerin arz ve talebi dengelemesine, stoksuzlukları en aza indirmesine ve tedarik sürelerini azaltmasına yardımcı olabilir. ML, TZY’de talep tahmin süreçlerini iyileştirmek için kullanılmıştır (İfraz vd., 2023). Geçmiş satış verileri, mevsimsel değişimler ve pazar eğilimleri analiz edilerek daha doğru talep tahminleri yapılabilir. Bu, envanter yönetimini optimize etmek ve stok seviyelerini en aza indirmek için önemli bir avantaj sağlamıştır.

Talep tahminine yönelik bir ML uygulamasında, Kanada Dökümhanelerinden alınan SCM verileri ve siparişleri simüle edilerek farklı tahmin teknikleri denenmiştir. Çalışma, boğa güreşi etkisi (*bullwhip effect*) olarak da bilinen çarpık talebi tahmin etmek için naif (*naïve*) kestirim, trend analizi, hareketli ortalamalar, doğrusal regresyon, tekrarlayan sinir ağları (*recurrent neural networks-RNN*) ve destek vektör makineleri gibi yöntemlerin etkinliğini karşılaştırmıştır. Sonuçlar, RNN ve destek vektör makinesi tekniklerinin regresyon yaklaşımıyla karşılaştırılabilir sonuçlar sunduğunu göstermektedir (Carbonneau vd., 2008).

Tedarik Zinciri Risk Yönetimi (TZRY) işletmeler için çok önemli bir sorundur. Aksaklıkları önceden tahmin etmek ve önlemek için AI destekli

ML yöntemlerinin kullanılması kapsamlı bir şekilde araştırılmıştır. Bununla birlikte, genellikle tahminlerin anlaşılabilirliği göz ardı edilerek tahmin doğruluğu üzerinde yoğunlaşmıştır (Ye vd., 2015; Zage vd., 2013). Bu durumu ele almak için hem tahmin doğruluğunu hem de anlaşılabilirliği dengeleyen veri odaklı bir yapay zekâ çerçevesi önerilmiştir. Bu yaklaşım, sevkiyat gecikmelerini tahmin etmek için gerçek tedarik zinciri verileri üzerinde test edilmiştir. Sonuçlar, doğruluk yerine anlaşılabilirliğe öncelik vermenin bazı tavizler vermeyi gerektirebileceğini ortaya koymaktadır (Baryannis vd., 2019).

Tedarik zinciri yönetiminde makine öğrenmesi uygulamaları beş temel alana dayanmaktadır: hammadde seçimi ve tedarigi, üretim, dağıtım (tüketiciye teslimat dahil), satış ve hizmet. Bu beş ana alanda makine öğrenmesini kullanan birçok çalışma bulunmaktadır. Bunlardan bazıları aşağıdaki gibi özetlenebilir. TZY’de tedarikçi değerlendirme ve seçim problemleri oldukça önemlidir. Tedarikçiler, perakendeci için nitelikli ve uygun olan doğru ürün veya hizmetin uygun miktarını doğru zamanda ve yerde teslim edebilir. Bu alanda veri bulmak genellikle zordur, bu nedenle küçük örneklem boyutları ve eksik veri gibi sorunlar yaygındır (Ni vd., 2020a). ML algoritmaları bu tür verilerle başa çıkmada başarılıdır. Örneğin, Lee ve arkadaşları (Lee vd., 2022) tedarik seçimi için takviyeli öğrenmeyi kullanmıştır. Endüstri 4.0’ın devreye girmesiyle birlikte ML teknikleri, veri kullanılabilirliği, yüksek işlem gücü ve geniş depolama kapasitesi nedeniyle üretim planlama ve kontrolünü ele almak için cazip bir yol haline gelmiştir (Usuga Cadavid vd., 2020).

Üretim planlaması için bulanık mantık ve karar ağaçları gerçek hayat uygulamalarında tatmin edici sonuçlar elde etmiştir (González Rodríguez vd., 2020). Ayrıca Garre ve diğerleri (Garre vd., 2020) makine öğrenmesi yöntemlerinin gıda endüstrisindeki üretim anormalliklerini geleneksel modellere göre daha iyi tespit edebildiğini göstermiştir. Tsolaki (Tsolaki vd., 2023) yük taşımacılığı, SC ve lojistik alanlarını kullanarak tüm makine öğrenmesi ve hibrit yöntemleri gözden geçirmiştir. Tüm makine öğrenmesi ve hibrit yöntemlerin bu alanda bilgi keşfini, planlamayı ve karar vermeyi desteklediğini belirtmişlerdir.

Talep ve satış tahmini çok değişkenli doğrusal olmayan ilişkilere sahiptir, bu nedenle ML tabanlı yöntemler geleneksel yöntemlerden daha iyi performans göstermiştir. Feizabadi (Feizabadi, 2022) talep tahmini için ARIMAX ve YSA modellerini karşılaştırmıştır ve YSA tahmin doğruluğu ARIMAX’tan daha yüksek bulunmuştur. Punia ve Shankar (Punia & Shankar, 2022) talep tahmini için hem geleneksel (ARIMA, ARIMAX), hem ML (RF,

ANN, LSTM) hem de hibrit ML modellerini kullanmış ve ML ve hibrit ML modellerinin geleneksel modellerden daha iyi performans ölçütlerine sahip olduğunu bulmuştur. Satış sonrası hizmetlerini iyileştirmek ve daha verimli TZY sağlamak için müşteri geri bildirimlerine ihtiyaç duyulmaktadır. Anketler, görüşmeler ve pazar araştırmaları yaygın olarak kullanılmaktadır, ancak bu yöntemlerin daha fazla sayıda örnekleme ulaşması gerekmektedir. Sosyal medya verileri, müşterilerin gerçek görüşlerine erişebilmek için son zamanlarda bu alanda kullanılmaya başlanmıştır. ML algoritmalarını Nesnelerin İnterneti (IoT) ile entegre etmek, büyük verileri işlemek için kullanılmıştır (Ni vd., 2020b). Khatua ve diğerleri (Khatua vd., 2021) de AI ile sosyal medya analizinin ürün performansı, marka itibarı, müşteri segmentasyonu, talep tahmini ve hizmet değerlendirmesi gibi belirli TZY alanlarının verimliliğini artırdığını belirtmiştir.

Endüstri 4.0'ın ortaya çıkmasıyla birlikte dijitalleşme (IoT), blok zinciri, büyük veri, bulut bilişim ve robotik gibi kavramlar da geliştirilmiştir. İşletmeler son zamanlarda TZY faaliyetlerini geliştirmek için bu yeni teknolojileri uygulamaya çalışmaktadır. Bu nedenle, büyük miktarda etkileşimli veriyi hızlı ve akıllı bir şekilde yorumlamak için büyük veri, AI ve ML yöntemlerini kullanılmıştır (Nagar vd., 2021; Tirkolace vd., 2021). Örneğin, araştırmacılar ML yöntemlerini kullanarak TZY'deki IoT ve standart uygulamalar arasındaki farkı araştırmışlardır (Li vd., 2022; Rizwan vd., 2022). Yazarlar, TZY'deki IoT uygulamalarının standart olanlardan daha iyi çalıştığını; standart olanlara kıyasla verimliliği ve kârı artırdığını bulmuşlardır. Öte yandan Cadavid ve arkadaşları (Usuga Cadavid vd., 2020), Endüstri 4.0 döneminde üretim, planlama ve kontrol problemleri için ML uygulamalarını incelemiştir. Üretim sistemindeki değişiklikleri hesaba katmak için makine öğrenmesi modelini güncelleme zorluğu ve veri toplamak için IoT teknolojilerini kullanmanın karmaşıklığı nedeniyle araştırmanın alanla ilgili bilgileri güncellemek için yetersiz olduğunu bildirmişlerdir.

TZY, kavramı ilk kez kullanıldığı 1980'li yıllardan bu yana yöneylem araştırmasında önemli bir konu olmuştur (Oliver & Webber, 1982; Pourhejazy & Kwon, 2016). Optimizasyon modelleri bu alanda en yaygın kullanılan yöneylem araştırması araçlarıdır (Barbosa-Póvoa vd., 2018). Bulanık teori ve esnek hesaplama, TZY optimizasyon problemlerini çözmek için çoğunlukla makine öğrenmesi yöntemlerini kullanmaktadır (Ni vd., 2020a). Abbasi ve arkadaşları (Abbasi vd., 2020) kan tedarik zinciri yönetimi optimizasyon problemini YSA, CART, K-NN ve RF kullanarak çözmüştür. Lin ve arkadaşları (Lin vd., 2022), performans optimizasyon problemini çözmek için yenilikçi bir makine öğrenmesi yöntemi kullanmıştır.

Tedarik Zinciri Yönetimi'nde makine öğrenmesini uygulamanın en önemli avantajlarından biri, işletmelerin verilerinden değerli çıkarımlar yapmalarına olanak sağlamasıdır. İkinci olarak, yöneticilere gerçek zamanlı içgörüler sağlayarak potansiyel riskleri en aza indirmesi sayılabilir. Son olarak da kapsamlı verilerin verimli bir şekilde analiz edilmesini sağlayarak tedarik zinciri hakkında tutarlı ve doğru bilgiler sağlamasıdır.

2. Veri Analitiği ve Türleri

Endüstri 4.0 dönemi, makinelerden ve internete bağlı ağ sensörlerinden gelen bol miktarda veri ile karakterize edilmektedir. Veri analitiği ve makine öğrenmesi, bu verileri anlamlandırmak ve makine zekasına ulaşmak için gerekli hale gelmiştir. Bu güçlü araçlar, izleme görevlerinden kritik karar alma süreçlerine kadar sürecin her yönüne nüfuz ederek geleneksel operasyonlarda ve TZY'de devrim yaratmıştır. Analitik sistemler, operasyonel verimliliği artıran, hataları ve kusurları en aza indiren, maliyetleri düşüren, kaliteyi geliştiren ve üretkenliği artıran sürdürülebilir karar verme süreçleri geliştirmek için bu verileri işleyerek entegre edilebilmektedir. Veri, akıllı ve sürdürülebilir üretim ve tedarik zinciri operasyonlarında girdi aynı zamanda da çıktıdır. Karar verme için hayati bir girdiyi ve gerçek zamanlı geri bildirim için temel bir çıktıyı temsil eder. Nihai hedefler geliri en üst düzeye çıkarmak ve maliyetleri en aza indirmekken, veri analitiği bu verileri bu hedeflere ulaşmaya yardımcı olabilecek anlamlı içgörülere dönüştürmek için hayati önem taşır.

Makine öğrenmesi yöntemleri uygulanarak çeşitli veri analizi aşamaları gerçekleştirilebilir. Bu aşamalar tahmine dayalı, tanısal, kuralcı ve açıklayıcı analitikleri içermektedir. Tahmine dayalı analitikler, potansiyel üretim çıktısını ve makinelerin kalan ömrünü tahmin etmek için istatistiksel modeller kullanır. Teşhis analitiği ise hem makine hem de ürünlerdeki arızaları tespit eder ve raporlar. Kural koyucu analitik, üretimi optimize etmek ve belirli eylemlerin etkisini belirlemek için öneriler sunar. Tanımlayıcı analitik, üretim sürecini ve ortamını çevreleyen koşulları incelemekte ve tanımlamaktadır (Kotsiopoulos vd., 2021).

2.1. Makine Öğrenmesi

Sayısal veriler ya da sayılarla ifade edilebilen gözlemler içeren her problemin bir çözüme ihtiyacı vardır. Günümüz koşullarında bu problemlerin klasik yöntemlerle manuel olarak çözülmesi mümkün olmadığından bilgisayarlar bu rolü üstlenmektedir. Bilgisayarların karmaşık problemlere algoritmalarla çözüm sunduğu bilinmektedir. Algoritmalar bilgisayarların çözüm elde etmesini ve konu hakkında bilgi edinmesini sağlar. ML, karmaşık sayısal

değerlere sahip problemleri makine diline çevirerek algoritmalar aracılığıyla çözümler.

Basit bir ifadeyle makine öğrenmesi, verilerden öğrenmek için bilgisayar sistemlerini kullanan bir teknolojidir. Bu teknolojinin temel amacı, TZY’de veri analizi, tahmin ve karar verme süreçlerini optimize etmektir. Makine öğrenmesinin temel ilkelerine aşina olmak, bu teknolojinin pratikte uygulanmasını sağlamak için çok önemlidir. ML, büyük miktarda veriyi analiz ederek kalıpları ve ilişkileri tanımlayabilmektedir. Bu yetenek, tedarikçi performansının değerlendirilmesi, talep tahminlerinin güncellenmesi ve envanter yönetiminin optimize edilmesi gibi TZY’nin birçok alanında kullanılabilir. Doğru ve güvenilir veri toplama süreçleri, ML uygulamalarının başarılı olmasını sağlamaktadır.

ML’in sınıflandırma, regresyon, kümeleme ve hibrit görevler dahil olmak üzere birçok kullanım alanı bulunmaktadır. Sınıflandırma, girdilerin önceden belirlenmiş sınıflara atanmasını içerirken, regresyon çeşitli özelliklere dayalı bir çıktı tahmin etmektedir. Temel fark, regresyonun sürekli bir çıktı değeri üretmesi, sınıflandırmanın ise ayrık bir çıktı değeri sunmasıdır. Kümeleme ise bir grup nesnenin daha küçük, benzer gruplara bölünmesini içermektedir. Hibrit görevler, her kategoriye sürekli değer biçiminde kategorik değerler atamak gibi sınıflandırma, regresyon veya kümelemeyi bir araya getirmektedir (Jo, 2021).

2.1.1. Makine Öğrenmesi Türleri

Sıklıkla kullanılan dört makine öğrenimi yöntemi vardır: denetimli öğrenme, denetimsiz öğrenme, yarı denetimli öğrenme ve pekiştirmeli öğrenme.

2.1.1.1. Gözetimli Öğrenme

Denetimli öğrenme girdi-çıkı verilerinden tahmin yapar. Bir öğretmenden öğrenildiğini varsayar. Modelin başarısı, eğitim aşamasında oluşturulan modelin yeni test verilerindeki çıktıyı ne kadar iyi tahmin ettiği ile ölçülür. Karar Ağaçları (DT), Naïve Bayes (NB), Destek Vektör Makinesi (DVM), k-en yakın komşu (KNN) ve denetimli yapay sinir ağları (ANN), TZY çalışmalarında sıklıkla kullanılan algoritmalarıdır (Tirkolae vd., 2021).

2.1.1.2. Denetimsiz Öğrenme

Denetimsiz öğrenme “öğretmensiz öğrenme” olarak da adlandırılmaktadır. Öğrenme sürecinde sadece girdi verileri vardır, çıktı verileri yoktur. Öğrenme yöntemi, çıktı verisi olmadan verinin karakteristik yapısını bulmayı

amaçlamaktadır. Denetimli öğrenmenin aksine, bu öğrenme yaklaşımlarında kesin bir başarı ölçütü yoktur. Ancak araştırmacılar en az farklılık gösteren ya da en çok benzeyen örüntüleri belirlemeye çalışmaktadır. K-ortalamlar kümeleme (k-means clustering), temel bileşen analizi (PCA) ve denetimsiz sinir ağı yöntemleri bu alanda önde gelen algoritmalar (Tirkolace vd., 2021).

2.1.1.3. Yarı Gözetimli Öğrenme

Denetimli öğrenme sadece verilerin etiketlenmesini gerektirir ki bu da uygulamalarda her zaman mümkün değildir. Yarı denetimli yöntemler, denetimli ve denetimsiz öğrenmeyi birleştirerek bu sorunu çözmektedir (Sammur & Webb, 2017). Kendi kendine eğitim ve transdüktif SVM yöntemleri, TZY uygulamalarında yaygın olarak kullanılmaktadır (Tirkolace vd., 2021).

2.1.1.4. Pekiştirmeli Öğrenme

Pekiştirmeli öğrenme, zaman içinde ödülü en üst düzeye çıkaracak şekilde nasıl hareket edileceğini bularak karmaşık sistemler için uyarlanabilir bir kontrolör olarak çalışabilir. Denetleyicinin (takviyeli öğrenme ajanı olarak da adlandırılır) amacı, uzun vadeli hedefleri en üst düzeye çıkarmak için dinamik sistemin yaşayabileceği her durumda en uygun kontrol eylemlerini öğrenmektir (Boute vd., 2022; Rolf vd., 2023). Klasik Q öğrenme yöntemleri, TZY problemlerinde en çok kullanılan algoritmalar (Rolf vd., 2023).

2.1.2. Makine Öğrenmesi Algoritmaları

Bu bölümde literatürde en çok kullanılan denetimli ve denetimsiz öğrenme algoritmaları özetlenmiştir. Regresyon, lojistik regresyon (LR), KNN, NB, DT, rastgele orman (RF), SVM, denetimli ANN ve Xgboost denetimli öğrenme algoritmalarıdır. K-means, hiyerarşik kümeleme, gürültülü uygulamaların yoğunluk tabanlı uzamsal kümelmesi (DBSCAN), PCA ve birliktelik kuralları ise denetimsiz öğrenme algoritmalarıdır.

1. **Regresyon** : Regresyon modellerinin iki türü vardır: doğrusal ve doğrusal olmayan. Doğrusal bir fonksiyona uyan girdi değişkenlerini kullanarak bir çıktı değişkenini modellemek mümkündür; bu süreç doğrusal regresyon olarak bilinir ve hesaplanması oldukça basittir. Bir model doğrusal olmayan bir fonksiyona uyduğunda doğrusal olmayan regresyon gerçekleştirmek zordur (Haykin, 2009).

2. **Lojistik regresyon (LR):** LR, sınıflandırma problemleri için kullanılır ve doğrusal regresyon tekniğini uyarlar (Hastie vd., 2008). Modelde cevap değişkeni kategoriktir ve sonuç için bir değer tahmin etmek yerine cevap değişkenlerinin kategorileri için olasılığı tahmin etmeye çalışır (Kantardzic, 2020). Girdi veri setleri hem nicel hem de kategorik olabilir.
3. **K-En Yakın Komşuluk (KNN):** KNN, regresyon ve sınıflandırma problemleri için kullanılır. Algoritma, yakınlıkta bulunan kümelenmiş veri setlerini bularak sınıflandırma ve regresyon problemlerini çözmeyi amaçlar. Bu nedenle eğitim seti K adet en yakın örneğe bölünür. Doğrusal ve doğrusal olmayan problemleri kullanan etkili ve uygulaması kolay bir algoritmadır, ancak performansı kullanıcının uygun K'yı seçmesine bağlıdır (Shyam vd., 2023).
4. **Naive Bayes (NB):** NB, Bayes teoremine dayanan sağlam bir sınıflandırma algoritmasıdır. Eğitim verilerinden sınıf üyelerinin olasılığı ve her bağımsız gösterge göz önüne alındığında sınıf üyeleri için koşullu olasılık olmak üzere iki tür olasılık hesaplanabilir. Her bir girdi göstergesinin bağımsız olduğunu varsaydığı için "naive" olarak adlandırılır. Hayati bir varsayım olmasına ve gerçek veriler için pratik olmamasına rağmen, yöntem çeşitli zorlu konularda inanılmaz derecede iyi sonuç vermektedir (Shobha ve Rangaswamy, 2018).
5. **Karar ağacı (DT):** DT'ler regresyon ve sınıflandırma kullanan etkili parametrik olmayan algoritmalarıdır. Test fonksiyonlarına sahip karar düğümlerini kullanan bir dizi özyinelemeli bölünmenin yerel bölgeyi tanımladığı hiyerarşik bir modeldir. Tipik olarak, arama uzayının belirli bir alanında bir çözüm bulmak için karar ağacı öğrenme sistemleri tarafından yukarıdan aşağıya bir yaklaşım kullanılır. Bu, basit bir ağacın bulunmasını sağlar; ancak bu en basit ağaç olmayabilir. Bir karar ağacındaki düğümler, nitelikleri test etmek için kullanılır (Kantardzic, 2020).
6. **Rastgele Orman (RF):** RF, regresyon ve sınıflandırma görevleri için topluluk karar ağacı tabanlı bir algoritmadır. Çok sayıda karar ağacından oluşur ve sonuçları oluşturmak için tahminlerini toplar. Eğitim sürecinde, birçok karar ağacı hem verilerin hem de özelliklerin rastgele seçimleri üzerinde eğitilir. Ormanda, her ağaç belirli bir girdi için bir tahmin üretir ve nihai tahmin, tüm ağaçların bireysel tahminlerinin ortalama (regresyonda) veya oylama (sınıflandırmada) yoluyla toplanmasıyla belirlenir. Aşırı uyumu azaltarak daha doğru bir tahmin elde edilmesini sağlar (Shyam vd., 2023).

7. **Destek Vektör Makinesi (SVM):** SVM, regresyon ve sınıflandırma problemleri için güvenilir ve iyi bilinen bir tekniktir. Marjı optimize ederken veri noktalarını farklı sınıflara verimli bir şekilde bölen bir hiper düzlem (yüksek boyutlu bir uzayda bir çizgi veya düzlem) bulmak SVM'lerin arkasındaki ana fikirdir. Hiper düzleme en yakın veri noktaları destek vektörleri olarak adlandırılır ve hiper düzlemin konumlandırılması üzerinde en önemli etkiye sahiptirler (Shyam vd., 2023).
8. **Yapay Sinir Ağları (ANN):** ANN'ler, beyin ve sinir sisteminin nasıl çalıştığını tahmin etmek için en etkili yöntemler arasındadır. Bu algoritmalar parametrik değildir ve derin öğrenme metodolojisi ile tüm makine öğrenimi yöntemlerini (denetimli, denetimsiz, yarı denetimli ve takviyeli öğrenme yöntemleri) kullanabilir. Sonuç olarak, gerçek dünya senaryolarında doğru tahminler üretebilir.
9. **Aşırı Gradyan Güçlendirme (Xgbost):** XGBoost, hem sınıflandırma hem de regresyon görevleri için karar ağaçlarının ilkelerini kullanan bir topluluk yöntemidir. Birkaç karar ağacını eğitmek için gradyan artırma metodolojisini kullanır. Random Forest, eğitim setinden rastgele alt kümeler seçerek karar ağaçları oluştururken, XGBoost daha önce oluşturulmuş ağaçların performansını göz önünde bulundurarak karar ağaçlarını sırayla oluşturur.
10. **K-Means:** K-Means algoritması, basitliği ve gücü nedeniyle popüler denetimsiz algoritmalarından biridir. Bir çeşit bölümlendirme kümeleme tekniğidir. Algoritma, her grubun merkezini (ortalamasını) kullanarak verileri kullanıcı tarafından belirtildiği gibi K kümeye ayırır. Daha sonra belirli koşullar sağlanana kadar gözlemleri kümelere yinelenmeli olarak atar. Kriterlerin birincil amacı, farklı kümeler arasındaki uzamsal ayrımı en üst düzeye çıkarırken bir küme içindeki gözlemler arasındaki uzamsal ayrımı azaltmaktır. Uzaklık hesaplamasında genellikle Öklid uzaklığı kullanılır (Hair, 2019).
11. **Hiyerarşik Kümeleme:** Hiyerarşik kümeleme, sabit k kümesi olmayan başka bir uygulanabilir ve pratik algoritmadır; grup sayısı iterasyonlarla değişir. Yukarıdan aşağıya bir yaklaşım kullanarak kümeleri dendrogramlarla görselleştirir. Başlangıçta, her örnek aynı kümenin bir parçasıdır; iterasyonlardan sonra, her örnek kendi kümesine aittir.
12. **DBSCAN:** Bölümlendirme ve hiyerarşik yaklaşımlar küresel şekle sahip kümeleri bulmayı amaçlar. Oval ve "S-şekilli" kümeler gibi keyfi

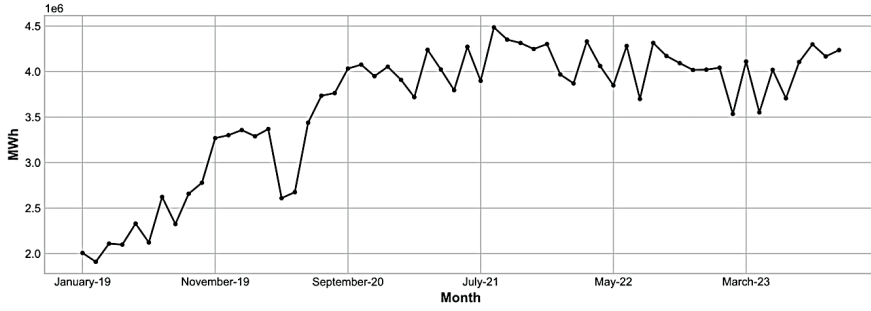
şekilli grupları bulmakta zorlanırlar. Kümelerdeki aykırı değerleri bulmada iyi değillerdir, bu nedenle yoğunluk tabanlı yöntemler kullanılır. DBSCAN çekirdek nesneleri (nesnenin yakınındaki birkaç nesne tarafından hesaplanan bir nesnenin yoğunluğunun durumu) tanımlamıştır; bunlar yoğun komşulukları olan nesnelerdir (Han vd., 2012).

13. Temel Bileşenler Analizi (PCA): PCA öncelikle boyut azaltma için kullanılır. Verileri temsil etmek için en uygun k adet n boyutlu ortogonal vektör bulur ($k < n$). Orijinal verilerin önemli ölçüde daha küçük bir uzaya yansıtılması nedeniyle boyutluluk azaltılır. PCA, sıralı ve sırasız göstergeler için idealdir ve seyrek ve çarpık verileri çözebilir (Han vd., 2012).

14. Birliktelik Kuralları: Birliktelik kuralları, denetimsiz öğrenme sistemlerini kullanarak yerel örüntüleri bulmayı amaçlar. En popüler ve önemli makine öğrenimi tekniklerinden biridir çünkü verideki kullanıcının önceden bilemeyeceği “altın bilgileri” dikkate alır. En çok kullanılan birliktelik kuralları teknikleri pazar sepeti ve apriori algoritmalarıdır (Kantardzic, 2020).

3. Tedarik Zinciri Yönetiminde Makine Öğrenmesi (Vaka Çalışması)

Çalışma, tedarik zinciri yönetiminde makine öğrenmesinin uygulanmasını ele almaktadır. Endüstri 4.0 bağlamında tedarikten üretime, üretimden pazarlamaya, son kullanıcılara teslimata ve satış sonrası hizmete kadar bir bilgi döngüsü oluşturan tedarik zincirindeki bu süreçlerin otomasyonunu vurgulamaktadır. Endüstrinin hızlı gelişimi, artan elektrik tüketimi ile kanıtlanmaktadır. Bu noktadan hareketle, Türkiye’de son dört yılda sektörde faturalandırılan elektrik tüketim verileri EPİAŞ web sitesinden derlenmiştir. Analiz, tüketim verileri kullanılarak Ocak 2019’dan başlayıp Ekim 2023’ten sona eren 58 aylık bir dönemi kapsamaktadır. Veriler, Şekil 3.1’de gösterildiği gibi MWh cinsinden aylık elektrik tüketimini göstermektedir.



