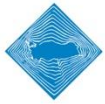


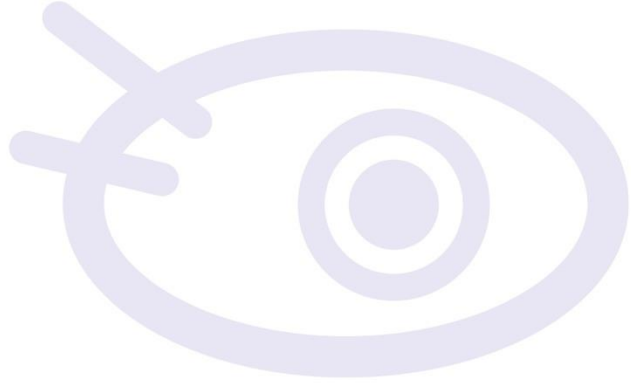
Proceedings Book



TURKISH STATISTICAL ASSOCIATION

international data science and statistics congress: **artificial intelligence**

15-17 October | Ankara



IDS
SS
2024
CC



IDSSC 2024 PROCEEDINGS BOOK

PUBLISHER:

Turkish Statistical Association

EDITOR:

Prof. Dr. Turhan Menteş

E-BOOK COVER and LAYOUT:

Fatih Cengiz Öz [foztasarim.com]