Şekil 3.1 Sanayi Elektrik Tüketimi (MWh)

Genellikle aylık olarak artan elektrik tüketimi, destek vektör regresyonu (SVR), doğrusal regresyon (LR), ikinci dereceden polinom regresyonu (QuPR), kübik polinom regresyonu (CPR) ve kuartik polinom regresyonu (QPR) yöntemleri kullanılarak tahmin edilmiştir. Tablo 3.1 tahmin edilen ve gerçek değerleri göstermektedir.

Tablo 3.1 Sanayide Öngörülen ve Gerçekleşen Elektrik Tüketim Değerleri (MWh)

Ay	SVR	LR	QuPR	CPR	QPR	Gerçek Değer
Ocak, 19	3.9046E+06	2.6812E+06	1.8949E+06	1.6910E+06	1.9361E+06	2.0080E+06
Şubat, 19	3.9046E+06	2.7141E+06	2.0106E+06	1.8496E+06	2.0087E+06	1.9109E+06
Mart, 19	3.9046E+06	2.7470E+06	2.1233E+06	2.0015E+06	2.0883E+06	2.1107E+06
...	...	...	...	...	...	...
Ağustos, 23	3.9046E+06	4.4918E+06	3.8681E+06	3.9899E+06	4.0768E+06	4.2996E+06
Eylül, 23	3.9046E+06	4.5247E+06	3.8212E+06	3.9821E+06	4.1412E+06	4.1674E+06
Ekim, 23	3.9046E+06	4.5576E+06	3.7713E+06	3.9752E+06	4.2202E+06	4.2364E+06

Makine öğrenmesi yöntemlerinin tahmin performansını gözlemlmek için MAPE değerleri hesaplanmış ve Tablo 3.2’de verilmiştir. Gözlemler sonucunda, QPR yönteminin sektördeki aylık elektrik tüketimini tahmin etmede diğer makine öğrenimi yöntemlerinden daha iyi performans gösterdiği görülmüştür. Tedarik zinciri yönetiminde QPR yönteminin kullanılması, sektördeki aylık elektrik tüketimi için yüksek doğrulukta tahminler yapılmasını sağlayabilir.

Tablo 3.2 MAPE değerleri

SVR	LR	QuPR	CPR	QPR
19.73%	12.39%	6.32%	5.96%	5.53%

Bu durumda makine öğrenmesi çoklu yöntemlerle uygulandığında başarılı tahminler vermiştir. Modelin daha başarılı sonuçlar verebilmesi için sanayi sektörünün ilgili dönem verilerinden daha kapsamlı bir veri seti ile daha yüksek performanslı sonuçlar elde edilmesi kaçınılmazdır. Bu bağlamda tedarik zincirinin ilk aşaması olan satın alma sürecinin makine öğrenmesi yöntemleri ile tahmin edilebileceğine ve modellenilebileceğine örnek teşkil etmektedir.

4. Sonuçlar

Teknolojik gelişmelerin mevcut tedarik zinciri yönetimine yepyeni bir bakış sunacağı alandaki çalışmaların artarak devam etmesinde de bu çalışmanın hazırlık sürecinde görülmüştür. Gerek AI gerekse ML yöntemleri işletmelerin yönetim asistanlığını üstleneceği öngörülmektedir. Bu bağlamda, tedarik zinciri yönetiminin çok boyutlu ele alınmasını ve bu süreçte de işletmelerin sektördeki pozisyon hedefleri doğrultusunda bu araçları kullanması kaçınılmazdır. Alanda inceleme yapmayı planlayan araştırmacılar için Tedarik Zinciri Yönetiminde Makine Öğrenmesi uygulamalarının daha kapsamlı bir okuması için web arşivini incelmesi tavsiye olunur.

Kaynakça

- Abbasi, B., Babaci, T., Hosseinifard, Z., Smith-Miles, K., & Dehghani, M. (2020). Predicting solutions of large-scale optimization problems via machine learning: A case study in blood supply chain management. *Computers & Operations Research*, 119, 104941. <https://doi.org/10.1016/j.cor.2020.104941>
- Barbosa-Póvoa, A. P., Da Silva, C., & Carvalho, A. (2018). Opportunities and challenges in sustainable supply chain: An operations research perspective. *European Journal of Operational Research*, 268(2), 399-431. <https://doi.org/10.1016/j.ejor.2017.10.036>
- Baryannis, G., Dani, S., & Antoniou, G. (2019). Predicting supply chain risks using machine learning: The trade-off between performance and interpretability. *Future Generation Computer Systems*, 101, 993-1004. <https://doi.org/10.1016/j.future.2019.07.059>
- Boute, R. N., Gijbrecchts, J., Van Jaarsveld, W., & Vanvuchelen, N. (2022). Deep reinforcement learning for inventory control: A roadmap. *European Journal of Operational Research*, 298(2), 401-412. <https://doi.org/10.1016/j.ejor.2021.07.016>
- Carbonneau, R., Laframboise, K., & Vahidov, R. (2008). Application of machine learning techniques for supply chain demand forecasting. *European Journal of Operational Research*, 184(3), 1140-1154. <https://doi.org/10.1016/j.ejor.2006.12.004>
- Deveci, T. P., Muhammet (Ed.). (2023). *Smart and Sustainable Operations and Supply Chain Management in Industry 4.0*. CRC Press. <https://doi.org/10.1201/9781003180302>
- Feizabadi, J. (2022). Machine learning demand forecasting and supply chain performance. *International Journal of Logistics Research and Applications*, 25(2), 119-142. <https://doi.org/10.1080/13675567.2020.1803246>
- Garre, A., Ruiz, M. C., & Hontoria, E. (2020). Application of Machine Learning to support production planning of a food industry in the context of waste generation under uncertainty. *Operations Research Perspectives*, 7, 100147. <https://doi.org/10.1016/j.orp.2020.100147>
- González Rodríguez, G., Gonzalez-Cava, J. M., & Méndez Pérez, J. A. (2020). An intelligent decision support system for production planning based on machine learning. *Journal of Intelligent Manufacturing*, 31(5), 1257-1273. <https://doi.org/10.1007/s10845-019-01510-y>
- Hair, J. F. (2019). *Multivariate Data Analysis*.
- Han, J., Kamber, M., & Pei, J. (2012). *Data Mining Concepts and Techniques* (3. bs). Morgan Kaufmann is an imprint of Elsevier.
- Hastie, T., Robert, T., & Friedman, J. (2008). *The Elements of Statistical Learning* (Second). Springer US.

- Haykin, S. (2009). *Neural networks and learning machines* (3rd ed). Prentice Hall.
- İfraz, M., Aktepe, A., Ersöz, S., & Çetinyokuş, T. (2023). Demand forecasting of spare parts with regression and machine learning methods: Application in a bus fleet. *Journal of Engineering Research*, 11(2), 100057. <https://doi.org/10.1016/j.jer.2023.100057>
- Jo, T. (2021). *Machine Learning Foundations: Supervised, Unsupervised, and Advanced Learning*. Springer International Publishing. <https://doi.org/10.1007/978-3-030-65900-4>
- Kantardzic, M. (2020). *Data Mining: Concepts, Models, Methods, and Algorithms* (3. bs). Wiley.
- Khatua, A., Khatua, A., Chi, X., & Cambria, E. (2021). Artificial Intelligence, Social Media and Supply Chain Management: The Way Forward. *Electronics*, 10(19), 2348. <https://doi.org/10.3390/electronics10192348>
- Kotsiopoulos, T., Sarigiannidis, P., Ioannidis, D., & Tzovaras, D. (2021). Machine Learning and Deep Learning in smart manufacturing: The Smart Grid paradigm. *Computer Science Review*, 40, 100341. <https://doi.org/10.1016/j.cosrev.2020.100341>
- Lambert, D. M., Cooper, M. C., & Pagh, J. D. (1998). Supply Chain Management: Implementation Issues and Research Opportunities. *The International Journal of Logistics Management*, 9(2), 1-20. <https://doi.org/10.1108/09574099810805807>
- Lee, C.-Y., Chou, B.-J., & Huang, C.-F. (2022). Data science and reinforcement learning for price forecasting and raw material procurement in petrochemical industry. *Advanced Engineering Informatics*, 51, 101443. <https://doi.org/10.1016/j.aci.2021.101443>
- Li, J., Zhang, R., Jin, Y., & Zhang, H. (2022). Optimal Path of Internet of Things Service in Supply Chain Management Based on Machine Learning Algorithms. *Computational Intelligence and Neuroscience*, 2022, 1-11. <https://doi.org/10.1155/2022/4844993>
- Lin, H., Lin, J., & Wang, F. (2022). An innovative machine learning model for supply chain management. *Journal of Innovation & Knowledge*, 7(4), 100276. <https://doi.org/10.1016/j.jik.2022.100276>
- Nagar, D., Raghav, S., Bhardwaj, A., Kumar, R., Lata Singh, P., & Sindhwani, R. (2021). Machine learning: Best way to sustain the supply chain in the era of industry 4.0. *Materials Today: Proceedings*, 47, 3676-3682. <https://doi.org/10.1016/j.matpr.2021.01.267>
- Ni, D., Xiao, Z., & Lim, M. K. (2020a). A systematic review of the research trends of machine learning in supply chain management. *International Journal of Machine Learning and Cybernetics*, 11(7), 1463-1482. <https://doi.org/10.1007/s13042-019-01050-0>

- Ni, D., Xiao, Z., & Lim, M. K. (2020b). A systematic review of the research trends of machine learning in supply chain management. *International Journal of Machine Learning and Cybernetics*, 11(7), 1463-1482. <https://doi.org/10.1007/s13042-019-01050-0>
- Oliver, R. K., & Webber, M. D. (1982). *Supply chain management: Logistics catches up with strategy*.
- Pourhejazy, P., & Kwon, O. (2016). The New Generation of Operations Research Methods in Supply Chain Optimization: A Review. *Sustainability*, 8(10), 1033. <https://doi.org/10.3390/su8101033>
- Punia, S., & Shankar, S. (2022). Predictive analytics for demand forecasting: A deep learning-based decision support system. *Knowledge-Based Systems*, 258, 109956. <https://doi.org/10.1016/j.knosys.2022.109956>
- Rizwan, A., Karras, D. A., Kumar, J., Sánchez-Chero, M., Mogollón Taboada, M. M., & Altamirano, G. C. (2022). An Internet of Things (IoT) Based Block Chain Technology to Enhance the Quality of Supply Chain Management (SCM). *Mathematical Problems in Engineering*, 2022, 1-12. <https://doi.org/10.1155/2022/9679050>
- Rolf, B., Jackson, I., Müller, M., Lang, S., Reggelin, T., & Ivanov, D. (2023). A review on reinforcement learning algorithms and applications in supply chain management. *International Journal of Production Research*, 61(20), 7151-7179. <https://doi.org/10.1080/00207543.2022.2140221>
- Sammut, C., & Webb, G. I. (Ed.). (2017). *Encyclopedia of Machine Learning and Data Mining*. Springer US. <https://doi.org/10.1007/978-1-4899-7687-1>
- Shobha, G., & Rangaswamy, S. (2018). Chapter 8—Machine Learning. *İçinde Handbook of Statistics* (C. 38, ss. 197-228).
- Shyam, K. P., Ramya, V., Nadiya, S., Parashar, A., & Gideon, D. A. (2023). Systems biology approaches to unveiling the expression of phospholipases in various types of cancer—Transcriptomics and protein-protein interaction networks. *İçinde Phospholipases in Physiology and Pathology* (ss. 271-307). Elsevier. <https://doi.org/10.1016/B978-0-443-15177-4.00016-9>
- Supply Chain Diagram | Supply Chain Flowcharts | Creately*. (2021, Kasım 3). <https://creately.com/usage/supply-chain-management-tools-examples/>
- Şahin, M. (2023). Machine Learning and Data Analytics. *İçinde Smart and Sustainable Operations and Supply Chain Management in Industry 4.0*. CRC Press. 10.1201/9781003180302-15
- Tirkolaei, E. B., Sadeghi, S., Mooseloo, F. M., Vandchali, H. R., & Acini, S. (2021). Application of Machine Learning in Supply Chain Management: A Comprehensive Overview of the Main Areas. *Mathematical Problems in Engineering*, 2021, 1-14. <https://doi.org/10.1155/2021/1476043>
- Tsolaki, K., Vafeiadis, T., Nizamis, A., Ioannidis, D., & Tzovaras, D. (2023). Utilizing machine learning on freight transportation and logistics appli-

- cations: A review. *ICT Express*, 9(3), 284-295. <https://doi.org/10.1016/j.icte.2022.02.001>
- Usuga Cadavid, J. P., Lamouri, S., Grabot, B., Pellerin, R., & Fortin, A. (2020). Machine learning applied in production planning and control: A state-of-the-art in the era of industry 4.0. *Journal of Intelligent Manufacturing*, 31(6), 1531-1558. <https://doi.org/10.1007/s10845-019-01531-7>
- Wenzel, H., Smit, D., & Sardesai, S. (2019). *A literature review on machine learning in supply chain management*. <https://doi.org/10.15480/882.2478>
- Ye, S., Xiao, Z., & Zhu, G. (2015). Identification of supply chain disruptions with economic performance of firms using multi-category support vector machines. *International Journal of Production Research*, 53(10), 3086-3103. <https://doi.org/10.1080/00207543.2014.974838>
- Younis, H., Sundarakani, B., & Alsharairi, M. (2021). Applications of artificial intelligence and machine learning within supply chains: systematic review and future research directions. *Journal of Modelling in Management*, 17(3), 916-940. <https://doi.org/10.1108/JM2-12-2020-0322>
- Zage, D., Glass, K., & Colbaugh, R. (2013). Improving supply chain security using big data. *2013 IEEE International Conference on Intelligence and Security Informatics*, 254-259. <https://doi.org/10.1109/ISI.2013.6578830>

Makine Öğrenmesi Yöntemleri ve Derin Sinir Ağlarıyla Trafik Kazalarında Hayatta Kalma Sonuçlarının Tahmin Edilmesi

Mustafa Bayram Gücen¹

Özet

Bu çalışma, trafik kazalarında hayatta kalma olasılıklarını tahmin etmeye yönelik farklı makine öğrenimi algoritmalarının ve derin sinir ağlarının uygulanması ve metodolojik değerlendirmesine odaklanmaktadır. Çalışmada, parametrik yöntemler (Lojistik Regresyon) ile parametrik olmayan yöntemler (k-En Yakın Komşular, Karar Ağaçları, Rastgele Orman) ve derin öğrenme temelli Derin Sinir Ağları geniş veri setleri üzerinde sistematik bir şekilde uygulanmıştır. Veri ön işleme, model eğitimi, hiperparametre optimizasyonu ve çapraz doğrulama gibi titiz metodolojik adımlar, tahmin performansını ölçmede kullanılan doğruluk, kesinlik, duyarlılık, F1 skoru ve ROC eğrisi altında kalan alan gibi metriklerle desteklenmiştir.

Analizler, her bir modelin veri setindeki yapısal ve istatistiksel özellikleri yakalama kapasitesinin farklılık gösterdiğini ortaya koymaktadır. Bu bağlamda, trafik kazalarında hayatta kalma olasılıklarının tahminine yönelik stratejilerin geliştirilmesi, modelin hesaplama verimliliği, yorumlanabilirliği ve genelleme yeteneği gibi faktörlerin dikkate alınmasıyla optimize edilebileceği vurgulanmaktadır. Çalışmanın bulguları, trafik güvenliği alanında veri odaklı yaklaşımların metodolojik temellerini güçlendirmesi ve ileri araştırmalar için sağlam bir referans oluşturması açısından önem taşımaktadır.

1. Giriş

Karayollarında meydana gelen trafik kazaları, dünya genelinde ciddi bir halk sağlığı sorunu oluşturmakta olup, her yıl milyonlarca insanın yaşamını kaybetmesine veya kalıcı yaralanmalara yol açmaktadır. Dünya Sağlık Örgütü (DSÖ) tarafından yayımlanan 2004 tarihli Karayollarında Trafik

1 Arş. Gör. Dr., Yıldız Teknik Üniversitesi Fen Fakültesi Matematik Bölümü, mgucen@yildiz.edu.tr, 0000-0002-9920-1747

Yaralanmaları Önleme Raporu, bu sorunun küresel çapta ne denli önemli olduğunu vurgulamış ve yol güvenliğinin insan yaşamı üzerindeki etkilerini detaylı bir şekilde ele almıştır. Trafik kazaları, özellikle insanların günlük hayatlarında en çok etkileşimde bulunduğu ve en karmaşık sistemlerden biri olan yol güvenliği ile doğrudan ilişkilidir. Dünya Sağlık Örgütü (2004) raporunda bu konuda yapılan küresel çabaların yetersiz kaldığını ve yol güvenliğinin artırılması için uluslararası düzeyde daha ciddi adımlar atılması gerektiğini belirtmiştir. Bu rapor, eğer mevcut politikalar güçlendirilmezse, trafik kazalarının önlenemeyeceği ve kazaların gelecekte daha da artacağı konusunda uyarılarda bulunmuştur. Bu bağlamda, trafik kazalarının önlenmesi amacıyla küresel bir çağrı yapılmış, dünya çapında ortak çözümler üretilmesi gerektiği ifade edilmiştir.

2023 yılı itibarıyla yayımlanan Dünya Sağlık Örgütü Küresel Yol Güvenliği Durum Raporu, trafik kazası ölümlerinde bir miktar azalma olduğunu gösterse de, hâlâ bu sorunun önüne geçilebilmiş değildir. Dünya Sağlık Örgütü (2023) raporunda 2023 yılı itibarıyla 1.19 milyon trafik kazası ölümünün kaydedilmesi, yol güvenliği konusunda yapılan iyileştirmelerin bir miktar etkili olduğunu göstermektedir. Ancak, bu rakamların hala yüksek olduğu ve trafik kazalarının hâlâ en önemli ölüm nedenlerinden biri olduğu gerçeği, çözümün ne kadar acil olduğunu ortaya koymaktadır. Özellikle düşük ve orta gelirli ülkelerde, trafik kazası ölümleri gençler ve çocuklar arasında önde gelen ölüm sebeplerinden biri olmaya devam etmektedir. Yaya, bisiklet sürücüsü ve motosiklet sürücüsü ölümleri, bu ülkelerde orantısız bir şekilde fazla olup, bu ölümlerin önlenmesi için özel stratejiler geliştirilmesi gerekmektedir. DSÖ, 2030 yılına kadar trafik kazası ölümlerini ve yaralanmalarını en az yarıya indirmeyi hedefleyen küresel bir hedef belirlemiştir. Bu hedefin gerçekleştirilmesi, yalnızca yerel değil, uluslararası iş birliğini ve küresel düzeyde koordinasyonun güçlendirilmesini gerektirmektedir. Bu amaçlarla çeşitli çalışmalar yapılmış, makine öğrenmesi ve derin öğrenme teknikleri kullanılmıştır. Behboudi vd. (2024), trafik kazası analizi ve tahmini konusunda makine öğrenimi tekniklerinin son yıllarda elde edilen önemli ilerlemelerini kapsamlı bir şekilde inceleyen bir derlemedir. Son beş yılda yapılan 191 çalışmayı ele alarak, kaza riskini, sıklığını, şiddetini, süresini tahmin etme ve kaza verilerinin genel istatistiksel analizine odaklanmıştır. Bu çalışmanın, kaza analizi ve tahminine ilişkin çok çeşitli alanlardaki son teknolojiye dair kapsamlı bir inceleme sunduğunu belirtmektedir. Derleme, farklı veri kaynaklarının ve gelişmiş makine öğrenmesi tekniklerinin entegrasyonunun tahmin doğruluğunu artırmada ve trafik verilerinin karmaşıklığını yönetmede etkinliğini vurgulamaktadır. Mevcut durumu haritalayarak literatürdeki boşlukları belirleyerek, bu

çalışma, Dünya Sağlık Örgütü'nün (DSÖ) hedeflerine paralel olarak, 2030 yılına kadar trafikle ilgili ölümleri ve yaralanmaları önemli ölçüde azaltmak amacıyla gelecekteki araştırmalara rehberlik etmeyi hedeflemektedir.

Makine öğrenimi ve Derin Sinir Ağları gibi ileri düzey teknolojiler, özellikle karmaşık verileri işleme ve verilerdeki gizli ilişkileri ortaya çıkarma konusunda önemli bir yere sahiptir. Bu teknolojiler, geleneksel yöntemlerin gözden kaçırabileceği önemli desenleri keşfetme yeteneğine sahip olup, büyük veri setlerini daha verimli bir şekilde analiz etme potansiyeline sahiptir. Theofilatos (2019), makine öğreniminin büyük veri setleri üzerinde uygulandığında, bu verilerdeki gizli ilişkileri ve karmaşık desenleri ortaya çıkarma gücüne sahip olduğunu göstermiştir. Benzer şekilde, Ballamudi (2019) ve Santos (2021) gibi araştırmacılar, derin sinir ağlarının özellikle yol güvenliği ve trafik kazalarının çözümünde önemli bir araç olduğunu vurgulamış, bu tekniklerin verilerdeki karmaşık ilişkileri anlamada üstün bir performans sergilediğini belirtmişlerdir. Kumar vd. (2021), araba kazalarının önlenmesi veya tespiti konusunda birçok çalışma ve araştırma yapıldığını belirtmişlerdir. Çoğu araştırma, kaza oluşturabilecek nesnelerin tespit edilmesine odaklanmaktadır. Önerilen sistemde, kazanın meydana geldiği durumları tespit etmek amacıyla, birbirine yakın hareket eden araçlardan uygun bilgilerin toplanarak makine öğrenimi araçlarıyla işlenmesi sağlanmıştır. Makine öğrenimi teknikleri, normal davranışlardan anormal davranışları ayırt etmede başarı sağlamıştır. Sistemin ana hedefi, trafik davranışlarını incelemek ve mevcut trafikten farklı hareket eden araçları yol kazası olarak değerlendirmektir.

Trafik kazalarının, emniyet kemeri kullanımı, çarpışma hızı ve yaş gibi faktörlerle ilişkili olduğu ve bu etmenlerin kazaların sonuçlarını önemli ölçüde etkilediği gerçeği, bu ileri düzey teknolojilerin trafik kazalarının anlaşılması ve çözülmesinde ne denli önemli bir rol oynayabileceğini göstermektedir.

Bu çalışmanın amacı, trafik kazalarında hayatta kalma sonuçlarını etkileyen faktörleri tahmin etmek ve analiz etmek için makine öğrenimi ve Derin Sinir Ağları tekniklerini kullanmaktır. Derin Sinir Ağları, verilerdeki karmaşık yapıları tanıma ve öğrenme yeteneği sayesinde, trafik kazalarının ve bu kazaların neden olduğu hayatta kalma sonuçlarının daha doğru bir şekilde analiz edilmesini sağlar. Bu yöntemler, kazaların oluşmasındaki etkenleri daha net bir biçimde ortaya koyabilmekte ve bu etkenler arasındaki ilişkilerin daha derinlemesine anlaşılmasını mümkün kılmaktadır. Özellikle, emniyet kemeri kullanımı, çarpışma hızı, yaş gibi faktörlerin kazaların sonuçlarını nasıl etkilediği üzerine yapılan analizler, kazaların önlenmesine yönelik stratejilerin geliştirilmesinde önemli bir katkı sağlayacaktır. Makine öğrenimi

ve Derin Sinir Ağları kullanılarak yapılan bu tür analizler, yol güvenliğinin artırılması adına daha etkili, hedeflenmiş ve bilimsel temellere dayalı stratejilerin oluşturulmasına olanak tanıyacaktır. Bu araştırma, trafik kazası ölümlerinin ve yaralanmalarının azaltılması, yol güvenliğinin artırılması ve hayatta kalma oranlarının iyileştirilmesi için daha doğru ve etkili çözümler geliştirmeyi hedeflemektedir. Sonuç olarak, bu çalışma, trafik kazalarının nedenlerini anlamada ve bu nedenlerin etkilerini analiz etmede, ileri düzey makine öğrenimi tekniklerinin kullanılmasının, güvenli yol sistemleri oluşturulmasına ve halk sağlığı sorunlarının çözülmesine katkıda bulunacak önemli bir adım olduğunu ortaya koymaktadır.

2. Literatür İncelemesi

Makine öğrenimi ve derin öğrenme, trafik kazalarıyla ilgili yaralanma şiddetini tahmin etmek ve trafik güvenliğini iyileştirmek için önemli araçlar olarak kullanılmaktadır. Literatürde, bu tekniklerin çeşitli yönleriyle kazaların şiddetini tahmin etme ve güvenlik önlemlerini geliştirme potansiyeli farklı araştırmacılar tarafından bir çok çalışmada araştırılmıştır.

Obasi ve Benson (2023) trafik kazalarının şiddetini tahmin etmek için Naive Bayes, Rastgele Orman, Lojistik Regresyon ve Yapay Sinir Ağları gibi makine öğrenimi modellerini kullanarak bir analiz çerçevesi önermiştir. Çalışmada, Birleşik Krallık'tan alınan 2005-2014 yıllarına ait trafik kazası verileri ile farklı makine öğrenimi modelleri karşılaştırılmış ve Random Forest ile Lojistik Regresyon modellerinin daha iyi sonuçlar verdiği gözlemlenmiştir. Bu çalışma, trafik kazası yönetimi ve kontrol önlemlerinin uygulanmasında makine öğrenimi tekniklerinin etkinliğini vurgulamaktadır.

Nandurge ve Dharwadkar (2017), yol trafik kazalarının temel faktörlerini belirlemenin, kaza verisi analizinin ana hedeflerinden biri olduğunu belirtmişlerdir. Yol kazası verilerinin heterojen yapısı nedeniyle analizlerin zorlayıcı olabileceği ifade edilmiştir. Bu heterojenliği aşmak için verilerin bölümlendirilmesi gerektiği vurgulanmıştır. Önerilen yöntem, yol kazası verilerinin segmentasyonunda ana görev olarak k-means kümeleme yöntemini kullanmaktadır. Ayrıca, k-means kümeleme algoritması ile tanımlanan kümelerin ve tüm veri setinin oluşumuyla ilgili durumları keşfetmek amacıyla ilişki kuralı madenciliği uygulanmıştır. K-means kümeleme ve ilişki kuralı madenciliğinin birleşik sonucu, önemli bilgilere ulaşılmasını sağlamaktadır. Patil vd. (2020) ise, Bengaluru'daki yol kazalarını analiz etmek için makine öğrenimi ve K-means algoritması kullanarak bir çalışmada kazaların yoğunlaştığı bölgeleri ve bu kazaların kök nedenlerini incelediler. Çalışma, trafik kazalarının gelişmiş ve gelişmekte olan ülkeler için büyük bir

tehdit olduğunu ve kazaların, hava koşulları, dik virajlar ve insan hataları gibi çeşitli faktörlerden kaynaklandığını vurgulamaktadır. Ayrıca, kazaların yol açtığı yaralanmaların bazen fark edilmeyen ancak sağlık üzerinde uzun vadeli etkiler yaratabilecek nitelikte olduğunu belirtmektedir. Zhengjing vd. (2021) ise, derin öğrenme tabanlı bir analitik çerçeve önererek, trafik kazası yaralanma şiddetini tahmin etmede kullanılan katkı sağlayan faktörlerin önemini analiz etmiştir. Çalışmada, CatBoost adı verilen makine öğrenimi yaklaşımı ile katkı sağlayan faktörlerin önemi ve bağımlılığı analiz edilmiş, K-means kümeleme kullanılarak veriler farklı sınıflara ayrılmıştır. Yüksek korelasyonlu faktörler kullanılarak derin öğrenme modeli ile yaralanma şiddeti tahmin edilmiştir. Sonuçlar, önerilen çerçevenin, trafik kazası yaralanma şiddetini tahmin etmede ve yayaların güvenliğini artırmada etkili olduğunu göstermektedir. Ahmed vd. (2021), trafik kazalarının ölüm ve yaralanmalara yol açmasının yanı sıra doğrudan ve dolaylı kayıplara neden olan küresel bir sorun olduğunu belirtmişlerdir. Çalışmada, büyük trafik verileri ve yapay zeka algoritmaları kullanılarak kaza şiddetinin tahmin edilmesi veya kazaların riskinin azaltılması için umut verici bir çözüm geliştirilmesi amaçlanmıştır. Mevcut çalışmalarda genellikle yol geometrisi, çevre ve hava koşulları gibi faktörler üzerinde durulmuş olsa da, alkol, ilaç kullanımı, yaş ve cinsiyet gibi insan faktörleri sıklıkla göz ardı edilmiştir. Çalışmada, kaza şiddetinin tahminine katkı sağlayan çeşitli faktörler incelenmiş ve tekli ve çoklu makine öğrenimi yöntemlerinin performansları, doğruluk, kesinlik, duyarlılık, F1 skoru ve ROC (receiver operator characteristic) eğrileri(AUC-ROC eğrisi altında kalan alan) gibi metriklerle karşılaştırılmıştır. Araştırma, kaza şiddetini sınıflandırma problemi olarak ele almış ve sonuçlar, Rastgele Orman(Random Forest) algoritmasının diğer yöntemleri geride bıraktığını ve her iki sınıflandırma türünde de en iyi sonuçları verdiğini göstermektedir. Araştırmada, Lojistik Regresyon ve kNN gibi tekli yöntemlerin benzer performanslar gösterdiği ancak ansambl yöntemlerinin, özellikle Rastgele Orman, XGBoost ve AdaBoost'un daha yüksek doğrulukla kaza şiddetini tahmin edebildiği tespit edilmiştir.

Derin sinir ağları, görüntü sınıflandırma görevlerinde gösterdiği yüksek başarı sayesinde tıbbi görüntüleme, otonom sürüş, güvenlik sistemleri ve endüstriyel kalite kontrol gibi alanlarda yaygın olarak kullanılmaktadır. Röntgen ve MR gibi tıbbi görüntülerde hastalık belirtilerini tespit etmenin yanı sıra, otonom araçlarda trafik işaretlerini ve yayaları tanımlayarak güvenli sürüşe katkı sağlayan derin öğrenme modelleri, geniş veri kümeleri üzerinde gerçekleştirilen derin analizlerle model performansını optimize etmekte ve hata oranlarını azaltmaktadır. Ayrıca, açıklanabilir yapay zeka yöntemleri ile karar süreçlerinin şeffaflaştırılması, kazalara yol açabilecek

nedenlerin belirlenmesi ve risk analizi gibi kritik görevlerde etkin çözümler sunmaktadır. Derin sinir ağları aynı zamanda tahminlemede ve kazalara yol açabilecek nedenlerin incelenmesinde kullanılmaktadır. Singh vd. (2020), yol kazalarını tahmin etmek için derin sinir ağları modelinin kullanılmasını önermektedir. Derin sinir ağları, bir veya daha fazla gizli katman ve çok sayıda düğüm içerir. Çalışmada, alınan kaza verileri ve çok sayıda kaza kaydı bulunmaktadır ve elde edilen sonuçlar, Derin sinir ağlarının yol kazalarının tahmininde geliştirilmiş bir performans gösterdiğini ortaya koymaktadır. Lin vd. (2021), yüksek kaza riski tahmin modeli geliştirmek için derin öğrenme yöntemlerini kullanarak trafik kazası verilerini analiz etmiş ve iyileştirme için öncelikli kavşakları belirlemiştir. Trafik kazası verilerinin derlendiği bir veri tabanı oluşturulmuş ve bu veriler üzerinden kazaların olabileceği yüksek riskli kavşakları tahmin edebilmek amacıyla çeşitli makine öğrenimi yöntemleri kullanılarak bir kavşak kaza riski tahmin modeli geliştirilmiştir. Vicent vd. (2025), dünya genelinde her yıl binlerce kişinin trafik kazalarında hayatını kaybettiğine dikkat çekmektedir. Trafik kazalarında tıbbi yardıma ihtiyaç duyulup duyulmadığının tahmin edilmesi, bu konuda önemli bir sorundur. Bu amaçla, yazarlar, tıbbi yardıma ihtiyaç duyulup duyulmadığını ayırt etmek için geliştirilmiş yeni bir derin öğrenme modeli sunmaktadır.

Vasavi (2018), yol kazaları verilerinde gizli desenleri ortaya çıkarmak için makine öğrenimi tekniklerinin kullanımını incelemiştir. Çalışmasında, K-medoids ve beklenti-maksimizasyon algoritmaları ile kümeler oluşturulmuş ve a priori algoritması kullanılarak kazalar arasındaki ortak özellikler keşfedilmiştir. Sonuçlar, makine öğrenimi tekniklerinin yol kazalarındaki gizli örüntüleri başarılı bir şekilde çıkarabildiğini göstermiştir. Kodepogu vd. (2023), yol güvenliği yönetimi kapsamında, trafik kazalarının şiddetini tahmin eden modellerin geliştirilmesinin önemine odaklanmaktadır. Çalışma, kaza şiddetini etkileyen çok yönlü faktörleri inceleyerek, hız ve araç boyutu gibi taşıt özellikleri ile yol tasarımı ve trafik hacmi gibi yol karakteristiklerini ele almıştır. Ayrıca, sürücü demografisi, deneyimi ve hava koşulları gibi dış etkenler de analiz edilmiştir. Veri analizi ve makine öğrenimi tekniklerindeki son gelişmeler, trafik kazası şiddetinin tahmin edilmesi konusunda önemli bir rol oynamaktadır. Geleneksel istatistiksel yöntemlerin aksine, makine öğrenimi modelleri karmaşık değişken etkileşimlerini daha iyi yakalayarak daha yüksek tahmin doğruluğu sunmaktadır. Ancak, bu modellerin etkinliği, kullanılan verilerin kalitesine ve kapsamına bağlıdır. Çalışma, standart veri toplama ve raporlama yöntemlerinin önemini vurgulamaktadır. Mevcut literatürün titizlikle analiz edilmesi sonucunda, çalışmada yol kazası tahmin modellerinin temel kavramları, teorik çerçeveleri ve veri kaynakları ele alınmıştır. Bulgular, yol kazalarının sıklığını ve ciddiyetini azaltmayı

amaçlayan gelişmiş modelleme tekniklerinin sürekli olarak geliştirilmesi ve iyileştirilmesi gerektiğini göstermektedir. Bu tür çalışmalar, trafik kontrolü ve kamu güvenliği önlemlerinin iyileştirilmesine önemli katkılar sunmaktadır. Bai vd. (2025), trafik kazalarının bireylere, topluma ve altyapıya önemli maliyetler getirdiğini vurgulayarak, kazaların maliyetlerini tahmin etmek için geliştirilmiş makine öğrenimi tekniklerini entegre eden kapsamlı bir çerçeve önerdiler. Çalışmada model, tarihsel kaza verileri, hava koşulları, yol altyapısı ve demografik faktörler gibi bağlamsal bilgilerle eğitilmiştir. Çalışma, bu yaklaşımın geleneksel yöntemlere göre üstün tahmin doğruluğu sağladığını ve farklı faktörlerin maliyet tahminine etkilerini değerlendiren duyarlılık analizleri sunduğunu belirtmektedir. Önerilen çerçeve, trafik kazası maliyetlerini tahmin etmek için güçlü ve verimli bir çözüm sunmakta ve toplum ve altyapı üzerindeki olumsuz etkileri hafifletmeye yönelik proaktif önlemleri kolaylaştırmaktadır.

Bu çalışmalar, makine öğrenimi ve derin öğrenme tekniklerinin trafik kazalarının tahmin edilmesi, kazaların şiddetinin analiz edilmesi ve trafik güvenliğinin iyileştirilmesi konusunda etkili bir araç olduğunu göstermektedir. Ayrıca, bu tekniklerin trafik kazası verilerindeki gizli desenleri ortaya çıkarma, önemli faktörleri belirleme ve yol güvenliği önlemlerini geliştirme açısından büyük potansiyel sunduğu vurgulanmaktadır. Bu tür teknolojilerin, daha güvenli bir trafik ortamı yaratmak ve kazaların önlenmesine yönelik stratejiler geliştirmek için önemli bir rol oynayabileceği belirtilmiştir. Ayrıca, bu tekniklerin çeşitli trafik kazası veri setleriyle uyumlu olarak uygulanabilir olması, gelecekteki trafik kazası yönetimi ve kontrol sistemlerinin daha verimli hale gelmesine katkı sağlayacaktır.

3. Veri Kümesi ve Yöntem

3.1. Veri Kümesi

Bu çalışmada kullanılan veri kümesi, çeşitli veri kaynaklarından alınan trafik kazalarından alınan veriler baz alınarak elde edilmiş trafik kazalarına dair detaylı kayıtları içermektedir. Veri kümesi; yaş, cinsiyet, çarpışma hızı, kask kullanımı, emniyet kemeri kullanımı ve hayatta kalma durumu gibi faktörleri içeren demografik, davranışsal ve durumsal özellikleri kapsamaktadır. Bu özellikler, trafik kazalarında hayatta kalma olasılıklarını etkileyen çeşitli faktörleri anlamaya yönelik önemli bilgiler sunmaktadır. Veri kümesi, yol kazalarının daha iyi analiz edilmesi için toplanmış olan değişkenleri içermektedir ve bu faktörlerin hayatta kalma üzerinde nasıl bir etkisi olduğuna dair önemli ipuçları sağlayabileceği düşünülmektedir. Veriler, veri analizi öncesi makine öğrenimi algoritmalarına uygun olacak şekilde

ön işleme tabi tutulacaktır. Veri kümesi, <https://www.kaggle.com/datasets/himelsarder/road-accident-survival-dataset> web adresinden elde edilmiştir.

Veri kümesinde kullanılan değişkenlere ait bilgiler Tablo 1'de verilmiştir.

Tablo 1. Değişkenlere Ait Bilgiler

Değişken Adı	Değişken Değerleri
Yaş (Age)	18-69
Cinsiyet (Gender)	Erkek (Male), Kadın (Female)
Çarpışma Hızı (Speed_of_Impact)	20-119 km/s (20-119 km/h)
Kask Kullanıldı mı (Helmet_Used)	Evet (Yes), Hayır (No)
Emniyet Kemeri Kullanıldı mı? (Seatbelt_Used)	Evet (Yes), Hayır (No)
Hayatta Kalma Durumu (Survived)	Evet (1), Hayır (0)

3.2. Veri Ön işleme Süreci

Veri ön işleme, makine öğrenmesi projelerinde modelin başarısını doğrudan etkileyen zorunlu bir adımdır. Gerçek dünya verileri genellikle eksik, hatalı veya düzensiz olduğundan, ham verinin doğrudan kullanılması modelin performansını düşürebilir. Bu nedenle, eksik verilerin doldurulması, kategorik değişkenlerin sayısallaştırılması ve veri ölçeklendirme gibi işlemler uygulanarak verinin analize uygun hale getirilmesi sağlanır.

Eksik verilerin yönetimi, modelin güvenilirliğini artırmak için yaygın olarak ortalama, medyan veya mod gibi istatistiksel yöntemlerle gerçekleştirilir. Kategorik değişkenler, One-Hot Encoding gibi tekniklerle sayısal formata dönüştürülerek modelin bu verileri işlemesi sağlanır. Ayrıca, farklı ölçeklerdeki değişkenlerin dengelenmesi için Min-Max Scaling veya Standardizasyon yöntemleri kullanılarak verinin uygun bir dağılıma getirilmesi sağlanır. Bu adımlar, modelin daha doğru ve genelleştirilebilir sonuçlar üretmesine yardımcı olur. Aşağıda, veri ön işleme sürecinin açıklamasını yaparak hangi yöntemlerin kullanıldığını detaylı şekilde ele alacağız.

3.2.1. Eksik Verilerin Doldurulması

Eksik verilerin doldurulması, modelin doğruluğunu artırmak ve analiz sürecini kesintisiz hale getirmek için yapılır. Eksik veriler şu şekilde doldurulmuştur:

Cinsiyet (Gender) Sütunu: Bu sütundaki eksik değerler, sütunun mod değeriyle doldurulmuştur. Mod, bir veri kümesindeki en sık tekrarlanan

değeri ifade eder. Yani, verilerin içinde en çok görülen kategori alınır ve eksik değerler bu kategoriyle doldurulur.

Çarpma Hızı (Speed_of_Impact) Sütunu: Bu sütundaki eksik veriler, sütunun ortalaması ile doldurulmuştur. Ortalama, verilerin toplamının, veri sayısına bölünmesiyle hesaplanır. Bu, eksik veriler için yaygın bir doldurma yöntemidir.

Kask Kullanımı (Helmet_Used) ve Emniyet Kemeri Kullanımı (Seatbelt_Used) Sütunları: Bu kategorik sütunlardaki eksik değerler de yine mod değeriyle doldurulmuştur. En sık görülen kategori seçilerek eksik değerler tamamlanır.

3.2.2. Kategorik Verilerin Sayısal Verilere Dönüştürülmesi

Makine öğrenmesi algoritmaları, kategorik verileri doğrudan işleyemez. Bu nedenle, One-Hot Encoding kullanılarak kategorik değişkenler sayısal verilere dönüştürülmüştür.

One-Hot Encoding yöntemiyle, her kategoriye karşılık gelen yeni bir sütun oluşturulur ve bu sütun, kategoriye ait her veri için 1 veya 0 değerleri alır. Örneğin, Cinsiyet sütununda Erkek ve Kadın gibi iki kategori varsa, bu iki kategori için ayrı sütunlar yaratılır. Verinin Erkek olması durumunda, Erkek sütunu 1, Kadın sütunu ise 0 olacaktır. Diğer kategori için de aynı işlem yapılır.

3.2.3. Veri Kümesinin Eğitim ve Test Olarak Ayrılması

Veri kümesindeki eğitim ve test kümeleri, modelin doğruluğunu değerlendirmek için ayrılır. Bu işlemin amacı, modelin eğitim verisiyle aşırı öğrenmesini (overfitting) engellemektir. Bu adımda stratifikasyon uygulanır:

Stratifikasyon (Stratification) veri kümesinin sınıflara göre oranlarını koruyarak eğitim ve test kümesine ayırma işlemidir. Yani, her iki küme de, bağımlı değişkenin (y) sınıflarına orantılı olarak ayrılır. Bu işlem, özellikle dengesiz sınıflara sahip veri setlerinde modelin daha dengeli bir şekilde eğitilmesini sağlar.

Stratifikasyon işlemi, her sınıfın veri kümesindeki oranını korur. Bu, eğitim ve test kümesinde de aynı sınıf oranlarının korunmasını sağlar.

3.2.4. Sayısal Verilerin Ölçeklendirilmesi

Sayısal veriler, farklı aralıklara ve büyüklüklere sahip olabilir. Bu, modelin öğrenme sürecini zorlaştırabilir. Bu nedenle, MinMaxScaler ve StandardScaler gibi ölçeklendirme yöntemleri kullanılmıştır. Her iki yöntem de verilerin belirli bir aralığa çekilmesini sağlar, ancak kullandıkları formüller farklıdır.

MinMaxScaler yöntemi, verileri belirli bir aralıkta (genellikle [0, 1]) yeniden ölçeklendirir. Verinin her bir değeri, minimum değeri çıkarılarak ve maksimum değere bölünerek normalize edilir. Bu işlem, tüm verilerin aynı aralığa çekilmesini sağlar.

MinMaxScaler, verileri [0,1] aralığına ölçeklendirmek için şu formülü kullanır:

$$X' = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

Burada: X' ölçeklendirilmiş değer, X orijinal değer, $X_{\min}$ veri kümesindeki minimum değer ve $X_{\max}$ veri kümesindeki maksimum değer olarak alınmaktadır.

StandardScaler yöntemi, verilerin ortalama ve standart sapma kullanılarak yeniden ölçeklendirilmesini sağlar.

$$X' = \frac{X - \mu}{\sigma} \quad (2)$$

Burada: X' standardize edilmiş değer, X orijinal değer, μ veri kümesinin ortalaması ve σ veri kümesinin standart sapması olarak alınmaktadır.

Bu yöntemle her verinin, veri kümesinin ortalamasından çıkarılması ve standart sapmaya bölünmesiyle normalize edilmesi sağlanır. Bu, verilerin sıfır ortalamaya ve birim standart sapmaya sahip olmasını sağlar.

3.2.5. Korelasyon Analizi

Korelasyon, bağımsız değişkenlerle bağımlı değişken arasındaki ilişkiyi gözlemlememizi sağlar. Tablo 2 bu amaçla yapılan bir korelasyon analizinin sonuçlarını sunmaktadır. Bu analizde, parametrik korelasyon yöntemlerinden olan Pearson korelasyonunun bir türü Point-Biserial korelasyon katsayısı (r_{pb}) ve non-parametrik korelasyon yöntemlerinden Spearman Korelasyon

Katsayısı (r_s , p_value) kullanılmıştır. Point-Biserial korelasyon katsayısı (r_{pb}), bir sürekli (nicel) değişken ile bir ikili (dichotomous) (örneğin, 0 ve 1 gibi) kategorik değişken arasındaki ilişkiyi ölçmek için kullanılır. Bu katsayı Pearson korelasyon katsayısının özel bir türüdür, yani hesaplama yöntemi Pearson korelasyon katsayısına dayanır. Formülü şu şekildedir:

$$r_{pb} = \frac{M_1 - M_0}{s} \times \sqrt{\frac{p_1 p_0}{n^2}} \quad (3)$$

Burada: M_1, M_0 , bağımsız değişkenin sürekli değişken için kategoriye göre ortalama değerleridir, s, sürekli değişkenin toplam standart sapmasıdır, p_1, p_0 , bağımsız değişkenin iki kategorisinin oranlarını ifade eder.

Nonparametrik korelasyon yöntemleri, verilerin normal dağılım göstermediği veya doğrusal olmayan ilişkilerin bulunduğu durumlarda, değişkenler arasındaki ilişkiyi değerlendirmek için tercih edilen güçlü araçlardır ve verilerin sıralanmasına dayalı olarak monotonik ilişkilerin kuvvetini ölçmede kullanılır Kendall (1948). Gibbons ve Chakraborti (2014) nonparametrik istatistiksel çıkarım yöntemlerini detaylı bir şekilde ele alarak, bu tekniklerin modern uygulamalardaki geçerliliğini ve pratik faydalarını ortaya koymaktadır. Bu yöntemler, özellikle parametrik testlerin varsayımlarının ihlal edildiği durumlarda, güvenilir sonuçlar üretebilmekte ve araştırmacılara verinin yapısal özelliklerini daha doğru bir şekilde yorumlama imkanı sağlamaktadır. Spearman Korelasyon Katsayısı (r_s), değişkenler arasındaki sıra korelasyonunu ölçen parametrik olmayan bir istatistiksel yöntemdir. Pearson korelasyon katsayısından farklı olarak, Spearman katsayısı değişkenlerin sıralarına dayalı olduğu için normal dağılım varsayımı gerektirmez ve değişkenler arasındaki doğrusal olmayan monoton ilişkileri de yakalayabilir Spearman (1904). Spearman Korelasyon Katsayısı (r_s) kullanılarak değişkenler arasındaki ilişki değerlendirilmiştir.

$$r_s = \frac{1 - 6 \sum d_i^2}{n(n^2 - 1)} \quad (3)$$

p-değeri (p-value), hipotez testinde korelasyon katsayısının istatistiksel olarak anlamlı olup olmadığını belirlemek için kullanılır. Düşük bir p-değeri ($p < 0,05$), iki değişken arasında istatistiksel olarak anlamlı bir ilişki olduğunu gösterirken, yüksek bir p-değeri ($p > 0,05$) bu ilişkinin tesadüfi olabileceğini işaret eder. Tablo 2, cinsiyet, yaş, emniyet kemeri kullanımı, çarpışma

hızı ve kask kullanımı gibi özelliklerin bağımlı değişkenle olan ilişkisini göstermektedir.

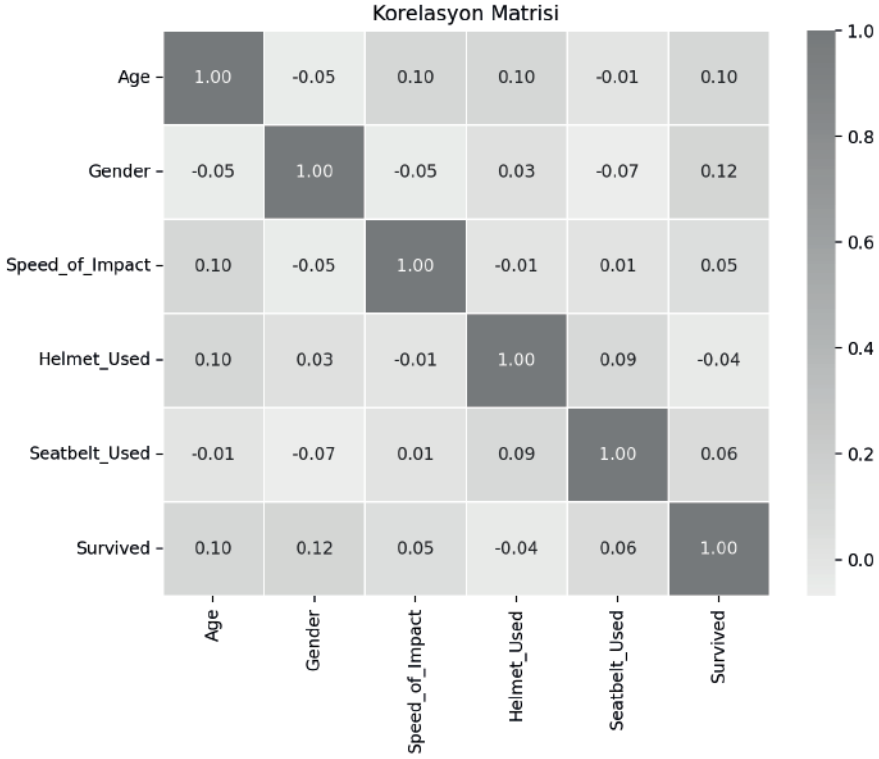
Şekil 1’de verilen korelasyon matrisi, veri setindeki değişkenler arasındaki doğrusal ilişkileri gösteren simetrik bir matristir. Bu matris, değişken çiftleri arasındaki korelasyon katsayılarını hesaplayarak, güçlü veya zayıf ilişkilerin belirlenmesine olanak tanır. Pozitif veya negatif korelasyon değerleri, değişkenlerin birbirleriyle nasıl etkileşime girdiğini anlamada önemli bir istatistiksel araçtır.

Korelasyon katsayıları genellikle düşük seviyelerde olup, çoğu ilişkinin Spearman Korelasyon Katsayısı açısından istatistiksel olarak anlamlı olmadığı gözlemlenmektedir. Özellikle p-değerleri 0,05’ten büyük olan özelliklerde, bu değişkenlerin bağımlı değişkenle güçlü bir ilişki kurmadığı söylenebilir. Bu sonuç, analiz edilen değişkenler arasında zayıf ilişkilerin varlığını ve çoğunun etkisinin sonuçlar üzerinde çok fazla olmadığını ortaya koymaktadır.

Tablo 2. Bağımlı Değişken ile Bağımsız Değişkenler Arasındaki Korelasyon Değerleri

Değişken Adı	r_{pb}	r_s	p-değeri (p-value)
Cinsiyet (Gender)	0,121845	0,121845	0,085657
Yaş (Age)	0,110090	0,103877	0,120695
Emniyet Kemeri Kullanımı (Seatbelt_Used)	0,059263	0,059263	0,404518
Çarpışma Hızı (Speed_of_Impact)	0,042198	0,046602	0,552989
Kask Kullanımı (Helmet_Used)	-0,041353	-0,041353	0,560969

Korelasyon ile ilgili değerler daha yaygın bir görsel ifade biçimi olan Şekil 1. de gösterimi verilen spearman korelasyon matrisi üzerinde daha ayrıntılı ve simetrik olarak incelenebilir.



Şekil 1. Spearman Korelasyon matrisi(Yaş (Age), Cinsiyet (Gender), Çarpışma Hızı (Speed_of_Impact), Kask Kullanımı (Helmet_Used), Emniyet Kemerini Kullanımı (Seatbelt_Used), Hayatta Kalma Durumu (Survived) olarak verilmektedir.)

3.2.6. Veri Artırma Teknikleri

Veri analizi ve modelleme sürecinde kullanılan çeşitli özellik mühendislik teknikleri, modelin doğruluğunu ve genelleme yeteneğini artırmada önemli rol oynamaktadır. Bu çalışmada, bir yol kazası verisinde hayatta kalma tahminlerini geliştirmek amacıyla bazı gelişmiş özellik mühendislik adımları uygulanmıştır. Özellikle, birinci derece ve ikinci derece özellikler oluşturulmuş ve bu özelliklerin kazaların sonuçları üzerindeki etkisi araştırılmıştır.

İlk olarak, ortalama türev(Average Derivative (AD)) özellikleri hesaplanmıştır. Bu özellikler, kazaların şiddetini belirlemede önemli olabilecek hız değişimlerini temsil eder. Çarpma Hızı sütunu üzerindeki farklar kullanılarak, her iki pencere boyutu için (2 ve 3) ortalama türev özellikleri oluşturulmuştur. Bu, çarpma hızındaki değişimin zaman içindeki etkisini analiz etmek için bir yöntemdir. Bu tür türevsel özellikler, özellikle

zaman serisi verilerinde modelin daha doğru tahminler yapmasına olanak tanıyabilir.