ISBN: 978-625-95613-0-1

IDS'2024

SPONSORS



Main



Panel



Student



Student



Student



Technology



Technology



Company



Company



Company



Training



Display



Display



Display



Display



GDPR



Legal



Web



Media

SUPPORTING ORGANIZATIONS



T.C. CUMHURBAŞKANLIĞI
DİJİTAL DÖNÜŞÜM OFİSİ



TÜRKİYE BİLİŞİM VAKFI



TÜRKİYE EKONOMİ KURUMU



Finansal Kurumlar Birliği



TRAI



Türkiye'nin Yapay Zeka Platformu



Türkiye Zeka Vakfı

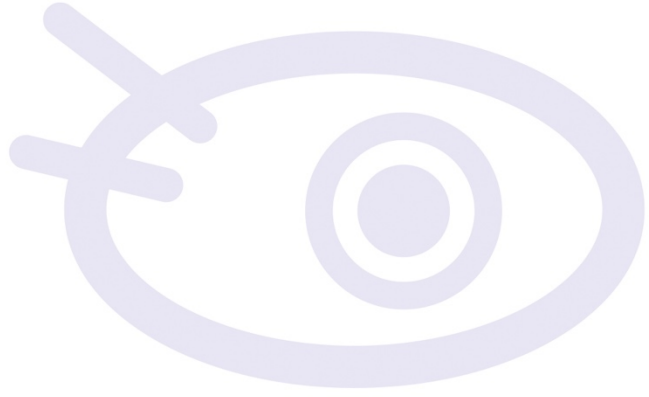


YAPAY ZEKA FABRİKASI

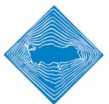


T.C. İÇİŞLERİ BAKANLIĞI

IDS
SS
2024CC



Committees



TURKISH STATISTICAL ASSOCIATION

HONORARY COMMITTEE

Dr. Fikri AKDENİZ

Dr. Orhan GÜVENEN

Dr. Ceyhan İNAL

Dr. Ömer GEBİZLİOĞLU

Dr. Cenap ERDEMİR

Dr. Ayşen APAYDIN

Dr. Bahadır ÇETİNEL

Belgin SAĞDIÇ

ADVISORY COMMITTEE

Rıfat HİSARCIKLIOĞLU

TOBB, The Union of Chambers and Commodity Exchanges of Türkiye; President

Prof. Dr. Faruk BİLİR

KVKK, Personal Data Protection Authority; President

İbrahim Ömer Gönül

Capital Markets Board of Turkey; Chairman

Erhan ÇETİNKAYA

TUIK, TURKSTAT; President

Prof. Dr. Necdet ÜNÜVAR

Ankara University; Rector

Prof. Dr. Muhammed ULUDAĞ

Galatasaray University; President

Prof. Dr. Musa YILDIZ

Gazi University; President

Prof. Dr. Mehmet Cahit GÜRAN

Hacettepe University; President

Prof. Dr. İsmail KOYUNCU

Istanbul Technical University; President

Prof. Dr. Ahmet YOZGATLIGİL

Middle East Technical University; President

Prof. Dr. Murat YÜLEK

Ostim Technical University; President

Prof. Dr. İhsan SABUNCUOĞLU

TED University; President

Prof. Dr. Fatih YILMAZ

Tokat Gaziosmanpaşa University; President

Furkan CİVELEK

Digital Transformation Office of the Presidency of Republic of Türkiye

Mehmet Ali TOMBALAK

Turkish Informatics Industry Association; Chairman

Faruk ECZACIBAŞI

Turkish Informatics Foundation; Chairman

Prof. Dr. Hasan KAZDAĞLI

Turkish Economic Institution; Chairman

Emrehan HALICI

Turkish Intelligence Foundation; Chairman

ORGANIZING COMMITTEE

Assoc. Prof. Dr. Yalçın TAHTALI
Tokat Gaziosmanpaşa University

Prof. Dr. Çağdaş Hakan ALADAĞ
Hacettepe University, TİD; Deputy Chairman

Prof. Dr. Adil DENİZLİ
Hacettepe University

EXECUTIVE COMMITTEE

Prof. Dr. Turhan MENTEŞ
TİD; Chairman

Prof. Dr. Çağdaş Hakan ALADAĞ
Hacettepe University, TİD; Deputy Chairman

Paşa YAŞAR
PyData Integrated Data Management Systems Inc.; Chairman

Prof. Dr. Altan ÇAKIR
ITU Artificial Intelligence Center Executive Board Member, AITR co-Chairman

Prof. Dr. Menemşe Gümüşdereliolu
Hacettepe University

Assoc. Prof. Dr. Yalçın Tahtalı
Tokat Gazi Osman Paşa University

Dr. Yüksel VARDAR
TİD; General Secretary

Levent KÜPELİ
TED University; General Secretary

Mustafa ERBİLLİ
KVKK, Personal Data Protection Authority; Head of Data Management Department

Assoc. Prof. Dr. Alaattin ECER
SPK; Head of Research Department

Assist. Prof. Mustafa GUZEL
Tokat Gazi Osman Paşa University

Lecturer Dr. Shiva SADIGHFARD
Tokat Gazi Osman Paşa University

Halil AKSU
Digitopia, TRAI; Founder

Remime KOÇOĞLU
ACCERT JSC.; Chairman

Hakan TARHAN
Datamind Informatics Inc.; Managing Partner

Nedim NAZLI
Tunesoft JSC.; Founding Partner

Hasan Çağlayan DÜNDAR
TOBB; The Union of Chambers and Commodity Exchanges of Türkiye; Advisor to the President

Fatih Cengiz ÖZ
Freelance Designer, Visual Editor

Levent UYANIKER
TİD; Press and Public Relations Representative

Süzülay HAZAR
TİD, Auditor

Dr. Tolga KASKATI
BYS GROUP; Co-founder

Oğuzhan AKYILDIRIM
TİD; Treasurer

Prof. Dr. Birdal ŞENOĞLU
TİD; Member

Prof. Dr. Ceylan YOZGATLIGIL
TİD; Member

Dr. Hilal GÜNEY
TİD; Member

Metin USLU
TİD; Member

ACADEMIC COMMITTEE

Academy of Economic Studies

Prof. Dr. Miruna Mazurencu Marinescu PELE

Ankara University

Prof. Dr. Olçay ARSLAN
Prof. Dr. Halil AYDOĞDU
Prof. Dr. Birdal ŞENOĞLU

Bilecik Şeyh Edebali University

Prof. Dr. Serpil TÜRKYILMAZ

Boğaziçi University

Prof. Dr. Necati ARAS

BYS Group

Assoc. Prof. Dr. Tolga KASKATI

Columbia University

Prof. Dr. Birol EMİR

Çanakkale Onsekiz Mart University

Dr. Berrin GÜLTAY

Dokuz Eylül University

Prof. Dr. Adil BAYKASOĞLU
Prof. Dr. Esin FİRUZAN
Prof. Dr. Burcu HÜDAVERDİ
Prof. Dr. Güçkan YAPAR