Ardından, polinom özellikleri oluşturulmuştur. Bu adımda, ikinci dereceden polinom özellikleri hesaplanmış ve her bir sayısal özelliğin kendisiyle ve diğer sayısal özelliklerle olan etkileşimlerinin etkisi analiz edilmiştir. Polinom özellikleri sınıfı kullanılarak, sayısal verilerdeki etkileşimler, eğilimler ve doğrusal olmayan ilişkiler dikkate alınmıştır. Bu özellikler, modelin doğrusal olmayan ilişkileri öğrenmesini sağlayarak tahmin gücünü artırabilir.

Yapılan bu polinom dönüşümü ile elde edilen yeni özellikler, modelin daha zengin bir veri seti üzerinde çalışmasını sağlar. Bu özelliklerin ismi, orijinal veri setinde yer alan sayısal sütunlardan türetilmiştir ve oluşturulan özelliklerin adı, polinom dönüşümü işlemiyle birlikte, her bir özelliğin etkileşimini ve seviyelerini ifade etmektedir. Bu işlem, modelin karmaşık ilişkileri daha iyi öğrenmesine yardımcı olur.

Son olarak, cinsiyet ve hayatta kalma gibi orijinal kategorik değişkenler, polinom özelliklerle oluşturulan yeni veri setine eklenmiştir. Bu adım, kategorik özelliklerin etkilerini modelin öğrenme sürecine dahil ederken, aynı zamanda verilerin bütünlüğünü korumaktadır. cinsiyet ve hayatta kalma değerlerinin olduğu sütunlar, kazalarda hayatta kalma tahminlerini etkileyebilecek önemli faktörlerdir ve bu nedenle, polinom özelliklerle birlikte değerlendirilmiştir.

Sonuç olarak, yapılan özellik mühendisliği işlemleri, yol kazası verisinde hayatta kalma tahminlerinin doğruluğunu artırmak için önemli bir adımdır. Ortalama türev ve polinom dönüşümleri gibi teknikler, modelin karmaşık ilişkileri öğrenmesine ve daha hassas tahminlerde bulunmasına olanak tanımaktadır. Bu yaklaşım, veri analizi ve makine öğrenmesi uygulamalarında etkili bir özellik mühendisliği stratejisi olarak değerlendirilebilir.

3.3. Yöntem

Bu çalışmada kullanılan veri ön işleme ve özellik mühendislik tekniklerine dair yapılan adımlar detaylı bir şekilde açıklanmıştır. Verilerin işlenmesi ve modelin doğruluğunu artırmak amacıyla çeşitli yöntemler uygulanmıştır. Bu adımlar, verinin anlamlı özellikler haline getirilmesini ve modelin genel performansının iyileştirilmesini hedeflemektedir. İlk adımda, veri setinden kullanılacak sayısal sütunlar seçilmiştir. Bu sütunlar, Yaş (Age), çarpma hızı (Speed of Impact), kask kullanımı (Helmet Used) ve emniyet kemerinin kullanımı (Seatbelt Used) olarak belirlenmiştir. Bu sütunlar, kazaların sonuçlarını anlamada kritik faktörler olarak kabul edilmiştir. Veri setinde

ortalama türev yani Average Derivative (AD) özellikleri oluşturulmuştur. Bu işlemde, hızdaki değişimin zaman içindeki etkisi dikkate alınmıştır. Çarpma hızı sütunu üzerinden hesaplanan farklar, belirli bir pencere uzunluğuna göre alınarak hız değişiminin ortalama türevini temsil eden yeni özellikler eklenmiştir. Pencere uzunlukları olarak 2 ve 3 kullanılmıştır. Bu işlem, kazaların etkisinin zamanla nasıl değiştiğini daha iyi anlamaya olanak sağlar. Bir sonraki adımda, Polinom Özellikleri oluşturulmuştur. Bu işlemde, orijinal sayısal özelliklerin ikili etkileşimleri ve ikinci dereceden polinomları hesaplanmıştır. Polinom Özellikleri sınıfı kullanılarak, ikinci dereceden polinom özellikler oluşturulmuş ve bu özelliklerin sayısal veri setine eklenmesi sağlanmıştır. Bu teknik, modelin doğrusal olmayan ilişkileri öğrenmesini kolaylaştırarak tahmin doğruluğunu artırabilir. Son olarak, orijinal veri setinden Cinsiyet (Gender) ve Hayatta Kalma Durumu (Survived) gibi kategorik değişkenler, oluşturulan yeni özelliklerle birlikte eklenmiştir. Bu kategorik değişkenler, kazaların hayatta kalma durumu üzerinde önemli etkiler yaratabilir. Dolayısıyla, modelin bu faktörleri de dikkate alarak tahminlerde bulunması sağlanmıştır.

Tüm bu adımların sonunda, özellik mühendisliği ile zenginleştirilmiş bir veri seti elde edilmiştir. Bu veri seti, kazaların sonuçlarını etkileyebilecek daha fazla bilgiye sahip olarak, tahmin modelinin doğruluğunu ve genelleme yeteneğini artırma amacını taşımaktadır. Bu yöntem, özellikle karmaşık ilişkilerin ve doğrusal olmayan etkilerin göz önünde bulundurulması gereken makine öğrenmesi uygulamaları için etkili bir strateji olarak öne çıkmaktadır.

Model kurulumu yapıldıktan sonra, farklı makine öğrenmesi algoritmaları ve derin sinir ağları kullanılarak model performansları detaylı bir şekilde değerlendirilmiştir. Model performansını ölçmek için doğruluk (accuracy), kesinlik (precision), duyarlılık (recall), F1 skoru (F1-score) ve ROC-AUC (Receiver Operating Characteristic - Area Under Curve) gibi temel sınıflandırma metrikleri kullanılmıştır. Bunlara ek olarak, algoritmaların hesaplama maliyetlerini değerlendirmek amacıyla işlem süresi (time taken) de analiz edilmiştir. Bu sayede, modellerin yalnızca doğruluğu değil, aynı zamanda verimliliği ve hesaplama süresi açısından da karşılaştırılması amaçlanmıştır.

Model eğitimi sırasında kullanılan veri setleri, ön işleme adımları uygulanarak temizlenmiş ve eksik veriler uygun yöntemlerle tamamlanmıştır. Kategorik değişkenler One-Hot Encoding veya Etiket Kodlama (Label Encoding) ile sayısal formata dönüştürülmüş, sürekli değişkenler ise Min-Max Scaling veya Standardizasyon ile ölçeklendirilmiştir. Bu işlemler, modelin daha kararlı ve genelleştirilebilir sonuçlar üretmesini sağlamak

amacıyla gerçekleştirilmiştir. Ayrıca, eğitim ve test veri kümelerinin dengeli dağılımı sağlanarak modelin aşırı öğrenme (overfitting) riskinin azaltılması hedeflenmiştir.

Çalışmada kullanılan algoritmalar arasında lojistik regresyon (Logistic Regression), karar ağaçları (Decision Trees), rastgele ormanlar (Random Forest), ve destek vektör makineleri (Support Vector Machines - SVM) yer almaktadır. Derin öğrenme tarafında ise Derin sinir ağları (Deep Neural Networks - DNN) gibi modeller test edilmiştir. Bu modeller farklı hiperparametre ayarlarıyla eğitilmiş ve sonuçlar karşılaştırmalı olarak analiz edilmiştir.

Son olarak, elde edilen metrikler ve işlem süresi değerleri karşılaştırılarak her bir modelin avantajları ve dezavantajları belirlenmiştir. Yüksek doğruluk oranı sağlayan modellerin işlem süresi bakımından daha maliyetli olduğu, bazı algoritmaların ise daha hızlı ancak düşük doğruluklu sonuçlar ürettiği gözlemlenmiştir. Bu analizler doğrultusunda, belirli uygulama senaryolarına en uygun model seçimi konusunda öneriler geliştirilmiştir.

3.3.1. Lojistik Regresyon

Lojistik regresyon, sınıflandırma problemlerinde yaygın olarak kullanılan doğrusal bir modeldir. Bu çalışmada, veri kümesindeki ilişkileri incelemek amacıyla lojistik regresyon analizi kullanılmıştır. Bu model, bağımsız değişkenlerin bağımlı değişkenin olasılığını nasıl etkilediğini tahmin etmemize olanak tanır. Lojistik regresyon modeli, bağımlı değişkenin başarı olasılığını

$$P(y=1|X)=\frac{1}{(1+e^{-z})} \quad (4)$$

şeklinde ifade eder. Bu dönüşüm, z 'yi 0 ile 1 arasında bir değere sıkıştırarak, gözlemin belirli bir sınıfa ait olma olasılığını verir. Burada doğrusal tahminci z , bağımsız değişkenlerin bir doğrusal kombinasyonudur ve $\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n$ şeklinde hesaplanır. Bu denklemde, β_0 sabit terim (intercept), $\beta_1, \beta_2, \dots, \beta_n$ modelin öğrenilen parametreleridir ve X kümesinden alınan yani veri kümesinden alınan $x_1, x_2, \dots, x_n$ bağımsız değişkenlerinin değerlerini ifade eder.

Lojistik regresyon modelinde, modelin tahmin performansı, doğruluk, duyarlılık, özgüllük, ROC eğrisi, AUC kullanılarak değerlendirilmiştir. Modelin katsayıları (β), bağımsız değişkenlerin bağımlı değişkenin olasılığı

üzerindeki etkisini gösterir. Pozitif katsayılar, bağımsız değişkenin değerindeki artışın bağımlı değişkenin başarı olasılığını artırdığını gösterirken, negatif katsayılar tersini gösterir. Katsayıların büyüklüğü, etkinin gücünü gösterir. Lojistik regresyon modelinin bazı varsayımları vardır: bağımlı değişkenin iki kategorili olması, bağımsız değişkenler ile bağımlı değişkenin logit dönüşümü arasında doğrusal bir ilişki olması, gözlemlerin birbirinden bağımsız olması ve çoklu doğrusallık olmaması (bağımsız değişkenler arasında yüksek korelasyon olmaması) (Hosmer vd., 2013).

Lojistik regresyon modelinde, optimal hiperparametre C 'nin belirlenmesi için bir arama işlemi gerçekleştirilmiştir. Bu işlemde, C değeri 0.001 ile 1000 arasında logaritmik bir şekilde değiştirilmiş ve model, eğitim verisi üzerinde 5 katlı K-Fold çapraz doğrulama (KFold(5)) kullanılarak değerlendirilmiştir. Her bir C değeri için çapraz doğrulama sonucu elde edilen F1 skoru hesaplanmış ve en yüksek ortalama F1 skoru ile en uygun C değeri belirlenmiştir. Sonuçlar, C değeri ile birlikte çapraz doğrulama skorları da sırasıyla yazdırılmıştır. En yüksek F1 skoruna sahip olan C değeri belirlenerek modelin performansı optimize edilmiştir (Pedregosa vd., 2011).

Son olarak, optimal C değeriyle Logistic Regression modeli yeniden eğitilmiş ve test verisi üzerinde modelin performansı, doğruluk, duyarlılık, özgüllük ve ROC eğrisi gibi metriklerle rapor edilmiştir. Zaman hesaplamaları da eklenerek modelin eğitilmesi ve değerlendirilmesi süresi kaydedilmiştir.

Lojistik regresyon modelinin bazı sınırlılıkları da bulunmaktadır. Öncelikle, model yalnızca iki kategorili bağımlı değişkenler için uygundur. Ayrıca, bağımsız değişkenlerle bağımlı değişken arasındaki ilişkinin doğrusal olduğu varsayılmaktadır. Bunun yanı sıra, lojistik regresyon modeli aykırı değerlere duyarlı olabilir, bu da modelin performansını olumsuz etkileyebilir.

Modelin sonuçları, bağımsız değişkenlerin bağımlı değişken üzerinde anlamlı bir etkiye sahip olduğunu göstermektedir. Bu bulgular, trafik kazalarında tahmin yaparken önemli çıkarımlar sunmakta ve söz konusu bağımsız değişkenlerin, kazaların oluşma olasılığı üzerindeki etkisini anlamamıza yardımcı olmaktadır.

3.3.2. k-En Yakın Komşu (kNN)

Bu çalışmada, veri kümesi üzerindeki sınıflandırma problemi için k-En Yakın Komşu (kNN) algoritması uygulanmıştır. kNN, veri noktalarını sınıflandırmak için en yakın komşuların sınıflarını kullanan basit ve etkili bir denetimli öğrenme yöntemidir. Algoritmanın başarısı, özellikle k parametresinin (komşu sayısı) doğru seçimine bağlıdır. Bu nedenle, modelin performansını en üst düzeye çıkarmak için, 1 ile 20 arasında değişen k

değerleri için 5 katlı çapraz doğrulama (KFold) kullanılarak en uygun k değeri belirlenmiştir. Çapraz doğrulama sürecinde, her bir k değeri için F1 skoru (sınıflandırma performansının bir ölçüsü) hesaplanmış ve en yüksek F1 skorunu veren k değeri optimal k olarak seçilmiştir. Bu optimizasyon süreci, modelin genelleme yeteneğini artırmayı ve aşırı uyum (overfitting) riskini azaltmayı amaçlamaktadır.

Optimizasyon sonucunda elde edilen optimal k değeri [optimal_k] olarak bulunmuştur. Bu değer kullanılarak, kNN modeli eğitim ve test verileri üzerinde tekrar eğitilmiştir. Modelin sınıflandırma performansı, sınıflandırma raporunu veren fonksiyon kullanılarak detaylı bir şekilde değerlendirilmiştir. Bu rapor, doğruluk(accuracy), kesinlik(precision), duyarlılık (recall) ve F1 skoru gibi metrikleri içermektedir. kNN algoritması, yeni bir veri noktasının sınıfını belirlemek için en yakın k komşusunun sınıflarını kullanır. Uzaklık hesaplaması genellikle Öklid uzaklığı ile yapılır:

$$d(x, y) = \sqrt{\sum (x_i - y_i)^2} \quad (5)$$

Burada $d(x, y)$, x ve y noktaları arasındaki uzaklığı, x_i ve y_i ise bu noktaların i . değerlerini temsil eder.

kNN algoritması trafik veri seti üzerindeki sınıflandırma problemi için başarılı bir şekilde uygulanmış ve optimal k parametresi kullanılarak modelin performansı artırılmıştır. Modelin sınıflandırma raporu, test verileri üzerinde tatmin edici sonuçlar verdiğini göstermektedir. Bu çalışma, kNN algoritmasının hiperparametre optimizasyonunun önemini ve modelin genel performansına olan etkisini vurgulamaktadır. Çapraz doğrulama gibi yöntemlerin kullanılması, modelin güvenilirliğini artırmakta ve gerçek dünya uygulamalarında daha iyi sonuçlar elde edilmesini sağlamaktadır (Cover ve Hart, 1967; Kuhn ve Johnson, 2013; Bishop, 2006; Hastie vd., 2009; Scikit-learn Documentation, 2025).

3.3.3. Karar Ağacı

Bu çalışmada, veri seti üzerinde sınıflandırma problemi için Karar Ağacı (Decision Tree) algoritması kullanılarak, veri odaklı karar verme süreçleri incelenmiştir. Karar ağaçları, özellikle karmaşık ve doğrusal olmayan ilişkileri modellemede etkili, denetimli öğrenme algoritmalarıdır (Breiman, 1986). Algoritma, veri setini özyinelemeli olarak alt kümelerle ayırarak, her bir alt küme için homojenlik kriterlerini maksimize etmeyi amaçlar. Bu süreçte, bilgi kazancı (information gain) veya Gini safsızlığı (Gini impurity) gibi metrikler kullanılarak optimal özellik seçimi yapılır. Bilgi kazancı, bir özelliğin veri setini ne kadar iyi ayırdığını ölçer ve şu şekilde ifade edilir:

$$IG(X, A) = H(X) - \sum \left[\left(\frac{|X_v|}{|X|} \right) * H(X_v) \right] \quad (6)$$

Burada $IG(X, A)$, IG (Information Gain) - Bilgi Kazancı, X veri setinde A özelliğinin bilgi kazancını, $H(X)$, X veri setinin entropisini, X_v , A özelliğinin v değerine sahip alt kümesini ve $| \cdot |$ kümenin eleman sayısını temsil eder. Entropi, veri setindeki sınıf dağılımının homojenliğini ölçer ve şu şekilde hesaplanır:

$$H(X) = - \sum \left[p_i * \log_2(p_i) \right] \quad (7)$$

Burada p_i , X veri setindeki i. sınıfının olasılığını temsil eder.

Modelin genelleme yeteneği, ağacın karmaşıklığını kontrol eden hiperparametrelerin doğru seçimine kritik derecede bağlıdır. Bu bağlamda, modelin aşırı uyumunu (overfitting) önlemek ve genelleme performansını artırmak amacıyla, k-katlı çapraz doğrulama (k-fold cross-validation) yöntemi kullanılarak hiperparametre optimizasyonu gerçekleştirilmiştir (Kohavi, 1995). Çapraz doğrulama, veri setini k eşit parçaya bölerek her bir parçayı test, diğerlerini eğitim verisi olarak kullanır ve bu işlemi k kez tekrarlar. Bu sayede, modelin farklı veri alt kümeleri üzerindeki performansı değerlendirilerek daha güvenilir sonuçlar elde edilir. Her bir hiperparametre kombinasyonu için sınıflandırma performansının bir ölçüsü olarak F1 skoru veya AUC-ROC hesaplanmış ve en iyi performansı veren kombinasyon optimal değer olarak belirlenmiştir.

Optimizasyon sonucunda elde edilen optimal hiperparametre değerleri kullanılarak, Karar Ağacı modeli eğitim ve test verileri üzerinde tekrar eğitilmiştir. Modelin sınıflandırma performansı, doğruluk (accuracy), kesinlik (precision), duyarlılık (recall), F1 skoru ve AUC-ROC gibi metriklerle değerlendirilmiştir (Powers, 2011). Karar ağaçları, verileri özellik değerlerine göre dallara ayırarak karar verir. Ağacın derinliği arttıkça, model daha karmaşık hale gelir ve eğitim verilerine aşırı uyum sağlayabilir. Bu nedenle, modelin karmaşıklığını kontrol eden hiperparametrelerin doğru seçimi modelin performansı için kritik öneme sahiptir.

Sonuç olarak, Karar Ağacı algoritması, veri seti üzerindeki sınıflandırma problemi için başarılı bir şekilde uygulanmış ve hiperparametre optimizasyonu ile modelin performansı artırılmıştır. Modelin değerlendirme metrikleri, test verileri üzerinde tatmin edici sonuçlar verdiğini göstermektedir. Bu çalışma, Karar Ağacı algoritmasının hiperparametre optimizasyonunun önemini ve modelin genel performansına olan etkisini vurgulamaktadır. Çapraz doğrulama ve bilgi teorisi tabanlı metriklerin kullanılması, modelin

güvenilirliğini artırmakta ve gerçek dünya uygulamalarında daha iyi sonuçlar elde edilmesini sağlamaktadır.

3.3.4. Rastgele Orman

Birden fazla karar ağacının çıktılarının ortalamasını alarak tahmin yapan Rastgele Orman modeli, sınıflandırma problemi için uygulanmıştır. Rastgele Orman, birden fazla karar ağacının bir araya gelerek oluşturduğu, topluluk öğrenme yöntemlerinden biridir. Bu algoritma, karar ağaçlarının aşırı uyum sorununu azaltarak daha sağlam ve genellenebilir modeller oluşturmayı amaçlar. Algoritmanın performansı, modelin karmaşıklığını ve hesaplama maliyetini doğrudan etkileyen bir parametrenin doğru seçimine bağlıdır. Bu nedenle, modelin en iyi performansını elde etmek için, bu parametre farklı değerler için çapraz doğrulama yöntemiyle optimize edilmiştir. Çapraz doğrulama, veri setini farklı parçalara bölerek her bir parçayı test, diğerlerini eğitim verisi olarak kullanır ve bu işlemi tekrarlar. Bu sayede modelin genelleme yeteneği daha güvenilir bir şekilde değerlendirilir. Her bir parametre değeri için sınıflandırma performansının bir ölçüsü hesaplanmış ve en yüksek performansı veren parametre değeri optimal değer olarak belirlenmiştir (Breiman, 2001).

Optimizasyon sonucunda elde edilen optimal parametre değeri kullanılarak, Rastgele Orman modeli eğitim ve test verileri üzerinde tekrar eğitilmiştir. Modelin sınıflandırma performansı, detaylı değerlendirme fonksiyonları ile analiz edilmiştir. Rastgele Orman algoritması, her bir karar ağacını farklı bir rastgele alt küme üzerinde eğiterek ve her ağaçta rastgele seçilmiş özellikler kullanarak çeşitlilik oluşturur. Bu çeşitlilik, modelin genelleme yeteneğini artırır ve aşırı uyum riskini azaltır. Her bir ağaç, tahminini yapar ve sınıflandırma sonucu, ağaçların çoğunluğunun verdiği karara göre belirlenir. Bu topluluk yaklaşımı, tek bir karar ağacına göre daha sağlam ve güvenilir sonuçlar üretir (Liaw ve Wiener, 2002).

Rastgele Orman'da, her bir ağaç en iyi özelliği seçmek için bilgi kazancı (information gain) veya Gini safsızlığı (Gini impurity) gibi metrikleri kullanır. Bilgi kazancı, bir özelliğin veri setini ne kadar iyi ayırdığını ölçer ve entropi kavramı üzerinden hesaplanır. Entropi, veri setindeki sınıf dağılımının homojenliğini ifade eder ve bu metrikler, her bir düğümde en iyi kararın verilmesini sağlar (Quinlan, 1986).

Topluluk öğrenme yaklaşımı, her bir karar ağacının tahminlerini birleştirerek nihai tahmini yapar. Sınıflandırma problemlerinde, her ağaç bir sınıf tahmini yapar ve nihai tahmin, ağaçların çoğunluğunun verdiği karara göre belirlenir. Regresyon problemlerinde ise ağaçların tahminlerinin

ortalaması alınır. Bu oylama veya ortalama alma süreci, modelin daha kararlı ve güvenilir tahminler yapmasını sağlar (Breiman, 2001).

Rastgele Orman, her bir karar ağacını farklı bir rastgele alt küme üzerinde eğiterek çeşitlilik oluşturur. Bu sürece bootstrap aggregating (bagging) denir. Bagging, veri setinden rastgele örnekler seçerek farklı alt kümeler oluşturur. Ayrıca, her bir ağaçta rastgele seçilmiş özellikler kullanılır. Bu rastgelelik, ağaçlar arasındaki korelasyonu azaltır ve modelin genelleme yeteneğini artırır (Breiman, 1996).

Modelin performansı, çapraz doğrulama (cross-validation) yöntemiyle değerlendirilir. Sınıflandırma problemlerinde, doğruluk(accuracy), kesinlik(precision), duyarlılık (recall), F1 skoru ve ROC eğrileri, AUC hesabı gibi metrikler kullanılır. Bu değerlendirme metrikleri, modelin ne kadar başarılı olduğunu sayısal olarak ifade eder (Kohavi, 1995).

Sonuç olarak, Rastgele Orman algoritması başarılı bir şekilde uygulanmış ve optimal parametre değeri kullanılarak modelin performansı artırılmıştır. Modelin değerlendirme metrikleri, test verileri üzerinde tatmin edici sonuçlar verdiğini göstermektedir. Bu çalışma, Rastgele Orman algoritmasının parametre optimizasyonunun önemini ve modelin genel performansına olan etkisini vurgulamaktadır. Çapraz doğrulama gibi yöntemlerin kullanılması, modelin güvenilirliğini artırmakta ve gerçek dünya uygulamalarında daha iyi sonuçlar elde edilmesini sağlamaktadır (Liaw & Wiener, 2002).

3.3.5. Derin Sinir Ağları (Deep Neural Networks, DeepNN)

Bu çalışmada, ikili sınıflandırma problemi için Derin Sinir Ağları (Deep Neural Networks, DeepNN) kullanılarak veri odaklı modelleme gerçekleştirilmiştir. Derin Sinir Ağları, çok sayıda gizli katmana sahip yapay sinir ağı modelleri olup, karmaşık desenleri ve ilişkileri öğrenme yeteneğiyle öne çıkar. Algoritmada, özel bir sınıf oluşturularak, TensorFlow ve Keras kütüphaneleri kullanılarak bir derin sinir ağı modeli tanımlanmıştır. Bu model, girdi katmanından sonra 5 adet gizli katman içermekte ve her bir gizli katmanında ReLU aktivasyon fonksiyonu kullanmaktadır. Çıktı katmanında ise sigmoid aktivasyon fonksiyonu kullanılarak ikili sınıflandırma problemi çözülmüştür. Model, Adam optimizasyon algoritması ve binary cross-entropy hata fonksiyonu ile derlenmiştir (Kingma ve Ba, 2015).

Modelin eğitimi, fit metodu aracılığıyla gerçekleştirilmiştir. Bu metot, girdi verileri (X) ve hedef değişkenleri (y) kullanarak ağı eğitir Chollet (2017). Eğitim süreci, belirli sayıda epoch (tekrar sayısı) ve batch size (toplu işleme boyutu) parametreleri ile kontrol edilmiştir. Modelin performansı, eğitim ve test verileri üzerinde ayrı ayrı değerlendirilmiştir. Eğitim verileri

üzerinde modelin doğruluğu (accuracy) hesaplanırken, test verileri üzerinde de aynı şekilde doğruluk (accuracy), duyarlılık (precision), geri çağırma (recall), F1 skoru, AUC-ROC ve log-loss gibi metrikler hesaplanmıştır (Sokolova ve Lapalme, 2009).

Modelin çalışma yapısı, her bir nöronun çıktısının bir aktivasyon fonksiyonu ile hesaplanmasına dayanır. Bir nöronun çıktısı şu şekilde ifade edilebilir:

$$y = f\left(\sum(w_i * x_i) + b\right) \quad (8)$$

Burada y , nöronun çıktısını, f , aktivasyon fonksiyonunu; w_i , girdilerin ağırlıklarını, x_i , girdileri ve b , bias terimini temsil eder. Aktivasyon fonksiyonları, ağırlık doğrusal olmayan ilişkileri modellemesini sağlar. Ağırlıkların güncellenmesi ise gradyan inişi algoritması ile gerçekleştirilmiştir ve şu şekilde ifade edilir:

$$w_i = w_i - \alpha * \partial L / \partial w_i \quad (9)$$

Burada α , öğrenme oranını; L , hata fonksiyonunu; ve $\partial L / \partial w_i$, L 'nin w_i 'ye göre türevini temsil eder.

Derin Sinir Ağları, yapay zeka ve makine öğrenmesi alanlarında güçlü bir modelleme aracıdır ve özellikle karmaşık veri setlerinde yüksek doğruluk oranları sağlayabilmektedir. Bu çalışmada, derin sinir ağı modelinin performansı, Python programlama dili ve Keras kütüphanesi kullanılarak değerlendirilmiştir. Modelin temel yapısı, birden fazla gizli katman içeren bir yapıyı benimsemektedir. Derin Sinir Ağı modelinde kullanılan katmanlar, verinin daha derin ve soyut özelliklerini öğrenebilmek için çeşitli aktivasyon fonksiyonları (ReLU ve sigmoid) kullanarak veri setinin özelliklerini sırasıyla dönüştürür. Bu çalışmada, modelin giriş katmanına 256 nöron eklenmiş, ardından sırasıyla 128, 64, 32 ve 16 nöronlu katmanlar eklenmiştir. Modelin çıktısı, ikili sınıflandırma problemini çözebilmek amacıyla sigmoid aktivasyon fonksiyonu ile elde edilmiştir.

Modelin eğitilmesinde, eğitim verisi üzerinde 100 epoch boyunca eğitim yapılmıştır. Eğitim sırasında, her bir veri örneği, modelin parametrelerini optimize etmek amacıyla Adam optimizasyon algoritmasıyla güncellenmiştir. Eğitim süreci, her bir epoch sonunda modelin doğruluğunu ve kayıp fonksiyonunu izleyerek gerçekleştirilmiştir. Eğitimin ardından, modelin performansı test verisi üzerinde değerlendirilmiş ve test doğruluğu hesaplanmıştır.

Modelin başarı metrikleri arasında, doğruluk (accuracy), kesinlik (precision), duyarlılık (recall), F1 skoru ve ROC eğrisi-AUC (Area Under

the Curve) gibi önemli performans göstergeleri yer almaktadır (Hand ve Till, 2001). F1 skoru, modelin sınıflandırma başarımını tek bir metrikle değerlendirebilmek için kullanılırken, AUC değeri, modelin sınıflandırma doğruluğunu daha kapsamlı bir şekilde ölçmek için önemli bir araçtır. Bu çalışmada elde edilen AUC skoru, modelin ikili sınıflandırmada ne kadar etkili olduğunu gösterirken modelin tahmin doğruluğunu sayısal bir şekilde ifade etmektedir.

Sonuç olarak, derin sinir ağları, özellikle yüksek doğruluk gerektiren sınıflandırma problemlerinde başarılı bir çözüm sunmaktadır (Goodfellow vd., 2016). Bu model, karmaşık verilerdeki soyut ilişkileri öğrenme yeteneği ile geleneksel makine öğrenmesi yöntemlerine kıyasla daha üstün performans gösterebilmektedir. Ancak, modelin doğruluğu ve genel performansı, kullanılan verinin kalitesine ve modelin hiperparametrelerinin doğru bir şekilde optimize edilmesine bağlıdır. Bu çalışmada, hiperparametrelerin belirlenmesi için çeşitli optimizasyon teknikleri kullanılmakta olup, elde edilen sonuçlar, derin sinir ağlarının uygulanabilirliğini ve etkinliğini desteklemektedir.

Sonuç olarak, Derin Sinir Ağı modeli başarıyla eğitilmiş ve test verileri üzerinde tatmin edici sonuçlar elde edilmiştir. Modelin değerlendirme metrikleri, modelin genelleme yeteneğinin yüksek olduğunu göstermektedir. Bu çalışma, Derin Sinir Ağlarının karmaşık sınıflandırma problemlerinde kullanılabileceğini ve veri odaklı karar verme süreçlerinde önemli bir rol oynayabileceğini göstermektedir. Bu sonuçlar, farklı makine öğrenmesi algoritmalarının ve Derin Sinir Ağlarının veri setine göre farklı performans sergileyebileceğini ortaya koymaktadır. Model seçiminde yalnızca doğruluk değil, diğer metrikler ve işlem süresi gibi faktörlerin de göz önünde bulundurulması gerekmektedir.

4. Bulgular

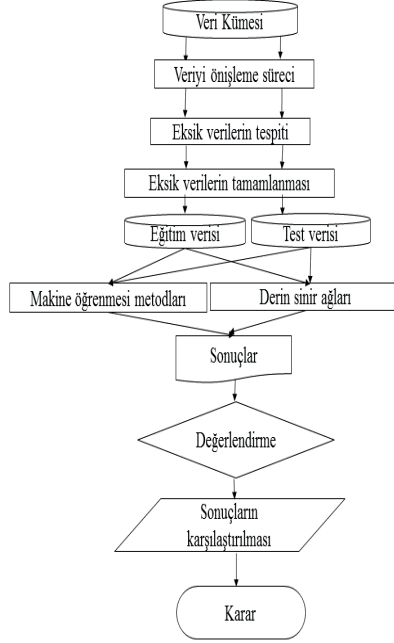
Bu çalışmada, trafik kazalarında hayatta kalma oranlarını tahmin etmek amacıyla farklı makine öğrenmesi modelleri ve Derin Sinir Ağları incelenmiş, modellerin performansları doğruluk(accuracy), kesinlik(precision), duyarlılık(recall) ve F1 skoru gibi metrikler açısından değerlendirilmiştir. Modellerin çalıştırılmasıyla elde edilen metrikler Tablo 2 de sunulmuştur. Modellerin ROC eğrileri-AUC grafiği Şekil 4 de verilmektedir. Elde edilen sonuçlar, kullanılan modeller arasında belirgin farklılıklar olduğunu göstermektedir. En yüksek doğruluk oranı %90,5 ile Rastgele Orman (Random Forest) modeli tarafından elde edilmiştir. Rastgele Orman

modeli, aynı zamanda %90,2 duyarlılık ve %90,99 F1 skoru ile en dengeli performansı sergileyen model olmuştur.

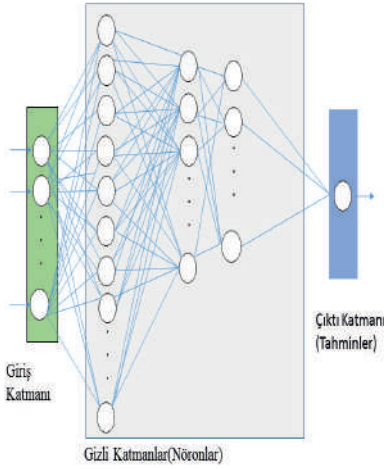
Karar Ağaçları (Decision Tree) modeli, %87,75 doğruluk oranı ile ikinci en iyi performansı göstermiştir. Bu modelin kesinlik ve duyarlılık değerleri de sırasıyla %89,2 ve %89,7 olarak hesaplanmıştır. Tek bir karar ağacı kullanan bu model, Rastgele Orman'a kıyasla biraz daha düşük performans göstermiş olsa da, hesaplama süresi açısından daha hızlı olduğu gözlemlenmiştir. k-En Yakın Komşular (KNN) algoritması ise %65,5 doğruluk oranı ile orta seviyede bir performans sergilemiştir. Özellikle yüksek değişkenlik gösteren veri setlerinde daha az başarılı olduğu ve işlem süresinin nispeten uzun olduğu belirlenmiştir.

Lojistik Regresyon modeli %54,75 doğruluk oranı ile en düşük performansı sergileyen model olmuştur. Özellikle doğrusal ayrılabilir olmayan veriler üzerinde düşük başarı göstermesi, bu modelin trafik kazalarında hayatta kalma tahminleri için yeterince uygun olmadığını düşündürmektedir. Derin Sinir Ağları (Deep Neural Networks) modeli ise %62 doğruluk oranı ile beklenenden düşük bir performans göstermiştir. Bu durum, modelin aşırı öğrenme (overfitting) veya yetersiz öğrenme (underfitting) problemleri yaşadığını düşündürmektedir. Modelin performansını artırmak için hiperparametre optimizasyonu ve daha büyük veri setleriyle eğitiminin yapılması gerekmektedir.

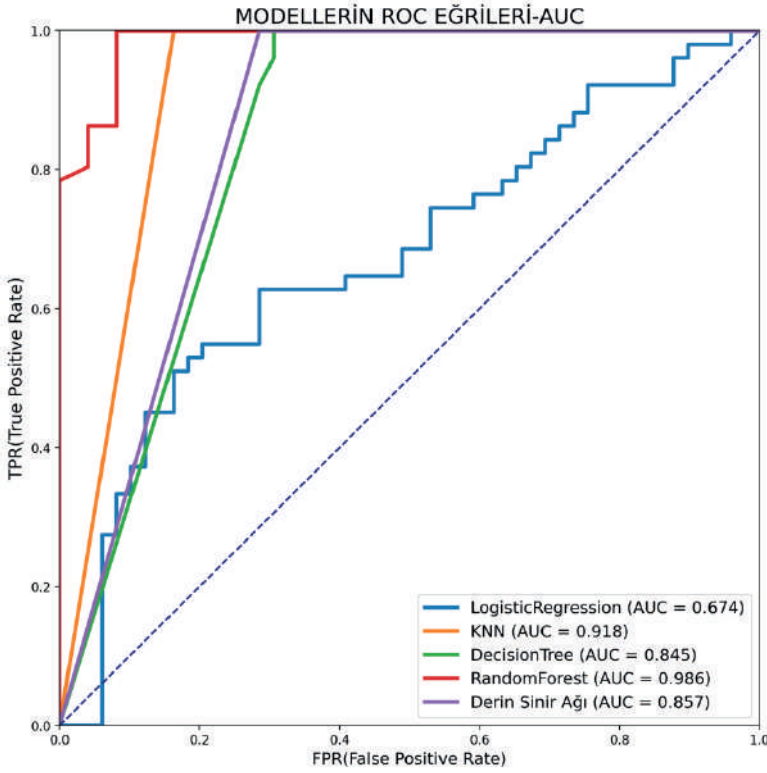
Genel olarak, ağaç tabanlı modellerin (Karar Ağaçları ve Rastgele Orman) diğer yöntemlere göre daha başarılı olduğu görülmüştür. Özellikle Rastgele Orman modeli, tüm performans metriklerinde en iyi sonuçları sunarak, trafik kazalarında hayatta kalma oranlarının tahmini için en uygun model olarak öne çıkmaktadır. Ancak, modelin işlem süresi ve karmaşıklığı dikkate alındığında, daha hızlı ancak nispeten düşük doğruluk oranına sahip olan Karar Ağaçları gibi modellerin de tercih edilebileceği sonucuna varılmıştır.



Şekil 2. Veri önleme ve veri analizi süreçlerinin gösterimi



Şekil 3. Derin sinir ağları



Şekil 4. ROC eğrileri ve AUC değerleri

Tablo 2. Algoritmalarından elde edilen sonuçlar

Model	Doğruluk (Accuracy)	Kesinlik (Precision)	Duyarlılık (Recall)	F1 Skoru (F1 Score)	Geçen Süre (Time Taken)
Rastgele Orman	0,9050	0,897952	0,902073	0,909905	38,184784
Karar Ağaçları	0,8775	0,892689	0,897317	0,894149	0,632241
k-En Yakın Komşular	0,6550	0,672424	0,652073	0,660337	0,928479
Derin Sinir Ağı	0,6200	0,640050	0,681341	0,636005	165,2073
Lojistik Regresyon	0,5475	0,554645	0,587683	0,569715	15,173153

5. Sonuç

Bu çalışmada, trafik kazalarında hayatta kalma oranlarını tahmin etmek amacıyla trafik verileri üzerinden çeşitli makine öğrenmesi modellerinin performansını kapsamlı bir şekilde değerlendirmiştir. Elde edilen sonuçlar, modeller arasında belirgin farklılıklar olduğunu ortaya koymaktadır. Özellikle Rastgele Orman modeli, yüksek doğruluk oranı (%90,5) ve düşük

işlem süresiyle öne çıkarken, ROC eğrileri de bu modelin diğer yöntemlere göre daha üstün sınıflandırma performansı sunduğunu göstermiştir. Karar Ağaçları modeli ise nispeten daha düşük doğruluk oranına sahip olmasına rağmen, işlem süresi açısından daha hızlı olması nedeniyle avantaj sağlamaktadır.

Derin Sinir Ağları, görsel sınıflandırma alanında elde ettikleri üstün performanslarla öne çıkmaktadır. Bu modeller, görüntülerdeki karmaşık desenleri ve yapıları otomatik olarak öğrenme kabiliyetleriyle, görüntü sınıflandırmada geleneksel yöntemlere kıyasla daha doğru ve verimli sonuçlar sunmaktadır. Ancak, bu çalışmada kullanılan Derin Sinir Ağları modeli beklenen performansı sağlayamamış, hem doğruluk hem de işlem süresi açısından daha düşük sonuçlar üretmiştir. Bu durum, modelin aşırı öğrenme (overfitting) veya yetersiz öğrenme (underfitting) problemleri yaşadığını göstermektedir. Gelecekte, modelin hiperparametre optimizasyonu yapılarak ve daha büyük veri setleri ile eğitilerek performansının artırılması sağlanabilir.

Genel olarak, ağaç tabanlı modellerin (Karar Ağaçları ve Rastgele Orman) diğer yöntemlere göre daha başarılı olduğu görülmüştür. Özellikle Rastgele Orman modeli, tüm performans metriklerinde en iyi sonuçları sunarak, trafik kazalarında hayatta kalma oranlarının tahmini için en uygun model olarak öne çıkmaktadır. Bununla birlikte, modelin işlem süresi ve hesaplama karmaşıklığı dikkate alındığında, nispeten daha düşük doğruluk oranına sahip ancak daha hızlı olan Karar Ağaçları gibi modellerin de tercih edilebileceği gözlemlenmiştir.

Bu bağlamda, çalışmanın bulguları, trafik kazalarında hayatta kalma oranlarını tahmin etme sürecinde daha karmaşık modellerin her zaman daha iyi sonuç vermediğini, veri yapısı ve hesaplama verimliliği gibi faktörlerin de göz önünde bulundurulması gerektiğini göstermektedir. Özellikle büyük ölçekli veri setleri kullanıldığında, hesaplama maliyetleri ve modelin genelleme yeteneği gibi faktörlerin kritik olduğu görülmektedir. Daha karmaşık modellerin her zaman en iyi performansı sunmaması, veri özelliklerine uygun model seçiminin önemini vurgulamaktadır.

Gelecekteki çalışmalar, farklı veri setleri ve ek parametreler kullanarak modellerin genel geçerliğini artırmaya yönelik kapsamlı analizlere odaklanabilir. Özellikle, değişken seçim yöntemleri ve veri ön işleme tekniklerinin model başarısına etkisi araştırılabilir. Ayrıca, derin öğrenme modellerinin performansını artırmak amacıyla transfer öğrenme ve ileri seviye optimizasyon tekniklerinin uygulanması faydalı olabilir. Son olarak, gerçek

zamanlı tahmin sistemleri geliştirmek için donanım destekli hızlandırıcıların kullanılması, modellerin pratik kullanım alanlarını genişletebilir.

Veri Kümesi Kaynağı ve Kullanılabilirliği

Bu çalışma, Kaggle platformunda yer alan Road Accident Survival Dataset veri kümesi kullanılarak gerçekleştirilmiştir. Veri kümesi, trafik kazalarında hayatta kalma olasılıklarını etkileyen demografik, davranışsal ve durumsal faktörler gibi çeşitli etmenler hakkında kapsamlı bilgiler sunmakta ve bu etmenlerin kazalarda hayatta kalma şansına nasıl etki edebileceği konusunda derinlemesine analiz yapılmasına imkân tanımaktadır. Veri kümesi, CC0: Public Domain lisansı altında yayımlanmaktadır; bu, verilerin telif hakkı kısıtlamalarından arındırıldığı ve herkese ücretsiz olarak kullanım hakkı verildiği anlamına gelmektedir. Kişisel veri içermemekte olup, kullanımına yönelik herhangi bir kısıtlama bulunmamaktadır. Bu nedenle, veriler akademik çalışmalar, araştırmalar ve diğer bilimsel analizler için serbestçe kullanılabilir. Veri kümesine ulaşmak için Kaynakça da verilen Kaggle web sitesinde datanın yayınlandığı web adresi için bağlantısı üzerinden erişim sağlanabilir.

Kaynakça

- Ahmed, S., Hossain, M. A., Bhuiyan, M. M. I., & Ray, S. K. (2021, December). A comparative study of machine learning algorithms to predict road accident severity. In 2021 20th International Conference on Ubiquitous Computing and Communications (IUCC/CIT/DSCI/SmartCNS) (pp. 390-397).
- Bai, M. L., Pamula, R., Subbarao, K., & Bharathi, S. (2025). Recommended System for Predicting Traffic Accident Costs using Enhanced Machine Learning Techniques. *Journal of Electrical Engineering & Technology*, 1-12.
- Ballamudi, K. R. (2019). Road accident analysis and prediction using machine learning algorithmic approaches. *Asian Journal of Humanity, Art and Literature*, 6(2), 185-192.
- Behboudi, N., Moosavi, S., & Ramnath, R. (2024). Recent advances in traffic accident analysis and prediction: a comprehensive review of machine learning techniques. *arXiv preprint arXiv:2406.13968*.
- Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
- Breiman, L. (1986). *Classification and regression trees*. Wadsworth & Brooks/Cole Advanced Books & Software.
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2), 123-140.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32.
- Chollet, F. (2017). *Deep learning with Python*. Manning Publications.
- Cover, T. M., & Hart, P. E. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21-27.
- Dünya Sağlık Örgütü (WHO). (2004). World report on road traffic injury prevention. <https://www.who.int/publications/i/item/9241562609>
- Dünya Sağlık Örgütü (WHO). (2023). Global status report on road safety 2023. <https://www.who.int/publications/i/item/9789240086517>
- Gibbons, J. D., & Chakraborti, S. (2014). *Nonparametric statistical inference: revised and expanded*. CRC press.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer. <https://doi.org/10.1007/978-0-387-84858-7>
- Hand, D. J., & Till, R. J. (2001). A simple generalisation of the area under the ROC curve for multiple class classification problems. *Machine Learning*, 45(2), 171-186. <https://doi.org/10.1023/A:1010920819831>
- Hosmer, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (3rd ed.). Wiley.

- Kendall, M. G. (1948). Rank correlation methods. Oxford University Press.
- Kingma, D. P. & Ba, J. (2015). Adam: A method for stochastic optimization. In *Proceedings of the 3rd International Conference on Learning Representations (ICLR)*.
- Kodepogu, K., Manjeti, V. B., & Siriki, A. B. (2023). Machine learning for road accident severity prediction. *Mechatron. Intell. Transp. Syst*, 2, 211-226.
- Kohavi, R. (1995). A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the 14th International Joint Conference on Artificial Intelligence*, 1137–1143.
- Kuhn, M., & Johnson, K. (2013). Applied predictive modeling. Springer. <https://doi.org/10.1007/978-1-4614-6849-3>
- Kumar, B., Basit, A., Kiruba, M. B., Giridharan, R., & Keerthana, S. M. (2021, July). Road accident detection using machine learning. In *2021 International Conference on System, Computation, Automation and Networking (ICSCAN)* (pp. 1-5).
- Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news*, 2(3), 18-22.
- Lin, D. J., Chen, M. Y., Chiang, H. S., & Sharma, P. K. (2021). Intelligent traffic accident prediction model for Internet of Vehicles with deep learning approach. *IEEE transactions on intelligent transportation systems*, 23(3), 2340-2349.
- Ma, Z., Mei, G., & Cuomo, S. (2021). An analytic framework using deep learning for prediction of traffic accident injury severity based on contributing factors. *Accident Analysis & Prevention*, 160, 106322. <https://doi.org/10.1016/j.aap.2021.106322>
- Nandurge, P. A., & Dharwadkar, N. V. (2017, February). Analyzing road accident data using machine learning paradigms. In *2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud) (I-SMAC)* (pp. 604-610).
- Obasi, I. C., & Benson, C. (2023). Evaluating the effectiveness of machine learning techniques in forecasting the severity of traffic accidents. *Heliyon*, 9(8), e17875. <https://doi.org/10.1016/j.heliyon.2023.e17875>
- Patil, J., Prabhu, M., Walavalkar, D., & Lobo, V. B. (2020, December). Road accident analysis using machine learning. In *2020 IEEE Pune Section International Conference (PuneCon)* (pp. 108-112).
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O & Duchesnay, É. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12, 2825-2830.
- Powers, D. M. W. (2011). Evaluation: From precision, recall and F-measure to ROC, informedness, markedness & correlation. *Journal of Machine Learning Technologies*, 2(1), 37–63.

- Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1), 81–106.
- Santos, D., Saias, J., Quaresma, P., & Nogueira, V. B. (2021). Machine learning approaches to traffic accident analysis and hotspot prediction. *Computers*, 10(12), 157.
- Sarder, H. (2025). Road accident survival dataset. Kaggle. <https://www.kaggle.com/datasets/himelsarder/road-accident-survival-dataset>
- Singh, G., Pal, M., Yadav, Y., & Singla, T. (2020). Deep neural network-based predictive modeling of road accidents. *Neural Computing and Applications*, 32, 12417–12426.
- Sokolova, M., & Lapalme, G. (2009). A systematic analysis of performance measures for classification tasks. *Information Processing & Management*, 45(4), 427–437. <https://doi.org/10.1016/j.ipm.2009.03.002>
- Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 72–101.
- Theofilatos, A., Chen, C., & Antoniou, C. (2019). Comparing machine learning and deep learning methods for real-time crash prediction. *Transportation research record*, 2673(8), 169–178.
- Vasavi, S. (2018). Extracting hidden patterns within road accident data using machine learning techniques. In *Information and Communication Technology: Proceedings of ICICT 2016* (pp. 13–22). Springer Singapore.
- Vicent, J. F., Curado, M., Oliver, J. L., & Pérez-Sala, L. (2025). A novel approach to predict the traffic accident assistance based on deep learning. *Neural Computing and Applications*, 37(7), 5343–5368.

Tüketicilerin Özel Günlerde Satın Alma Davranışlarının Rastgele Orman Yöntemiyle İncelenmesi

Mert Ersen¹

Melik Masarifoğlu²

Kemal Cem Söylemez³

Özet

Bu çalışma, özel günlerin perakende satışlar üzerindeki etkisini analiz ederek, firmaların stok ve üretim planlamalarında karşılaştıkları belirsizlikleri azaltmak amacıyla makine öğrenmesi temelli bir satış tahmin modeli geliştirmeyi amaçlamaktadır. Özellikle Yılbaşı, Kara Cuma ve Noel gibi dönemlerde yaşanan ani talep artışları, doğru tahmin modelleri kullanılmadığında işletmeler için operasyonel riskler doğurabilmektedir. Bu kapsamda, Amazon'un OLED televizyon satış verileri kullanılarak Rastgele Orman algoritmasıyla bir tahmin modeli oluşturulmuş ve bu modelin özel günlerdeki satış davranışlarını anlamada nasıl kullanılabileceği gösterilmiştir. Çalışmanın bulguları, makine öğrenmesi tabanlı modellerin satış tahmin süreçlerinde etkin bir şekilde kullanılabileceğini ortaya koymakta ve işletmelerin veri temelli karar alma süreçlerine katkı sunmaktadır.

- 1 Dr. Bağımsız Araştırmacı, mert_9034@hotmail.com, ORCID ID: <https://orcid.org/0000-0001-5643-4690>
- 2 Dr. Bağımsız Araştırmacı, memasarif@hotmail.com, ORCID ID: <https://orcid.org/0000-0001-8910-0486>
- 3 Dr. Bağımsız Araştırmacı, kemalcemsoylemez@gmail.com, ORCID ID: <https://orcid.org/0000-0003-1607-7208>

1. Giriş

Satış promosyonu, bir üreticinin tüketicileri belirli bir markayı satın almaya yönlendirmek ve satış ekibini daha agresif bir şekilde satış yapmaya teşvik etmek amacıyla sunduğu her türlü teşvik olarak tanımlanabilir (Yemez & Sağır, 2021). Bu promosyonlar, tüketicileri alım yapmaya motive etmek için pazarlama stratejilerinin önemli bir parçasını oluşturur ve bir ürünün gelirene önemli ölçüde katkı sağlayabilir (Shimp, 2003). Bu bağlamda, tüketiciler, daha az harcama yapma veya akıllı bir alışveriş yapan birey olma duygusunu yaşamak amacıyla promosyonlu fiyatlara olumlu tepki verirler (Lennon vd., 2011).

Tüketicilerin ilgisini çekmek ve reklamı yapılan ürünlere yönlendirmek amacıyla, “Süper Fırsat”, “Özel Gün İndirimi”, “Büyük Tasarruf” gibi promosyon ifadeleri sıklıkla kullanılmaktadır. Bu tür ifadeler, aynı zamanda tüketicilerin teklife ilişkin algılarını şekillendirmeye yönelik de etkili bir araçtır. Pazarlamacılar, tüketicilerin satın alma davranışlarını artırmak için genellikle kısa süreli ve aralıklı kampanyalar düzenlerler. Bu tür teklifler, genellikle yıl boyunca sürekli olarak sunulmaz ve belirli özel günlerde (anneler günü, babalar günü, yılbaşı, sevgililer günü, bayramlar vb.) daha yoğun olarak karşımıza çıkar. Bu strateji, yeni müşteri kazanımını artırırken, mevcut müşterilerin de daha fazla alışveriş yapmalarına yol açar (Karaca vd., 2018).

Promosyonlar hem fiyatları düşürerek hem de duygusal etkiler yoluyla tüketicilerin ürün satın alırken elde edeceği ekonomik faydayı ciddi şekilde değiştirebilir (Raghubir vd., 2004). Promosyon teklifleri hem parasal hem de parasal olmayan kazançlar ve kayıplar içerebilir. Örneğin, indirimli bir ürün almak maddi bir kazanç sağlarken, tüketici daha az para harcarsa da, o ürünü bulmak ve almak için harcanan zaman ve gösterilen çaba maddi olmayan bir kayıp oluşturur (Fogel ve Thornton, 2008). Ayrıca, promosyonlar, bilinçli alışveriş yapan tüketicilerde olumlu duygular uyandırırken, uygun olmayan zamanlarda alışveriş yapmak gibi maddi olmayan kayıplara da yol açabilir. Bu tür “iki ucu keskin” etkiler, özellikle Black Friday gibi etkinliklerde daha belirgindir. Örneğin, sabahın erken saatlerinde alışveriş yapmak ya da stokların tükenmesi nedeniyle promosyonlu ürünleri bulamamak sinir bozucu olabilir (Lennon vd., 2011). Bu yüzden, indirim peşinde koşan tüketiciler çoğu zaman beklediklerinden fazla çaba harcar ve sonunda kazandıkları paradan çok, yitirdikleri zaman ve enerjinin etkisini hisseder. Özellikle uzun kuyruklar, tıklım tıklım mağazalar ya da çevrim içi platformlarda yaşanan yoğunluk, promosyonun cazibesini gölgede bırakabilir.

Tüketici davranışının altında yatan nedenlerin anlaşılması, işletmeler için son derece önemli olup, bu konuda çeşitli modeller geliştirilmiştir. Farklı disiplinler, tüketicilerin bir ürün ya da hizmet satın alma kararı verirken hangi faktörlerin etkili olduğunu açıklamaya çalışmışlardır. Örneğin, iktisat alanına göre tüketiciler, satın alma kararlarında faydalarını maksimize etmeyi hedefler. Ancak bu açıklama, psikoloji, sosyoloji ve antropoloji gibi disiplinler tarafından genellikle kabul edilmez. Çünkü bu disiplinler, bireyin deneyimleri, motivasyonları, algıları, kişilik özellikleri, sosyal grupları, aile yapısı ve inançları gibi faktörlerin göz ardı edildiğini belirtir. Oysa, tüketicinin karar verme sürecinde bu özelliklerin ve aidiyetlerinin de önemli bir rolü vardır (Mucuk, 1998). Pazarlama ve satış faaliyetleri de bu nedenle yalnızca ekonomik gerekçelere değil, tüketicinin davranışsal ve psikolojik özelliklerine de odaklanmak zorundadır (Uraz, 1978). Bu çerçevede pazarlama disiplini, tüketici kararlarını açıklamak için “kara kutu” kavramını kullanır. Çünkü satın alma süreci, birçok faktörün etkileşimiyle şekillenen karmaşık bir yapıya sahiptir ve bu sürecin tam olarak nasıl işlediği hâlâ net olarak bilinmemektedir. Kara kutu, bu bilinmezliği ve öngörülemezliği temsil eder (Böge, 1994).

Tüketici davranışı, yalnızca yukarıda belirtilen faktörlerle şekillenmez. İşletmelerin pazarlama faaliyetleri ve çabaları da bu davranış üzerinde önemli bir etkiye sahiptir. Özellikle özel gün ve haftaların tüketim üzerindeki etkileri, bu açıdan ele alınmalıdır. Tüketici davranışları, genellikle “kara kutu” olarak kabul edilse de, bu davranışların arkasında insanın zihinsel, duygusal ve fiziksel tepkileri ile ihtiyaçları yatar. Pazarlama faaliyetleri, bu yönleri uyandırarak tüketicilerin satın alma davranışını tetiklemeyi hedefler. Özel günler yaklaştıkça artan yoğun ve çoğu zaman agresif reklam kampanyalarının, en rasyonel tüketicileri bile etkilediği görülmektedir. Bu sayede işletmeler, söz konusu günleri adeta toplumsal bir tüketim çılgınlığına dönüştürebilmektedir. Tüketici davranışlarında duyguların ne denli güçlü bir rol oynadığının fark edilmesi, pazarlama stratejilerinin de bu duygular üzerine yoğunlaşmasına yol açmıştır. Sevgi, özlem, hatırlanma gibi temalar üzerinden verilen mesajlar aracılığıyla daha fazla tüketim teşvik edilmekte, böylece tüketim kültürünün yaygınlaşması ve bireylerin zamanla birer “tüketim toplumu” üyesine dönüşmesi kolaylaştırılmaktadır (Bayuk, 2009).

Günümüzde üretim süreçlerinin dijitalleşmesiyle birlikte tüketim davranışları da önemli ölçüde değişmiş, özellikle çevrim içi alışverişin yaygınlaşması tüketim hacmini ciddi oranda artırmıştır. Tüketicilerin yalnızca tek bir tıklamayla alışveriş yapabilmesi, bu artışın başlıca nedenlerinden biridir. Bu hızlı tüketim ortamında, satışların doğru biçimde tahmin edilmesi ve bu tahminlere dayalı olarak stok yönetimi ile üretim planlamasının yapılması

büyük önem taşımaktadır. Sağlıklı tahminler olmaksızın, işletmelerin gelir-gider dengesini öngörmeleri neredeyse imkânsızdır. Üretim planlama, satın alma, satış ve pazarlama gibi birçok departman, kararlarını bu öngörüler doğrultusunda şekillendirir (Nacar ve Erdebilli, 2021). Uzun veya kısa vadeli satış tahminlerinin doğru yapılmaması, müşteri memnuniyetsizliğine, para kaybına, hammadde eksikliğine ve benzeri sorunlara yol açabilir. Ayrıca, stratejik planlama sürecinde tahmin eksikliği, işletmenin hedef ve politikalarını belirlemesini engeller. Bu durum sadece işletmenin kendisini değil, aynı zamanda tedarik zincirinde yer alan üreticiler, perakendeciler, tedarikçiler ve nihai tüketiciler dâhil tüm paydaşları olumsuz etkileyebilir. Olası tahmin hataları zincirleme biçimde tüm süreci sekteye uğratabilirken, doğru yapılan tahminler yalnızca işletme kârlılığını artırmakla kalmaz, aynı zamanda tüm tedarik zincirinde verimlilik sağlar. Böylelikle işletmeler sektördeki rekabetçiliğini korurken, tedarik zincirindeki diğer paydaşlar için de sürdürülebilir bir kârlılık ortamı yaratılmış olur (Nacar ve Erdebilli, 2021; Yücesan, 2018).

Bununla birlikte, kapasite planlaması, işletmeler için kritik bir kavramdır çünkü işletmeler yalnızca sahip oldukları kapasiteye göre üretim yapabilirler. Bu nedenle, kârlı olabilmek için satışlarını da bu kapasiteye uygun şekilde yönetmek zorundadır. İşletmelerin sahip oldukları kapasiteyi en verimli şekilde kullanmaları gerekir ve bu da doğru bir kapasite yönetimiyle mümkündür. Kapasite planlaması, işletmelerin belirli bir zaman diliminde kullanacakları kapasiteyi belirlemek için yapılır. Bu planlamanın hatalı yapılması, işletmeleri fazla veya yetersiz kapasite sorunuyla karşı karşıya bırakabilir. Böyle bir durumda, işletmenin verimliliği ve karlılığı beklenen seviyeye ulaşamayacaktır. Kapasite planlaması yapılırken, mevcut kapasitenin yanı sıra, işletmenin tahmin edilen satış miktarları da dikkate alınmalı ve planlama bu verilere göre yapılmalıdır. Bu nedenle, güvenilir satış tahminleri, işletmelere pek çok açıdan fayda sağlar. Güvenilir satış tahminleri, yalnızca kapasite planlamasını değil; iş stratejisinin oluşturulmasını, maliyetlerin azaltılmasını ve kârın artırılmasını da doğrudan etkiler. Dolayısıyla satış tahminleri, işletmenin büyümesi ve rekabetçi bir pozisyon elde etmesi açısından karar alma süreçlerinde temel bir parametre olarak öne çıkar (Chen vd., 2010).

Uygulamada, birçok işletme satış tahminini doğrudan yapma yoluna gitmektedir. Ancak, bu yaklaşım yerine satış tahminlerinin; pazar potansiyeli ve işletmenin satış potansiyeline dair önceden yapılacak değerlendirmelere dayandırılması daha sağlıklı sonuçlar verecektir. Pazarlama planlaması işletmenin planlama işlevinin önemli bir aşamasını oluşturmakla birlikte

işletmenin genel stratejisi diğer departmanların faaliyetleri ile uyum içinde olmak zorundadır (Eroğlu, 1996).

Satış tahmini yapmak, tüketici davranışlarının belirsizliği ve özelleştirilmiş ürünlere artan ilgi nedeniyle oldukça zorlaşmıştır. Günümüzde, herhangi bir yöntem kullanılmadan büyük ölçekli satış tahminleri yapmak işletmeler açısından neredeyse imkânsızdır. Geleneksel yöntemler, büyük veri kümeleriyle başa çıkmakta yetersiz kalmakta; esneklikten uzak olmaları ve özel durumları göz ardı etmeleri sebebiyle doğru tahminler üretmekte zorlanmaktadır. Bu noktada, gittikçe yaygınlaşan yapay zekâ teknolojileri bu zorlukların çözümünde kullanılmaya başlamıştır (Acı ve Doğanşoy, 2022; Erol ve İnkaya, 2024). Yapay zekanın yalnızca robotik ya da telekomünikasyon gibi alanlarda kullanıldığına dair yaygın bir önyargı bulunsun da bu anlayış giderek değişmekte ve yapay zekâ, veriyle ilgili tüm alanlarda yaygınlaşmaktadır. Bu dönüşümün önemli bir parçası da makine öğrenmesidir (Nacar ve Erdebili, 2021). Makine öğrenmesi, pek çok mühendislik probleminde hızlı ve etkili çözümler sunma potansiyeline sahiptir. Henüz gelişmekte olan bir alan olması nedeniyle literatürde çeşitli varsayımlar yer almaktadır. Bununla birlikte, birçok işletme veri biliminin sunduğu avantajların, geleneksel yöntemlerden daha hızlı ve isabetli olduğunu henüz fark etmiş değildir. Bu çalışma, söz konusu farkındalığın artmasına katkı sağlayarak işletmelere yol gösterici olmayı amaçlamaktadır.

Bu çalışmada Rastgele Orman yöntemini kullanılarak satış tahmini için model oluşturulmuş ve tahmin yöntemlerinde makine öğrenmesi algoritmalarının kullanılmasının gerekliliği üzerinde durulmuştur.

Çalışmanın ikinci ve üçüncü bölümünde sırasıyla makine öğrenmesi ve satış tahmini konuları ile ilgili literatür araştırması yapılmıştır. Dördüncü bölümde metod kısmında rasgele orman yöntemi anlatılmış ardından beşinci bölümde Amazon'daki OLED televizyon satış verileri örnek bir satış tahmini çalışması gerçekleştirilmiştir. Son bölümde ise tartışma ve sonuç bölümüne yer verilmiştir. Çalışmada özellikle özel günlerde yetersiz stok nedeniyle işletmelerin uğradıkları zararın gösterilmesi amaçlanmıştır. Makine öğrenmesi yöntemleriyle doğru satış tahminlerinin yapılmasının işletmeler için çok önemli olduğu gösterilmeye çalışılmıştır.

2. Makine Öğrenmesi

Dijitalleşen dünyada; nesnelerin interneti (IoT), bulut bilişim, büyük veri (Big Data), siber güvenlik, yapay zekâ ve makine öğrenmesi gibi kavramlar giderek daha fazla önem kazanmaktadır. Özellikle büyük veri kavramı, günümüz bilgi yönetimi açısından belirleyici bir rol oynamaktadır. Geçmişte

veriler fiziksel ortamda, defterler aracılığıyla kaydedilirken; günümüzde veri miktarının katlanarak artması, bu yöntemi sürdürülemez hâle getirmiştir. Verinin bu denli hızlı çoğalması, büyük veri yığınları içinden anlamlı ve doğru bilgilere ulaşmayı oldukça karmaşık bir süreç hâline getirmiştir. Bu noktada devreye giren veri madenciliği, büyük veri setleri içerisindeki örüntüleri, eğilimleri ve anlamlı ilişkileri ortaya çıkarmayı amaçlar. Bu sayede, veri yığınları sadece depolanan bir içerik olmaktan çıkıp, stratejik kararlar için kullanılabilir bilgiye dönüşür (Savaş vd., 2012).

Veriden anlamlı bilgi çıkarılması, yalnızca mevcut durumu analiz etmeye değil, aynı zamanda gelecekteki olaylara dair öngörülerde bulunmaya da olanak tanır. Bu öngörü süreci, özellikle karar alma süreçlerinde büyük önem taşır. İnsanlar, yaşam deneyimleri ve eğitimleri doğrultusunda tahminlerde bulunabilir; ancak karar verirken duygusal etkenlerin devreye girmesi, rasyonel seçimler yapmayı zorlaştırabilir (Tosunoğlu vd., 2021). Ayrıca, veri miktarının fazla olduğu durumlarda bazı önemli ayrıntıların gözden kaçması olasıdır ve bu da yanlış kararlarla sonuçlanabilir. Buna karşın, eğitilmiş makineler tüm verileri hızlı bir şekilde analiz edip, çok kısa bir sürede mantıklı kararlar alabilir. Makine öğrenmesinde, insan düşünme biçimini ve karar alma yeteneğini doğru bir şekilde modellemek oldukça önemlidir (İnik ve Ülker, 2017). Bu bağlamda makine öğrenmesi, insan zekâsını taklit etmeye yönelik hesaplama algoritmalarının bir dalı olarak tanımlanmakta; aynı zamanda veriye dayalı problemlerin uygun algoritmalar aracılığıyla modellenmesini ifade etmektedir (El-Naqa ve Murphy, 2015; Atalay ve Çelik, 2017). Bununla birlikte, büyük veri; çeşitlilik, hız ve hacim açısından önemli zorluklar doğurur (Suthaharan, 2014). Makine öğrenmesi algoritmaları, bu büyük ve karmaşık veri yığınlarını doğru şekilde sınıflandırarak söz konusu zorlukların üstesinden gelmede etkili çözümler sunmaktadır (Ersen vd., 2022; Keskin, 2025).

Makine öğrenmesi ilk kez 1959 yılında, Arthur Samuel tarafından bir dama oyununu analiz etmek amacıyla kullanılmıştır. Samuel, makine öğrenmesini bilgisayarlara açıkça programlanmaksızın öğrenme yetisi kazandıran bir araştırma alanı olarak tanımlamıştır (Samuel, 1959).

1970'li yıllarda, Japonya'da bu alanda yapılan araştırmaların ardından ilk insansı robot olan Wabot-1 geliştirilmiştir (İnal, 2021). 1980'li yıllarda ise istatistiksel yöntemler ve veri madenciliğindeki ilerlemeler, makine öğrenmesinin teorik temellerini güçlendirmiş; bu dönemde uzman sistemler ve karar destek sistemleri gibi yapay zekâ uygulamaları yaygınlık kazanmaya başlamıştır (Foote, 2021).

1980'lerden itibaren hızla gelişen makine öğrenmesi, 1990'lı yıllarda veriye dayalı yöntemlere odaklanmaya başlamıştır. Bu dönemin en dikkat çekici gelişmelerinden biri, IBM tarafından geliştirilen 'Deep Blue' adlı satranç programının, 1997 yılında Dünya Satranç Şampiyonu Garry Kasparov'u yenerek kamuoyunun büyük ilgisini çekmesidir (Campbell vd., 2002).

1990'lar ve 2000'ler, makine öğrenmesinin hızla geliştiği ve uygulama alanlarının genişlediği dönemler olmuştur. Bu yıllarda büyük veri kavramının ortaya çıkması, veri erişimini önemli ölçüde artırmış; bu da makine öğrenmesi algoritmalarının daha verimli çalışmasına olanak tanımıştır (Foote, 2021). Aynı dönemde destek vektör makineleri, rastgele ormanlar ve derin öğrenme gibi birçok önemli yöntem geliştirilmiş, alandaki ilerlemeye ivme kazandırmıştır.

2010'lu yıllarda ise işlem gücünü artıran donanımların yaygınlaşması, veri akışının daha da hızlanması ve derin öğrenme modellerinin evrilmesiyle birlikte makine öğrenmesinde büyük bir sıçrama yaşanmıştır. Özellikle derin öğrenme ve sinir ağı tabanlı yaklaşımlar, görüntü tanıma, doğal dil işleme ve metin analizi gibi alanlarda çarpıcı başarılar elde etmiştir. Bu gelişmeler, endüstriden sağlığa, otomasyondan eğitime kadar pek çok sektörde köklü değişimlere yol açmıştır (Koch, 2022).

Makine öğrenmesi, temelde olasılık, istatistik ve optimizasyon gibi matematiksel disiplinlere dayanır. Bu süreçte, geçmiş verilere dayalı olarak bir model oluşturulur ve bu model aracılığıyla geleceğe yönelik tahminler yapılması hedeflenir. Bu özellik, makine öğrenmesini diğer yazılım mühendisliği yöntemlerinden ayıran önemli bir farktır. Geleneksel yöntemlerde, yazılım geliştiricilerinin veri setlerini manuel olarak inceleyip belirli kurallar ve kalıplar oluşturması beklenir. Ancak bu yaklaşım, genellikle geliştiricinin önyargılarına dayanır ve gerçekçi olmayan sonuçlar doğurabilir (Grigorev, 2020).

Makine öğrenmesi, giderek daha geniş bir alanda uygulanmakta ve bu yayılımın önümüzdeki yıllarda da artarak devam etmesi beklenmektedir. Bu eğilimin birçok nedeni vardır. İlk olarak, sembolik makine öğrenmesi, hesaplamalı öğrenme teorisi, sinir ağları, istatistik ve örüntü tanıma gibi farklı araştırma alanlarında çalışan topluluklar zamanla birbirlerini keşfetmiş ve aralarında iş birlikleri gelişmeye başlamıştır. Bu disiplinler arası etkileşim, makine öğrenmesinin gelişimini hızlandırmıştır. İkinci olarak, makine öğrenmesi yöntemleri yalnızca konuşma tanıma, yüz tanıma, el yazısı tanıma ve tıbbi tanı gibi geleneksel problemlere değil; aynı zamanda bilgi keşfi, doğal dil işleme, robot kontrolü ve birleşik optimizasyon gibi daha yeni

ve karmaşık problem alanlarına da başarıyla uygulanmaktadır (Dietterich, 1997)

Makine öğrenmesinin temel türleri, gözetimli ve gözetimsiz öğrenme olarak ikiye ayrılır. Gözetimli öğrenme, daha önce belirlenmiş doğru sınıflandırmalara dayalı olarak, veri kaynağından örneklerin eğitilmesini sağlar. Gözetimsiz öğrenme ise, etiketlenmemiş verilerdeki gizli desenleri tespit etmek için kendi kendini organize eden sinir ağlarını kullanmayı öğrenir. Gözetimsizlik, çözümü değerlendirecek bir hata sinyali olmadan bilgi öğrenme ve organizasyonu yapabilme yeteneğini ifade eder (Sathya ve Abraham, 2013). Gözetimli öğrenme ise, sınıflandırma ve regresyon olmak üzere iki alt türe ayrılır. Veri setine bağlı olarak farklı algoritmaların kullanılması mümkündür (Nacar ve Erdebilli, 2021).

3. Satış Tahmininde Makine Öğrenmesi Uygulamaları

Satış tahminleri, tedarik zincirindeki önemli araştırma konularından biridir (Ingle vd., 2021). Bu alanda kullanılan yöntemler istatistiksel yöntemler ve yapay zekâ yöntemleri olmak üzere iki başlıkta incelenebilir (Ingle vd., 2021). İstatistiksel yöntemler arasında doğrusal regresyon, hareketli ortalama, ağırlıklı ortalama, üstel düzeltme ve ARIMA gibi teknikler bulunmaktadır (Ingle vd., 2021; Acı ve Doğanşoy, 2022; Erol ve İnkaya, 2024). Yapay zekâ yöntemleri ise, makine öğrenmesi ve derin öğrenme olmak üzere iki ana gruba ayrılabilir. Makine öğrenmesi yöntemleri, k-en yakın komşuluk, destek vektör makineleri, karar ağaçları ve yapay sinir ağları gibi teknikleri içerirken; derin öğrenme yöntemleri arasında çok katmanlı derin sinir ağları (DNN), evrimsel sinir ağları (CNN), tekrarlayan sinir ağları (RNN) ve uzun kısa süreli bellek ağları (LSTM) yer almaktadır (Ingle vd., 2021; Acı ve Doğanşoy, 2022; Erol ve İnkaya, 2024).

Bir araştırmaya başlamadan önce, önceki çalışmaların gözden geçirilmesi büyük önem taşımaktadır. Bu, yapılacak olan çalışmanın mevcut literatüre uygun olmasını ve daha önceki araştırmalarla tutarlı bir şekilde desteklenmesini sağlar. Satış tahminlerine yönelik yapılan literatür taramaları, veri madenciliği tekniklerinin sıkça kullanıldığını ortaya koymaktadır. Çünkü talep tahmini, satış tahminiyle doğrudan ilişkili bir kavramdır ve bu nedenle araştırmalarda “talep tahmini” ve “satış tahmini” kavramları genellikle birbirinin yerine kullanılmıştır.

Carlson ve Umble (1980), Amerika’da beş farklı otomobil segmentinde (standart ve lüks otomobiller) beş yıllık talep tahminlerini regresyon analizi kullanarak yapmışlardır. Çalışmada, bağımsız değişkenler olarak; benzin fiyatları, benzin kıtlığının piyasa üzerindeki etkisi, taşıt fiyatları, tüketici gelir

düzeyleri ve sektördeki işçi grevleri gibi faktörler incelenmiştir (Carlson ve Umble, 1980).

Gavcar vd. (1999), Türkiye'deki kâğıt ve karton talebini tahmin etmek amacıyla bir çalışma yapmışlardır. Çalışmada, kâğıt ürünleri, ithalat ve ihracat hacimleri, eşya fiyat endeksi, gayrisafi milli hasıla ve nüfus gibi değişkenler analizde kullanılmıştır. Talep tahmini, çoklu regresyon analiziyle yapılmış ve modeldeki değişkenler arasındaki ilişkiler korelasyon analizi ile incelenmiştir. Sonuç olarak, 21. yüzyılda talep edilecek kâğıt miktarına yönelik bir tahmin yapılmıştır (Gavcar vd., 1999).

Aksel (2000), İstanbul ve Türkiye genelinde uzun dönem elektrik talep tahmini yapmıştır. 2001, 2003 ve 2005 yıllarına ait yük enerji tüketim değerlerini tahmin etmek için çok değişkenli regresyon analizi ve geriye yayılım algoritmasına sahip, iki gizli katmanlı yapay sinir ağı yöntemlerini kullanmıştır. Çalışmada kullanılan değişkenler; nüfus, kişi başına düşen gayri safi milli hâsıla, gelişim hızı, sanayi üretim endeksi, petrol varil fiyatı ve elektrik enerjisi birim satış fiyatı olmuştur. Elde edilen tahminler, çok değişkenli regresyon analizi ve geriye yayılım algoritmasının sonuçları, ağırlıklı ortalama yöntemiyle tek bir sonuca dönüştürülmüştür (Aksel, 2000).

Efendigil vd. (2009), beyaz eşya tedarik zinciri için talep tahmini yapmışlardır. 24 aylık veri, her perakendeciden anket yoluyla toplanmış ve bu veriler 79 eğitim ve 17 test verisi olarak ikiye ayrılmıştır. Çalışmada kullanılan değişkenler; ürün fiyatı, ürün kalitesi ve promosyonlar olmuştur. Tahmin yöntemi olarak Bulanık Mantık ve Yapay Sinir Ağları kullanılmıştır. Çalışmanın sonucunda, Bulanık Mantık uygulamasında test edilen Gauss üyelik fonksiyonunun daha etkili bir çözüm sunduğu gözlemlenmiştir (Efendigil vd., 2009).

Demirel (2009), elektrik enerjisi talep tahmini üzerine bir tez çalışması yapmıştır. Çalışmasında kullanılan değişkenler; Gayri Safi Milli Hâsıla, üretilen enerji, tüketilen enerji, nüfus ve kurulu güç verileri olmuştur. Tahmin edilen sonuçlar, ilgili kurumun öngörü değerleriyle karşılaştırılmıştır (Demirel, 2009).

Yücesoy (2011), kâğıt talebi tahmini üzerine bir çalışma yapmıştır. Bu çalışmada Yapay Sinir Ağları (YSA) ve Çoklu Doğrusal Regresyon modelleri kullanılmıştır. Çalışmada yer alan değişkenler; üretim miktarı, ithalat hacmi, kâğıt basım sanayi toptan fiyat endeksi, GSYH, nüfus verileri ve okuryazarlık oranıdır. Tahmin sonuçları incelendiğinde, YSA modeli %98,89 başarı oranı ile öne çıkarken, regresyon modelinin başarı oranı ise %61,89 olarak belirlenmiştir (Yücesoy, 2011).

Demirbaş (2011), kombi satış adetlerini tahmin etmeyi amaçlayan bir çalışma yapmıştır. Çalışmanın hedefi, talep tahmini yaparak müşteri kaybı riskini azaltmaktır. Tahminler, zaman serisi analizi yöntemiyle gerçekleştirilmiştir ve 2008-2010 yılları arasındaki satış verileri kullanılmıştır. Araştırma sonuçlarına göre, en başarılı tahminlerin en küçük kareler tahmincileri tarafından yapıldığı bulunmuştur (Demirbaş, 2011).

Karahan (2011), talep tahmini yapmak amacıyla Yapay Sinir Ağları (YSA) kullanmıştır. Modelde, bir gizli katmanlı, ileri beslemeli ve geri yayılım gösteren bir sinir ağı yapılandırılmıştır. Çalışmada 2004-2010 yılları arasındaki veriler kullanılmıştır. Oluşturulan modelde beş girdi değişkeni ve bir çıktı değişkeni bulunmaktadır. Çıktı değişkeni, kuru kayısı talebini tahmin etmektedir. Bağımsız değişkenler arasında geçmiş talep verileri, tarih, döviz kuru, aylık ortalama fiyat, pazar sayısı ve mevsimsel etkiler yer almıştır. Ağın eğitimi için momentum ve öğrenme katsayısı 0,1 olarak belirlenmiştir. Yapılan tahminlerin doğruluğu %86 olarak bulunmuş ve gerçek verilerle tahmin edilen değerler arasındaki farkın oldukça düşük olduğu gözlemlenmiştir. Çalışmanın amacı, tahmin edilen verilere dayanarak firmanın üretim planlamasını önceden yapabilmesidir (Karahan, 2011).

Hicham ve Bouhorma (2012), ambalaj satışlarını tahmin etmek amacıyla bir çalışma yapmıştır. Çalışmada kullanılan değişkenler; imalat tüketici endeksi, rekabet endeksi ve geçmiş satış verileridir. Tahminler, Delphi yöntemi ve Yapay Sinir Ağları (YSA) ile gerçekleştirilmiştir. Sonuçlar, MAPE ve RMSE oranlarıyla değerlendirilmiş ve Delphi yönteminin YSA yöntemine kıyasla daha başarılı sonuçlar verdiği sonucuna varılmıştır (Hicham ve Bouhorma, 2012).

Karaatlı vd. (2012), otomobil satış tahmini yapmak için Yapay Sinir Ağları (YSA) yöntemini kullanmışlardır. Veri seti, Ocak 2007 ile Haziran 2011 arasındaki aylık verilere dayanmaktadır. Çalışmada kullanılan değişkenler; gayri safi yurt içi hasıla (GSYH), reel kesim güven endeksi (RKGE), yatırım harcamaları (YH), tüketim harcamaları (TH), tüketici güven endeksi (TGE), döviz kuru gibi bağımsız değişkenler ile satılan toplam otomobil sayısı bağımlı değişken olarak alınmıştır. Modelin tahminleri Yapay Sinir Ağları ile yapılmış ve MAPE değeri kullanılarak doğruluk analizi gerçekleştirilmiştir. Gerçek ve tahmin edilen değerler karşılaştırıldığında, yapılan tahminlerin başarılı olduğu sonucuna ulaşılmıştır (Karaatlı vd., 2012).

Azadeh vd. (2013), demir çelik sektöründe talep tahmini yapmışlardır. Çalışmada kullanılan değişkenler; milli gelir, büyüme oranı, enflasyon ve çelik üretim verileri olmuştur. Tahminler, Bulanık Mantık ve Yapay Sinir Ağları (YSA) yöntemleriyle gerçekleştirilmiştir. Araştırmacılar, önerdikleri

mükemmeliyet modelinin esnek olduğunu ve bu modelin hem küçük hem de büyük ölçekli şirketlerde uygulanabileceğini vurgulamışlardır (Azadeh vd., 2013).

Koochakpour ve Tarokh (2016) tarafından yapılan bir çalışmada, bir işletmenin satış tahminleri yapılmıştır. Bu tahmin için kullanılan değişkenler arasında milli gelir, döviz kuru, enflasyon ve geçmiş satış verileri yer almaktadır. Çalışmada, tahmin yöntemleri olarak Bulanık Mantık ve Yapay Sinir Ağları kullanılmıştır. Araştırmanın bulgularına göre, Bulanık Mantık modelleri, Yapay Sinir Ağlarına göre daha başarılı sonuçlar elde etmiştir (Koochakpour ve Tarokh, 2016).

Yılmaz (2018) tarafından gerçekleştirilen bir çalışmada, aylık konut satışları tahmin edilmiştir. Çalışmada, Yapay Sinir Ağları (YSA) ve regresyon yöntemleri kullanılmıştır. Veri seti, 2013-2017 yılları arasında aylık olarak toplanmıştır. Tahmin için; konut fiyatları, kredi durumu, nüfus sayısı, GSYİH, işsizlik oranı ve ailedeki kişi sayısı gibi değişkenler değerlendirilmiştir. Yapılan analiz sonucunda, tahmin edilen veriler 2017 yılındaki gerçek verilerle karşılaştırılmış ve YSA modelinin gerçek verilere daha yakın tahminler sağladığı gözlemlenmiştir (Yılmaz, 2018).

Rincon-Patino vd. (2018), avokado satışlarını inceleyen bir çalışma gerçekleştirmiştir. Hava koşullarının etkisi göz önünde bulundurularak yapılan bu çalışmada, tarım sektörüne ait veriler makine öğrenimi yöntemleriyle analiz edilmiştir. Çalışmada, Doğrusal Regresyon, Yapay Sinir Ağları'ndaki çok katmanlı algılayıcı (MLP), Destek Vektör Makineleri ve Çok Değişkenli Regresyon modelleri kullanılmıştır. Elde edilen sonuçlar değerlendirildiğinde, Destek Vektör Makineleri ve Çok Değişkenli Regresyon modellerinin daha başarılı tahminler sunduğu tespit edilmiştir (Rincon-Patino vd., 2018).

Ecemiş (2018), paslanmaz çelik satışlarını tahmin etmek amacıyla bir çalışma yapmıştır. Bu çalışmada, hammadde fiyatları, USD/TRY kuru, ÜFE ve sanayi üretim endeksi gibi değişkenler kullanılmıştır. Tahmin modelleri olarak Yapay Sinir Ağları ve Destek Vektör Makineleri (SVM) test edilmiştir. Elde edilen sonuçlar, MAPE, MAE ve RMSE gibi hata oranlarına göre karşılaştırılmıştır. Yapay Sinir Ağları yöntemi, makine sektörü dışında kalan tüm sektörlerde Destek Vektör Makineleri'ne kıyasla daha başarılı sonuçlar elde etmiştir (Ecemiş, 2018).

Kılıç vd. (2018), yemekhane talebini yapay zekâ teknikleriyle tahmin etmeye yönelik bir çalışma yapmıştır. Bu çalışmanın amacı, yemek üretiminin, talebi karşılayacak şekilde doğru miktarda yapılmasını sağlamaktır.

Üniversitelerde talep miktarı zaman zaman değişkenlik göstermektedir, bu yüzden doğru bir talep tahmini yaparak bu durumu çözmeyi hedeflemiştir. Kullanılan değişkenler arasında; tüketici türleri, yemekhaneden yararlanan öğrenci ve personel sayıları yer almıştır. Çalışmada, Yapay Sinir Ağı, Destek Vektör Makineleri ve Lincer Regresyon teknikleri kullanılmıştır. Sonuçlar değerlendirildiğinde, Destek Vektör Makineleri'nin bu analizde daha başarılı tahminler sunduğu belirlenmiştir (Kılıç vd., 2018).

Pavlyshenko (2019), çalışmasında Rossmann mağazalarının satış taleplerini incelemiştir. Araştırma, R programı ve Jupyter Notebook kullanılarak yapılmıştır. Çalışmada, Yapay Sinir Ağları, Rastgele Orman Algoritması ve Ağaç tabanlı modeller gibi çeşitli modeller kullanılmıştır. Yapılan analizler sonucunda, satış tahminlerinde regresyon modellerinin zaman serisi modellerine kıyasla daha doğru sonuçlar verdiği tespit edilmiştir (Pavlyshenko, 2019).

Türk ve Kıanı (2019), beyaz eşya sektöründe satış tahminleri yapmışlardır. Çalışmalarında beyaz eşya satış miktarı (SM) bağımlı değişken olarak kullanılırken, tüketici güven endeksi (TGE), reel kesim güven endeksi (RKGE), gayri safi yurt içi hasıla (GSYH), sanayi üretim endeksi (SÜE), tüketim harcamaları (TH), hane halkı maddi durum beklentisi (HMDDB), ekonomi güven endeksi (EGE) ve evlenme istatistikleri bağımsız değişkenler olarak yer almıştır. Tahmin için Yapay Sinir Ağları (YSA) ve regresyon yöntemleri kullanılmıştır. Sonuçlar, YSA modellerinin regresyon analizlerine kıyasla daha başarılı olduğunu göstermiştir. Ayrıca, tahmin edilen değerler ile gerçek veriler karşılaştırıldığında, elde edilen sonuçların gerçek verilere oldukça yakın olduğu gözlemlenmiştir (Türk ve Kıanı, 2019).

Catal vd. (2019), satış tahminleri üzerine bir araştırma yapmıştır. Araştırmada, klasik zaman serisi yöntemleri ile makine öğrenmesi algoritmalarının karşılaştırılması yapılmıştır. Kullanılan yöntemler arasında Bayes, Yapay Sinir Ağları ve Karar Ağacı Yöntemleri yer almaktadır. Tahmin güvenilirlikleri, ortalama kare hata ve kök ortalama kare hata sonuçlarıyla değerlendirilmiştir. Başlangıçta, karar ağacı regresyon modeli en düşük hatayı verirken, tahmin sürecinin sonunda arttırılmış karar ağacının en düşük hatayı sağladığı gözlemlenmiştir. Bu bulgular ışığında, çalışmada regresyon modellerinin klasik zaman serisi yöntemlerinden daha iyi tahminler sunduğu sonucuna varılmıştır (Catal vd., 2019).

Weng vd. (2019), tedarik zinciri için satış tahminleri yapmak amacıyla Light Gradient Boosting Machine (LightGBM) ve Long Short-Term Memory (LSTM) yöntemlerini birleştiren bir model önermişlerdir. Çalışmalarında, Kaggle platformunda yer alan Corporación Favorita Market

Sale ve Rossman mağaza satışları ile Jollychic E-ticaret platformundaki tedarik zinciri veri setleri kullanılmıştır. Modelin başarısı için değişken seçim sürecinin büyük bir öneme sahip olduğu vurgulanmıştır. Veri setini analiz etmek için, veri görselleştirme tekniklerinden biri olan dağılım grafiği tercih edilmiştir. Değişken mühendisliği aşamasında, verilerin satış verisi olup olmadığına göre eksik veriler 0 veya -1 ile tamamlanmış, alakasız değişkenler ise veri setinden çıkarılmıştır. Ayrıca, verilerde herhangi bir anomali olup olmadığı da incelenmiştir. Modelin doğruluğunu değerlendirmek için Normalleştirilmiş Ağırlıklı Kök Ortalama Kare Logaritmik Hata (NWRMSLE) metriği kullanılmıştır. Çalışmada, Ridge, Destek Vektör Makineleri, Rastgele Orman, XGBoost, LightGBM ve LSTM yöntemleri ile birleşik modelin doğruluğu karşılaştırılmıştır. Sonuçlar, önerilen birleşik modelin üç farklı veri seti üzerinde başarılı bir şekilde çalıştığını ortaya koymuştur (Weng vd., 2019).

Helmini vd. (2019), çalışmalarında ARIMA gibi istatistiksel tahmin modellerinin genellikle yaygın olarak kullanıldığını, ancak doğrusal olmayan problemleri analiz etmekte yetersiz kaldığını belirtmişlerdir. Bu nedenle, doğruluğu diğer yöntemlere göre daha yüksek olan makine öğrenmesi tekniklerinin kullanılması gerektiğini vurgulamışlardır. Bu amaçla, yazarlar özgün bir Long Short-Term Memory (LSTM) modeli geliştirmiş ve bu modele hiperparametre optimizasyonu uygulamışlardır. Geliştirilen LSTM modeli, XGBoost ve Rastgele Orman modelleriyle karşılaştırılmıştır. Modeli test etmek için Kaggle platformundaki Rossmann mağaza satış veri seti kullanılmıştır. Değerlendirme için RMSE ve MAE hata fonksiyonları kullanılmıştır. Çalışmada, geliştirilen LSTM modelinin doğruluğunun, diğer makine öğrenimi yöntemlerine göre çok daha üstün olduğu sonucuna varılmıştır (Helmini vd., 2019).

Poyraz (2020), eczaneler için ilaç talep tahmini yapmıştır. Çalışmada, makine öğrenmesi yöntemleri kullanılarak gelecekteki yılların satış miktarları tahmin edilmiştir, böylece etkili bir stok yönetimi sağlanması hedeflenmiştir. Kullanılan makine öğrenme algoritmaları arasında Linear Regresyon, Gaussian Process, M5Rules, Yapay Sinir Ağları'nın çok katmanlı algılayıcı modeli, SMOreg, M5P ve Random FOREST yer almıştır. Veri seti, 2015-2019 yılları arasındaki haftalık verilerden oluşmaktadır. Sonuçlar, MAPE değeriyle karşılaştırılmış ve M5Rules algoritması en düşük hata oranına sahip olarak belirlenmiştir. Ayrıca, Gaussian Processes algoritması ise üç hafta boyunca en iyi sonuçları elde eden algoritma olmuştur (Poyraz, 2020).

Verstraete vd. (2020), taktiksel satış tahmini yapmak için Lasso yöntemini kullanarak bir çalışma gerçekleştirmiştir. Çalışmada, tedarik zincirinin tüm

unsurları için taktiksel satış tahmininin önemli olduğu vurgulanmıştır. Yazarlar, geleneksel satış tahmin yöntemlerinin bazı makroekonomik faktörlere uyum sağlamakta yetersiz kaldığını belirtmişlerdir. Makroekonomik değişiklikler öngörülemediği için, genellikle uzman görüşleri veya çalışanların yaptığı varsayımlar kullanılır, ancak bu tür yaklaşımlar hem pahalıdır hem de analitik açıdan doğrulanabilir değildir. Bu yüzden, makroekonomik değişiklikleri tahmin etmek amacıyla Lasso yöntemi tercih edilmiştir. Çalışmada kullanılan veri seti, otomotiv sektöründeki bir firmadan alınan 62 aylık verilerden oluşmaktadır. Modelin hata oranı, ortalama mutlak yüzde hata kullanılarak hesaplanmıştır. Uzun vadeli tahmin çalışması sonucunda, klasik tahmin yöntemlerine göre hata oranının %54,5 azaldığı gözlemlenmiştir (Verstraete vd., 2020).

Antipov ve Pokryshevskaya (2020), Gradient Boosting Machine (GBM), Rastgele Orman ve ElasticNet yöntemlerini karşılaştırarak, satış tahminlerinde kullanılabilecek genel bir yaklaşım geliştirmişlerdir. Çalışmada, yorumlanabilir makine öğrenmesi teknikleri kullanılarak, eyleme geçirilebilir tahminler elde edilmiştir. Veri seti, 156 haftalık bir dönemi kapsamakta olup, bir perakende analitik şirketi tarafından temin edilmiştir (Antipov ve Pokryshevskaya, 2020).

Tozak (2021), bir tekstil firmasının satış verilerini tahmin etmek amacıyla veri madenciliği programlarını kullanmıştır. Veriler, yaklaşık dört yıl boyunca aylık olarak toplanmıştır ve karar ağacı algoritması ile tahmin edilmiştir. Çalışmada yer alan değişkenler arasında; müşteri sayısı, üretim yeri, fabrika, yıl, ay, sevk adedi ve birim fiyat bulunmaktadır. Elde edilen tahminler sonucunda, firmanın yıllık üretim kapasitesinin ne kadar olması gerektiği belirlenmiş ve bu doğrultuda üretici için doğru bir planlama yapma hedeflenmiştir (Tozak, 2021).

Nacar ve Erdebili (2021), uçtan uca bir makine öğrenmesi proje sürecini ele almış ve bir makine öğrenmesi projesinin aşamaları ile veriye yaklaşım yöntemini tanıtmışlardır. Uygulama kısmında, doğrusal regresyon, Ridge, Lasso, Elastic Net, K-en yakın komşu ve Rastgele Orman gibi makine öğrenmesi algoritmaları kullanılarak gerçek bir veri seti üzerinde satış tahmin modeli geliştirilmiştir. Bu modelde, en düşük hata oranını veren algoritma olarak Rastgele Orman belirlenmiştir (Nacar ve Erdebili, 2021).

Demirer-Polat (2022) çalışmasında, beyaz eşya sektöründe faaliyet gösteren bir firmanın aylık bazda toplanan verileri incelenmiştir. Bağımlı değişken olarak sevk verileri kullanılırken, bağımsız değişkenler arasında reel kesim güven endeksi, gayri safi yurtiçi hasıla, sanayi üretim endeksi, evlenme oranı, tüketici güven endeksi ve ekonomi güven endeksi yer almaktadır.

Çalışmanın amacı, veri madenciliği teknikleri olan yapay sinir ağları, destek vektör makineleri ve çoklu doğrusal regresyon yöntemlerini kullanarak satış tahminleri yapmaktır. Ayrıca, yapılan tahminler birbirleriyle karşılaştırılarak aralarındaki farklılıklar incelenmiş ve satış süreçlerini optimize ederek maksimum kâr sağlamak ve zaman kaybını en aza indirmek hedeflenmiştir. Veriler, Weka programı aracılığıyla test edilmiş ve satış tahminleri üç farklı yöntemle ayrı ayrı değerlendirilerek sonuçlar karşılaştırılmıştır (Demirer-Polat, 2022).

Acı ve Doğanşoy (2022), yerel bir süpermarketin 2019-2020 yıllarına ait iki yıllık e-ticaret verilerini (web sitesi ve mobil uygulama) ve ürün satışını etkileyen faktörleri (tatil günleri, TÜFE değeri ve işsizlik oranı) göz önünde bulundurarak, belirli ürün kategorileri için talep tahminleri yapmışlardır. En doğru talep tahmin modelini elde etmek amacıyla, altı farklı yapay zeka algoritmasından (Derin Öğrenme, Yapay Sinir Ağları, Gauss Süreç Regresyonu, Regresyon Ağacı, Destek Vektör Regresyonu ve Topluluk Öğrenme) oluşan yirmi dört farklı yöntem kullanılmıştır. Tahmin sonuçları, korelasyon katsayısı (R), Kök Ortalama Kare Hata (KOKH) ve Ortalama Mutlak Hata (OMH) gibi kriterlerle değerlendirilmiştir. En yüksek başarı, Long Short-Term Memory (LSTM) kullanılarak elde edilmiştir (KOKH=0,0353; OMH=0,0164; R=0,9742). Bu bulgular, doğru ürün tedarikinin sağlanmasında, satış kampanyalarında ve pazarlama stratejilerinde kullanılarak e-perakende satışlarının verimliliğini artırabilir (Acı ve Doğanşoy, 2022).

Peköz ve İnkaya (2023), COVID-19 pandemisi sırasında bir üçüncü parti lojistik firmasındaki talep tahminini doğru bir şekilde yapmak amacıyla bir çalışma gerçekleştirmişlerdir. Çalışmada, firma tarafından Haziran 2020 ile Aralık 2020 arasında kaydedilen sevkiyat verileri incelenmiş ve tahmin problemi tek değişkenli bir zaman serisi problemi olarak ele alınmıştır. Araştırma kapsamında, derin öğrenme temelli bir talep tahmini modeli önerilmiştir. Bu modelde, evrişimli sinir ağı (CNN) ve uzun kısa dönem hafıza (LSTM) ağı birleştirilerek hibrit bir model oluşturulmuştur. CNN, özniteliklerin çıkarılmasını sağlarken, LSTM ağı uzun dönemli bağımlılıkların yakalanmasına olanak tanımaktadır. Bu model, hibrit CNN-LSTM olarak adlandırılmaktadır. Önerilen modelin tahmin başarısı, klasik tahmin yöntemleri ve diğer makine öğrenmesi ile derin öğrenme teknikleriyle karşılaştırılmış ve değerlendirilmiştir. Tüm yöntemlerin parametre değerleri deneysel olarak belirlenmiştir. Deneysel bulgulara göre, hibrit CNN-LSTM modelinin tahmin performansı diğer yöntemlerden daha yüksek olmuştur. Bu yaklaşım, lojistik talebinin doğru bir şekilde tahmin edilmesine olanak

tanılarak, işgücü ve kaynak planlaması süreçlerinde önemli bir katkı sağlamaktadır (Peköz ve İnkaya, 2023).

Erol ve İnkaya (2024), satış tahmini problemini çözmek için çeşitli derin öğrenme yöntemlerinin performanslarını karşılaştırmışlardır. Çalışmada, derin sinir ağı (DNN), derin otokodlayıcı (Deep AE), evrişimli sinir ağı (CNN), tekrarlayan sinir ağı (RNN), uzun kısa süreli bellek ağı (LSTM), çift yönlü LSTM (Bi-LSTM), kapılı tekrarlayan birim (GRU), CNN-LSTM ve evrişimli LSTM (ConvLSTM) gibi farklı yöntemler kullanılmıştır. Deneysel çalışmalar, çeşitli sektörlerden alınan satış verileri üzerinde gerçekleştirilmiş ve hiperparametre optimizasyonu yapılmıştır. Elde edilen sonuçlar, tahmin doğruluğu ve eğitim süreleri açısından karşılaştırılmış ve istatistiksel anlamlılık analizi yapılmıştır. Çalışmanın bulguları, LSTM ve GRU modellerinin yüksek tahmin doğruluğu sağladığını, CNN modelinin ise eğitim sürelerini önemli ölçüde kısalttığını göstermiştir (Erol ve İnkaya, 2024).

Ürün satışını veya talebini belirleyen faktörler temelinde yapılan çalışmalar incelenmiştir. Yapılan çalışmalarda kullanılan, satışlara etki eden ekonomik faktörler olarak genellikle, ülke milli geliri, sanayi üretim endeksi, ekonomi güven endeksi, enflasyon, döviz kuru, sektörün üretimleri, ürün fiyatları ve geçmiş satış verileri olduğu görülebilir.

4. Yöntem

Bu çalışmada OLED tv satışlarındaki sipariş miktarlarının tahmini güçlü makine öğrenmesi yöntemlerinden Rastgele Orman yöntemi ile gerçekleştirilmiştir.

4.1. Rastgele Orman Yöntemi

Rastgele orman (RF), bir topluluk öğrenme algoritması olarak tanımlanabilir (Breiman, 2001). Bu algorithmada, her karar ağacı düğümünde en uygun bölme, tüm elemanlar arasından rastgele seçilen bir alt küme üzerinden belirlenir. Düğümler arasındaki en iyi bölmenin sağlanabilmesi için ise Gini indeksi kullanılır. N sınıf içeren bir veri seti için Gini indeksi, Eşitlik 1'de gösterildiği gibi hesaplanır.

$$Gini(t) = 1 - \sum_j [p(j|t)]^2 \quad (1)$$

Formülasyondaki j ve t sırasıyla sınıfı ve düğümü temsil etmektedir. Ayrıca $p(j|t)$, t düğümündeki j sınıfının göreceli olasılığını ifade etmektedir. Gini indeksi, aynı zamanda değişkenlerin homojenliği hakkında önemli bilgiler sunar. Gini indeksinin sıfıra yakın olması, verinin daha homojen olduğunu

gösterirken, yüksek bir değer ise verinin heterojenliğe eğilimli olduğunu ifade eder (Ekmekcioğlu vd., 2020).

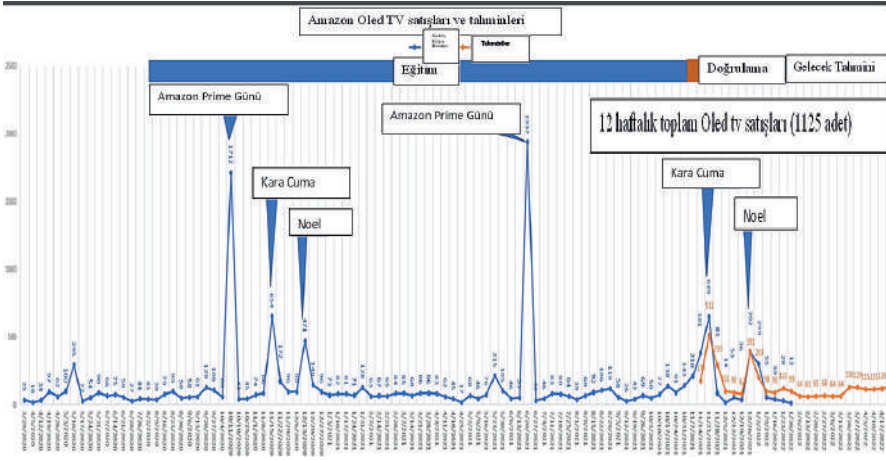
RF algoritması, örneklem büyüklüğü ve ağaç sayısı gibi hiperparametreler içerir. Kullanıcı, karar ormanını N adet ağaçla kurar ve böylece her sınıflandırma işleminde veri seti N farklı karar ağacından geçirilir. Ayrıca, eğitim ve test setlerinin belirlenmesi RF yöntemi için kritik bir aşamadır (Yu ve Abdel-Aty, 2014). Bir RF modelinin ortalama hata oranı hesaplanır ve hata oranı en düşük olan ağaç, en yüksek ağırlıkla değerlendirilir. Hata oranı en yüksek olan ağaca ise en düşük ağırlık verilir. Genel olarak, her ağacın bir oyu bulunduğu için, nihai tahmin ve en uygun sınıflandırıcıyı belirlemek amacıyla majority voting yapılır (Koyuncu ve Ekmekcioğlu, 2024).

5. Uygulama

Bu çalışma ile dünyanın en büyük e-ticaret ve perakende şirketlerinden biri olan Amazon'un tv satışları tahmin edilmeye çalışılmış ve bu tahminler ışığında üretim için gerekli stok maliyetlerinin planlanması hedeflenmiştir. Firmaların kar etme amacı ile maliyetlerini minimum da tutmaları gerekmektedir. Çalışmadaki tüm veri seti 29.03.2020 ile 30.01.2022 tarihleri arasındaki haftalık OLED tv satışlarından oluşmaktadır. Bu tarihler arasında satışların zirve yaptığı bazı özel günler ise Tablo 1'de verilmiştir.

Tablo 1. Bazı Özel Günler

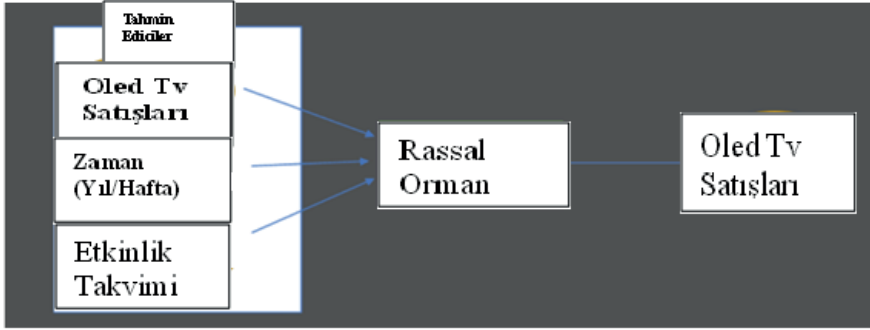
Tarih	Olay
11.10.2020	Amazon Prime Günü
15.11.2020	Kara Cuma
13.12.2020	Noel
20.06.2021	Amazon Prime Günü
21.11.2021	Kara Cuma
25.12.2021	Noel



Şekil 1. 29.03.2020 ile 30.01.2022 tarihleri arasındaki OLED TV sipariş edilen miktarlar ve tahminler

Şekil 1’de Tablo 1’de verilen bazı önemli günlerde OLED TV satışlarında sipariş edilen miktarlarda zirve yaptığı görülmektedir. Bu nedenle firmaların doğru stok miktarlarını ellerinde bulundurarak sipariş edilen miktarları karşılamaları oldukça önemlidir.

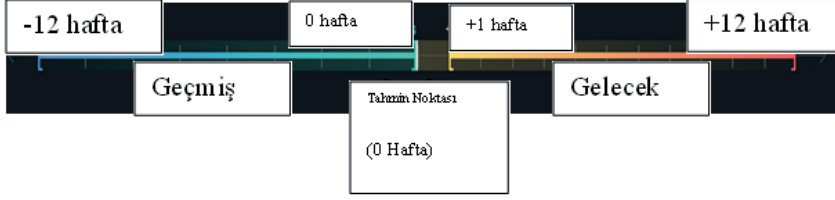
Bunun gerçekleşebilmesi için ise iyi bir planlama ve bir sonraki yılı yüksek oranda tahmin ederek yol izlemek önemlidir. Şekil 2’de OLED TV satış tahmini için oluşturulan tahmin modeli verilmiştir.



Şekil 2. Tahmin Modeli (OLED TV satışları)

Çalışmada 14.11.2021 ile 30.01.2022 tarihleri arasında 12 haftalık OLED tv satışları için sipariş edilen miktarlar kullanılmıştır. Bu tarihler arasında kara cuma ve Noel haftaları olması nedeniyle özellikle seçilmiştir. Gelecek 12 hafta için ise 06.02.2022 ile 24.04.2022 arasında satışlar tahmin

edilmeye çalışılmıştır. Şekil 3’de Amazon OLED tv satışlarında kullanılan veri setinin tahmin zamanı verilmiştir.



Şekil 3. OLED TV satış verileri için tahmin zamanı

Çalışmada Rastgele Orman yöntemi ile elde edilmiş olan 12 haftalık tahmin sonuçları Tablo 2’de verilmiştir.

Tablo 2. OLED TV satış tahminleri

Sipariş Tarihi	Sipariş Edilen Miktar	Tahminler Toplamı	Ortalama Mutlak Yüzde Hata (MAPE) (Rastgele Orman Modeli)	Temel Tahminler Toplamı	Ortalama Mutlak Yüzde Hata (MAPE) Temel Model
14.11.2021	381	179	53%	69	82%
21.11.2021	649	512	21%	50	92
28.11.2021	83	293	253%	77	7
05.12.2021	14	101	621%	138	886
12.12.2021	53	90	70%	91	72
19.12.2021	36	82	128%	143	297
26.12.2021	392	391	0%	210	46
02.01.2022	299	203	32%	381	27
09.01.2022	55	98	78%	381	593
16.01.2022	39	90	131%	83	113
23.01.2022	28	122	336%	14	50
30.01.2022	12	99	725%	53	342
Toplam	2041	2260	11%	1690	17%

Modellerin öngörü doğruluğunun (tahmin performansının) ölçümünde kullanılan istatistikler; en küçük değere sahip hata kareler ortalaması (MSE),

ortalama hata (ME), hata kareler ortalamasının karakökü (RMSE), hatanın mutlak ortalaması (MAE), ortalama yüzde hata (MPE) ve ortalama mutlak yüzde hata (MAPE) şeklindedir. Ancak hangi istatistiğin kullanılması gerektiği konusunda belirli bazı kriterler takip edilmektedir.

Eğer hata değerleri birbirine yakın büyüklüklerdeyse, MSE istatistiği kullanılabilir. Ancak tahmin sonuçlarından biri ya da birkaçı, ortalamadan önemli ölçüde sapma gösteriyorsa, MSE uygun bir ölçüt olmayabilir. Çünkü MSE, hataları karesini alarak değerlendirdiği için büyük sapmalara aşırı duyarlıdır ve bu da sonuçları olduğundan daha yüksek gösterebilir. Bu tür durumlarda, daha dengeli sonuçlar sunan MAE tercih edilebilir. Öte yandan, bir modelin yansız (biasız) olup olmadığını değerlendirmek gerekirse, tahminlerin sistematik olarak gerçek değerlerin altında ya da üstünde kalıp kalmadığına bakılır. Bu tür analizlerde MPE kullanılır. Ayrıca, tahmin yöntemlerinin kullandığı değerlerin birimlerinde farklılık varsa örneğin biri ham verilerle, diğeri logaritmik dönüşüm yapılmış verilerle çalışıyorsa karşılaştırma için MAPE gibi oransal ölçütlere başvurulması daha uygundur (Yılmaz ve Çilengiroğlu, 2022).

MAPE istatistiği, farklı birimlere sahip tahmin modelleri arasında karşılaştırma yaparken ortaya çıkabilecek olumsuzlukları ortadan kaldırma avantajı sağlar. Öngörü hatalarını yüzde cinsinden ifade etmesi, MAPE'yi diğer istatistiksel ölçütlerden ayıran ve tek başına anlam taşıyan önemli bir özelliğidir. Bu yönüyle, MAPE birçok durumda tercih edilen bir performans göstergesi olarak öne çıkar. Değerlendirme açısından bakıldığında, MAPE değeri %10'un altında olan modeller "yüksek doğruluk" düzeyine sahip olarak kabul edilirken, %10 ile %20 arasında kalan modeller ise genellikle "doğru tahmin yapan" modeller şeklinde sınıflandırılır (Yılmaz ve Çilengiroğlu, 2022). MAPE değeri Eşitlik 2'de verildiği hesaplanmaktadır.

$$MAPE = \frac{\sum_{t=1}^n \frac{|u_t|}{Y_t}}{n} * 100 \quad (2)$$

Tablo 2'de hesaplanan MAPE değeri %11 değeriyle %10-%20 değerleri arasında kaldığı için doğru tahmin edilen model olarak öngörülmektedir.

Çalışmada henüz verisi elde edilmemiş olan 06.02.2022 tarihleri ile 24.04.2024 tarihleri arasındaki Rastgele Orman yöntemiyle elde edilen 12 haftalık OLED TV sipariş miktarlarını tahmin sonuçları ise Tablo 3'de verilmiştir.

Tablo 3: Gelecek 12 Haftalık Satış Tahminleri

Sipariş Tarihi	Tahminler Toplamı
06.02.2022	64
13.02.2022	61
20.02.2022	65
27.02.2022	68
06.03.2022	64
13.03.2022	64
20.03.2022	130
27.03.2022	129
03.04.2022	115
10.04.2022	115
17.04.2022	120
24.04.2022	130

Çalışmada son olarak model kullanılarak Rastgele Orman metodunu kullanarak ne kadar hasılat farkı elde edeceklerine ilişkin Tablo 4’de özet bir tahmin değerleri tablosu hazırlanmıştır.

Tablo 4. Özet Tahmin Değerleri

	Sipariş Edilen Miktar	Model Tahmin Miktarı	Göz atma sayısı	Stok Dışı Görünümler	Ürün Detay Sayfası (PDP)	Kayıp Miktar	Ortalama Satış Fiyatı
14.11.2021-30.01.2022	2041	2260	605.706	119.143	0,34%	405	€ 1091

Tablo 4’de 119.143 tane stok dışı görünüm olduğu görülmüştür. Bu değeri 0,34 PDP değeri ile çarptığımızda kayıp miktar değerini elde ederiz.

$$\text{Model kullanılmadan kayıp miktar} = 119.143 * 0,0034 = 405$$

$$\text{Model kullanılarak kayıp miktar} = 2041 + 405 - 2260 = 186$$

$$\text{Gelir Farkı} = (405 - 186) * € 1091 = € 238.929$$

Tablo 4’deki değerleri hesaplayarak Rastgele Orman modeli yöntemini kullanarak Gelir farkı değerini elde ettiğimizde € 238.929 daha fazla hasılat elde edeceğimizi görmekteyiz. Bu nedenle firmalar ellerindeki stok miktarlarını planlarken zaman serileri ve makine öğrenmesi gibi yöntemleri kullanmalarının kendileri için oldukça önemli olduğu görülmektedir. Bununla birlikte tahminleme yaptığımız 14.11.2021 ile 30.01.2022 tarihleri arasında 21.11.2021 tarihinin kara cuma ve 25.12.2021 tarihinin ise noel haftası olduğu görülmektedir. Bu tarihlerden Tablo 2’de kara cuma

gününde 649 adet OLED tv siparişi verildiği ve modelin 512 adet tahminde bulunduğu görülmektedir. Diğer yandan Noel haftaları olan 26.12.2021 tarihinde 392 adet OLED tv siparişi verildiği modelin ise 391 adet olarak tahminde bulunduğu ve 02.01.2022 tarihinde ise 299 adet sipariş verildiği ve modelin 203 adet OLED tv siparişi tahmin ettiği gözlemlenmiştir. Özellikle bunun gibi özel günlerde siparişlerin zirve yapması nedeniyle doğru bir modelleme tahmini ile toplamda firmaların ellerindeki stok miktarlarını doğru ayarlayarak daha çok hasılat edebilecekleri görülmektedir.

6. Tartışma ve Sonuç

Son yıllarda giderek daha fazla önem kazanan ve “yılın özel günleri” olarak adlandırılan Yılbaşı, Sevgililer Günü, Anneler Günü, Ramazan ve Kurban Bayramları ile Kara Cuma gibi dönemler, taşıdıkları kültürel ve dini anlamların ötesinde, tüketici davranışlarını önemli ölçüde etkilemektedir. Bu özel günlerde, işletmelerin yürüttüğü kampanya ve pazarlama stratejileri sayesinde talepte belirgin artışlar, hatta zaman zaman talep patlamaları yaşanabilmektedir. Özellikle perakende ve e-ticaret sektörlerinde bu dönemsel hareketlilik, satış hacimlerinin ciddi şekilde yükselmesine katkı sağlamaktadır. Bu bağlamda, talep tahmini işletmeler açısından yalnızca operasyonel değil, aynı zamanda stratejik bir gereklilik hâline gelmiştir. Pazardaki konumunu korumak, rekabet avantajı elde etmek ve sürdürülebilir büyümeyi sağlamak isteyen firmalar için doğru talep öngörülleri, üretim süreçlerinin temelini oluşturur. Etkili bir üretim planlaması, hammadde, yarı mamul, insan kaynağı, makine-teçhizat ve finansman gibi unsurların zamanında ve doğru şekilde belirlenmesini sağlar. Aynı zamanda, talep doğrultusunda üretimin yapılması, gereksiz stok birikimini önleyerek kaynakların daha verimli kullanılmasına olanak tanır. Özetle, isabetli tahminlerle üretim, satış ve stok yönetimi arasında sağlıklı bir denge kurulabilir.

Bu çalışmanın temel amacı, işletmelerin üretim planlaması süreçlerinde stratejik bir bakış açısı geliştirmelerine katkı sağlamaktır. Bu doğrultuda geliştirilen tahmin modeli, firmaların insan kaynağı, ekipman ve diğer üretim kaynaklarını önceden planlamalarına yardımcı olmakta; elde edilen çıktılar, üreticiler, araştırmacılar ve pazarlama profesyonelleri için karar destek mekanizması işlevi görmektedir. Satış miktarlarının önceden öngörülebilmesi, firmalara hammadde maliyetlerini optimize etme, üretimi daha düşük maliyetle gerçekleştirme ve stok yönetimini etkinleştirme gibi önemli avantajlar sunar. Ayrıca, doğru tahmin sonuçlarıyla üretim kapasitesinin artırılma potansiyeli değerlendirilebilir, bu da yeni müşteri segmentlerine ulaşma ve satış hacmini genişletme imkânı yaratır. Türkiye’de

giderek yaygınlaşan lisanslı depoculuk uygulamaları açısından da bu tür tahminlerin yönlendirici etkisi olabilir.

Tüm bu kazanımlar neticesinde firmalar, minimum maliyetle maksimum kâr elde etmeye yönelik stratejiler geliştirerek sürdürülebilirliklerini artırabilirler. Çalışma bulguları, doğru modelleme teknikleri kullanıldığında kârlılığın önemli ölçüde artırılacağını göstermektedir. Nitekim üretim planlamasının etkin bir şekilde gerçekleştirilebilmesi, büyük ölçüde güvenilir talep tahminlerine bağlıdır. Günümüzde, doğru öngörüler işletmelerin mevcut kaynaklarını daha etkin ve sürdürülebilir biçimde kullanmaları açısından kritik bir rol oynamaktadır. Özellikle son yıllarda artan ekonomik belirsizlikler, üretim planlamasının stratejik değerini daha da artırmıştır. Tüm bu kazanımların sonucunda, firmalar minimum maliyetle maksimum kâr elde etmeye yönelik stratejiler geliştirerek sürdürülebilirliklerini artırabilir. Çalışmada elde edilen bulgular, doğru modelleme teknikleri kullanıldığında kârlılığın önemli ölçüde artırılacağını ortaya koymaktadır. Bu doğrultuda, işletmeler geleneksel tahmin yöntemlerinin ötesine geçerek, daha isabetli sonuçlar sunabilen analitik ve sayısal yöntemlere yönelmektedir. Bu yöntemler arasında öne çıkan makine öğrenmesi temelli analizler, talep tahmini sürecinde yaygın olarak kullanılmaya başlanmıştır.

Bu çalışmanın uygulama bölümünde, makine öğrenmesi yöntemlerinden biri olan Rastgele Orman (Random Forest) algoritması kullanılarak bir satış tahmini gerçekleştirilmiştir. Elde edilen model, yalnızca belirli bir sektöre değil, farklı sektörlerdeki firmalara da uygulanabilecek niteliktedir. Ayrıca, diğer makine öğrenmesi algoritmalarının entegrasyonu ile modelin kapsamı daha da genişletilebilir. Bu bağlamda, yalnızca sınırlı uygulama alanlarına sahip olduğu düşünülen makine öğrenmesi yöntemlerinin, gerçek dünya problemlerine başarıyla uyarlanabileceği gösterilmiştir.

Gelecekte yapılacak çalışmaların, bu tür tahmin modellerini farklı sektörlerde test ederek çok daha geniş bir uygulama alanına yaymaları ve endüstrinin ihtiyaç duyduğu öngörü modeli çalışmalarına katkı sunmaları beklenmektedir.

Kaynakça

- Acı, M., & Doğanşoy, G. A. (2022). Demand forecasting for e-retail sector using machine learning and deep learning methods, *Journal of the Faculty of Engineering and Architecture of Gazi University*, 37(3), 1325-1339.
- Aksel, F. (2000). *Regresyon Analizi ve Yapay Sınır Ağı Yöntemleri ile Uzun Dönem Yük Tahmini*. Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.
- Antipov, E. A., & Pokryshevskaya, E. B. (2020). Interpretable machine learning for demand modeling with high-dimensional data using Gradient Boosting Machines and Shapley values. *Journal of Revenue and Pricing Management*, 19(5), 355–364.
- Atalay, M., & Çelik, E. (2017). Büyük Veri Analizinde Yapay Zeka Ve Makine Öğrenmesi Uygulamaları. *Mehmet Akif Ersoy Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 9(22), 155-172.
- Azadeh, A., Neshat, N., Mardan, E., & Saberi, M. (2013). Optimization of steel demand forecasting with complex and uncertain economic inputs by an integrated neural network–fuzzy mathematical programming approach. *The International Journal of Advanced Manufacturing Technology*, 65, 833-841.
- Bayuk, M. N. (2009). Özel Gün ve Haftaların Tüketim Üzerindeki Etkileri. *Kamu İş Dergisi*, 11(1), 185-197.
- Böge, E. (1994). *Tüketici Davranışı ve Karar Verme*. Ankara: Gazi Yayınları.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.
- Campbell, M., Hoane, A. J. & Hsu, F. (2002). Deep blue, *Artificial Intelligence*, 134, 57–83.
- Carlson, R. L., & Umble, M. M. (1980). Statistical Demand Functions for Automobiles and Their Use for Forecasting in an Energy Crisis. *The Journal of Business*, 53(2), 193–204.
- Catal, C., Ece, K., Arslan, B., Akbulut, A. (2019). Benchmarking of Regression Algorithms and Time Series Analysis Techniques for Sales Forecasting. *Balkan Journal of Electrical and Computer Engineering*, 7(1), 20-26.
- Chen, C. Y., Lee, W. I., Kuo, H. M., Chen, C. W., & Chen, K. H. (2010). The study of a forecasting sales model for fresh food. *Expert Systems with Applications*, 37(12), 7696-7702.
- Demirbaş, F. P. (2011). *Kombi Üretiminde Talep Tahmin Yöntemlerinin Uygulanması*. Yüksek Lisans Tezi, Kocaeli Üniversitesi, Fen Bilimleri Enstitüsü, Kocaeli.
- Demirel, Ö. (2009). *ANFIS ve ARMA modelleri ile elektrik enerjisi yük tahmini*. Yüksek Lisans Tezi, Marmara Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

- Demirer-Polat, E. (2022). *Beyaz Eşya Sektöründe Satış Tahmini: Bir Veri Madencilği Uygulaması*, Yüksek Lisans Tezi, Uludağ Üniversitesi, Sosyal Bilimler Enstitüsü, Bursa.
- Dietterich, T. G. (1997). Machine-Learning Research. *AI Magazine*, 18(4), 97-136.
- Ecemiş, O. (2018). Model ağaç yöntemiyle satış tahmini: Paslanmaz çelik sektöründe bir uygulama. *Akademik Sosyal Araştırmalar Dergisi*, 6(84), 336-350.
- Efendigil, T., Önüt, S., & Kahraman, C. (2009). A decision support system for demand forecasting with artificial neural networks and neuro-fuzzy models: A comparative analysis. *Expert Systems with Applications*, 36(3), 6697-6707.
- Ekmekcioğlu, M., Başakın, E. E., & Özger, M. (2020). Tree-based nonlinear ensemble technique to predict energy dissipation in stepped spillways. *European Journal of Environmental and Civil Engineering*, 26(8), 3547-3565.
- El-Naqa, I., & Murphy, M. J. (2015). What is machine learning?. In: *Machine Learning in Radiation Oncology* (pp. 3-11). Springer, Cham.
- Eroğlu, A. H. (1996). Pazarlama Planlaması Sürecinde Satış Tahminleri Ve Çoklu Regresyon Ve Korelasyon Analizinin Etkinliği. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 1(1), 99-116.
- Erol, B., & İnkaya, T. (2024). Satış Tahmini için Derin Öğrenme Yöntemlerinin Karşılaştırılması. *Uludağ Üniversitesi Mühendislik Fakültesi Dergisi*, 29(2), 535-554.
- Ersen, M., Büyüklü, A. H., & Taşabat, S. E. (2022). Data Mining as a Method for Comparison of Traffic Accidents in Şişli District of Istanbul. *Journal of Contemporary Urban Affairs*, 6(2), 113- 141
- Fogel, S. O., & Thornton, C. G. (2008). What a hassle! consumer perceptions of costs associated with sales promotions, *Journal of Promotion Management*, 14, 31-44.
- Foot, K. D. (2021). A Brief History of Machine Learning. Available online: <https://www.dataversity.net/a-brief-history-of-machine-learning/> (accessed on: 15.09.2024).
- Gavcar, E., Şen, S., & AYTEKİN, A. (1999). Türkiye’de Kullanılan Kağıt-Karton Türlerinin Talep Tahminlerinin Belirlenmesi. *Turkish Journal of Agriculture and Forestry*, 23(2), 203 - 211.
- Grigorev, A. (2020). *Machine Learning Bookcamp MEAP V06* (A. Books (ed.)). Copyright 2020 Manning Publications. Available online: <https://www.manning.com/books/machine-learning-bookcamp> (accessed on: 17.09.2024).

- Helmini, S., Jihan, N., Jayasinghe, M., & Perera, S. (2019). Sales forecasting using multivariate long short term memory network models. *PeerJ PrePrints*, 7, c27712v1.
- Hicham, A. & Bouhorma, M. (2012). Hybrid intelligent system for sale forecasting using delphi and adaptive fuzzy back-propagation neural networks. *(IJACSA) International Journal of Advanced Computer Science and Applications*, 3(11), 122-130.
- Ingle, C., Bakliwal, D., Jain, J., Singh, P., Kale, P., & Chhajed, V. (2021) Demand forecasting: Literature review on various methodologies, *In 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 1-7.
- İnal, G. (2021). *Yapay zekâ, veri madenciliği ve büyük veri kavramlarına karşı okul yöneticilerinin algılarının belirlenmesi*. Yayınlanmamış Yüksek Lisans Tezi, Akdeniz Üniversitesi, Eğitim Bilimleri Enstitüsü, Antalya.
- İnik, Ö. & Ülker, E. (2017). Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri. *Gaziosmanpaşa Bilimsel Araştırma Dergisi*, 6(3), 85-104.
- Karaatlı, M., Helvacioğlu, Ö. C., Ömürbek, N., Tokgöz, G. (2012). Yapay Sinir Ağları Yöntemi ile Otomobil Satış Tahmini. *Uluslararası Yönetim İktisat ve İşletme Dergisi*, 8(17), 87-100.
- Karaca, Ş., Özbakır U., M. & Yemez, İ. (2018). Özel Gün İndirimlerinin Tüketicilerin Satın Alma Davranışları Üzerindeki Etkisinin İncelenmesi. *İğdır Üniversitesi Sosyal Bilimler Dergisi*, 16, 413-436.
- Karahan, M. (2011). *İstatistiksel Tahmin Yöntemleri: Yapay Sinir Ağları Metodu ile Ürün Talep Tahmini Uygulaması*, Doktora Tezi, Selçuk Üniversitesi, Sosyal Bilimler Enstitüsü, Konya.
- Keskin, A. (2025). Predicting and Reducing Patient Waiting Times in Dental Clinics Using Machine Learning: A Case Study from Türkiye. *Black Sea Journal of Engineering and Science*, 8(1), 243-248.
- Kılıç, F., Akkaya, M. R., & Memili, N. (2018). Yemekhane için Yapay Zeka Teknikleri Kullanımı ile Günlük Talep Tahmini. *Avrupa Bilim ve Teknoloji Dergisi*, 13, 65-71.
- Koch, R. (2022). History of Machine Learning—A Journey through the Timeline. Available online: <https://www.clickworker.com/customer-blog/history-of-machine-learning/> (accessed on: 15.09.2024).
- Koochakpour, K. & Tarokh, M. J. (2016). Sales budget forecasting and revision by adaptive network fuzzy base inference system and optimization methods. *Journal of Computer & Robotics*, 9(1), 25-38.
- Koyuncu, Z., & Ekmekcioğlu, Ö. (2024). İstanbul İçin CBS Tabanlı Makine Öğrenmesi İle Sel Duyarlılık Haritasının Oluşturulması. *Doğal Afetler Ve Çevre Dergisi*, 10(1), 1-15.

- Lennon, S. J., Johnson, K. K. P., & Lee, J. (2011). A perfect storm for consumer misbehavior: Shopping on Black Friday. *Clothing & Textiles Research Journal*, 29(2), 119–134. <https://doi.org/10.1177/0887302X11401907>
- Mucuk, İ. (1998). Pazarlama İlkeleri ve Örnek Olaylar. 9. Basım, İstanbul: Türkmen Kitabevi.
- Nacar, E. N., & Erdebilli, B. (2021). Makine Öğrenmesi Algoritmaları ile Satış Tahmini. *Endüstri Mühendisliği*, 32(2), 307-320.
- Pavlyshenko, B. M. (2019). Machine-Learning Models for Sales Time Series Forecasting. *Data*, 4(1), 1–11.
- Peköz, A. Z., & İnkaya, T. (2023). Derin öğrenme ile talep tahmini: Bir üçüncü parti lojistik firması için COVID-19 döneminde vaka analizi. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, 29(4), 331-339.
- Poyraz, İ. (2020). *Makine Öğrenmesi Algoritmaları ile Eczaneler için İlaç Talep Tahmini*, Yüksek Lisans Tezi, İstanbul Aydın Üniversitesi, Lisansüstü Eğitim Enstitüsü, İstanbul.
- Raghubir, P., Inman, J. J., ve Grande, H. (2004). The three faces of consumer promotions. *California Management Review*, 46, 23-42.
- Rincon-Patino, J., Lasso, E., & Corrales, J. C. (2018). Estimating Avocado Sales Using Machine Learning Algorithms and Weather Data. *Sustainability*, 10(10), 3498.
- Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 3(3), 210–229.
- Sathya, R., & Abraham, A. (2013). Comparison of supervised and unsupervised learning algorithms for pattern classification. *International Journal of Advanced Research in Artificial Intelligence*, 2(2), 34–38.
- Savaş, S., Topaloğlu, N., & Yılmaz, M. (2012). Veri madenciliği ve Türkiye’deki uygulama örnekleri. *İstanbul Ticaret Üniversitesi Fen Bilimleri Dergisi*, 21, 1-23.
- Shimp, T.A. (2003). *Advertising, promotion, and supplemental aspects of integrated marketing communications*, 6th ed., United States of America: Thomson SouthWestern, Boston, MA.
- Suthaharan, S. (2014). Big data classification: Problems and challenges in network intrusion prediction with machine learning. *ACM SIGMETRICS Performance Evaluation Review*, 41(4), 70-73.
- Tosunoğlu, E., Yılmaz, R., Özeren, E., Sağlam, Z. (2021). Eğitimde Makine Öğrenmesi: Araştırmalardaki Güncel Eğilimler Üzerine İnceleme. *Abmet Keleşoğlu Eğitim Fakültesi Dergisi*, 3(2), 178-199.
- Tozak, E. Y. (2021). *Veri Madenciliği Programları Kullanılarak Bir Tekstil Firmasının Satış Verilerinin Değerlendirilmesi*, Yüksek Lisans Tezi, Karabük Üniversitesi, Lisansüstü Eğitim Enstitüsü, Karabük.

- Türk, E., & Kıran, F. (2019). Yapay Sinir Ağları ile Talep Tahmini Yapma: Beyaz Eşya Üretim Planlama Örneği. *İstanbul Sabahattin Zaim Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 1(1), 30-37.
- Uraz, Ç. (1978). Temel pazarlama bilgileri. Ankara: Kalite Matbaası.
- Verstraete, G., Aghezzaf, E. H., & Desmet, B. (2020). A leading macroeconomic indicators' based framework to automatically generate tactical sales forecasts. *Computers and Industrial Engineering*, 139, 106169.
- Weng, T., Liu, W., & Xiao, J. (2019). Supply chain sales forecasting based on lightGBM and LSTM combination model. *Industrial Management and Data Systems*, 120(2), 265–279.
- Yemcz, İ., & Sağır, S. (2021). İndirim Kampanyalarının Tüketicilerin Satın Alma Davranışı Üzerindeki Etkisi: Black Friday Örneği. *Avrasya Sosyal Ve Ekonomi Araştırmaları Dergisi*, 8(2), 308-331.
- Yılmaz, H. (2018). *Antalya ili Konut Satışının Çoklu Doğrusal Regresyon ve Yapay Sinir Ağları ile Tahmini*, Yüksek Lisans Tezi, Akdeniz Üniversitesi, Sosyal Bilimler Enstitüsü, Antalya.
- Yılmaz, M. B., & Çilengiroğlu, Ö. V. (2022). Talep Tahminleme Değişkenlerinin Üssel Düzeltme Yöntemi ile Belirlenmesi. *Euroasia Journal of Mathematics, Engineering, Natural & Medical Sciences*, 9(22), 92–103.
- Yu, R., & Abdel-Aty, M. (2014). Analyzing crash injury severity for a mountainous freeway incorporating real-time traffic and weather data. *Safety Science*, 63, 50–56.
- Yücesan, M. (2018). YSA, ARIMA ve ARIMAX Yöntemleriyle Satış Tahmini: Beyaz Eşya Sektöründe bir Uygulama. *İşletme Araştırmaları Dergisi*, 10(1), 689–706.
- Yücesoy, M. (2011). *Temizlik Kağıtları Sektöründe Yapay Sinir Ağları ile Talep Tahmini*. Yüksek Lisans Tezi, İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

Nicel Karar Vermede Çok Kriterli Yaklaşımlar ve Makine Öğrenmesi Çalışmaları

Editör:

Dr. Öğr. Üyesi Hasan Aykut KARABOĞA