Ege University

Prof. Dr. Gözde ULUTAGAY

Erciyes University

Prof. Dr. Gökmen ZARARSIZ

Eskişehir Technical Üniversitesi

Prof. Dr. Yeliz Mert KANTAR
Prof. Dr. İlhan USTA
Assoc. Prof. Dr. Şükrü ACITAŞ

Faculty of Administration and National Security University

Dr. Sylwia SZYBOWSKA

Gazi University

Prof. Dr. Mustafa ALKAN
Prof. Dr. Akif BAKIR

Giresun University

Prof. Dr. Erol EĞRİOĞLU

Hacettepe University

Prof. Dr. Çağdaş Hakan ALADAĞ
Prof. Dr. Serpil AKTAŞ ALTUNAY
Prof. Dr. Selim ÇAĞLAR
Prof. Dr. Meral ÇETİN
Prof. Dr. Gıyasettin DEMİRHAN
Prof. Dr. Adil DENİZLİ
Prof. Dr. Menemşe GÜMÜŞDERELİOĞLU
Prof. Dr. Ergun KARAAĞAOĞLU
Prof. Dr. Ebru Akçapınar SEZER
Prof. Dr. Kasırğa YILDIRAK
Assoc. Prof. Dr. Harun ARTUNER
Assoc. Prof. Dr. Derya ERSEL
Assoc. Prof. Dr. Duygu İÇEN
Assoc. Prof. Dr. Yasemin KAYHAN
Assoc. Prof. Dr. Ayfer Ezgi YILMAZ

Higher School of Euroregional Economy

Iwonka FLOREK

Icahn School of Medicine

Assoc. Prof. Dr. Umut ÖZBEK

Invivyd Inc.

Dr. İlker YALÇIN

İskenderun Technical University

Dr. Baki ÜNAL

İstanbul Technical University

Prof. Dr. Altan Çakır

İstanbul Ticaret University

Assoc. Prof. Dr. Seda BAĞDATLI KALKAN

Karabük University

Prof. Dr. Canan HAMURKAROĞLU

Karadeniz Technical University

Prof. Dr. Orhan KESEMEN

Kırıkkale University

Prof. Dr. Güvenç ARSLAN

Koç University

Prof. Dr. Deniz YÜRET

Louisiana State University

Assoc. Prof. Dr. Evrim ORAL

Marmara University

Prof. Dr. Deniz İNAN
Prof. Dr. Ufuk YOLCU

Memorial Sloan-Kettering Cancer Center

Prof. Dr. Mithat GÖNEN

Mimar Sinan Fine Arts University

Prof. Dr. Barış AŞIKGİL
Prof. Dr. Gülay BAŞARIR KIROĞLU

Microsoft

Assoc. Prof. Dr. Erdal COSGUN

Muğla Sıtkı Koçman University

Assoc. Prof. Dr. Serdar DEMİR
Assoc. Prof. Dr. Eralp DOĞU

Necmettin Erbakan University

Prof. Dr. Aşır GENÇ

New Hampshire University

Asist. Prof. Dr. Burcu EKE RUBINI

Ondokuz Mayıs University

Prof. Dr. Yüksel TERZİ

Middle East Technical University

Prof. Dr. Ayşen AKKAYA
Prof. Dr. İnci BATMAZ
Prof. Dr. Özlem İLK DAĞ
Prof. Dr. Ceylan YOZGATLIGİL
Prof. Dr. Vilda PURUTÇUOĞLU
Prof. Dr. Pınar KARAGÖZ
Prof. Dr. İsmail Sengör ALTINGÖVDE
Prof. Dr. Cem İYİĞÜN
Prof. Dr. A. Sevtap SELÇUK KESTEL
Prof. Dr. Tuğba TEMİZEL TAŞKAYA
Assoc. Prof. Dr. Hande ALEMDAR
Assoc. Prof. Dr. Fulya GÖKALP YAVUZ

RMIT University

Dr. Haydar DEMİRHAN

Temple University

Prof. Dr. Recai YÜCEL

The University of Tennessee

Prof. Dr. Hamparsum BOZDOGAN

TOBB Economy and Technology University

Prof. Dr. Tahir HANALIOĞLU

Tokat Gaziosmanpaşa University

Assoc. Prof. Dr. Yalçın TAHTALI

Yıldız Technical University

Prof. Dr. Filiz KARAMAN

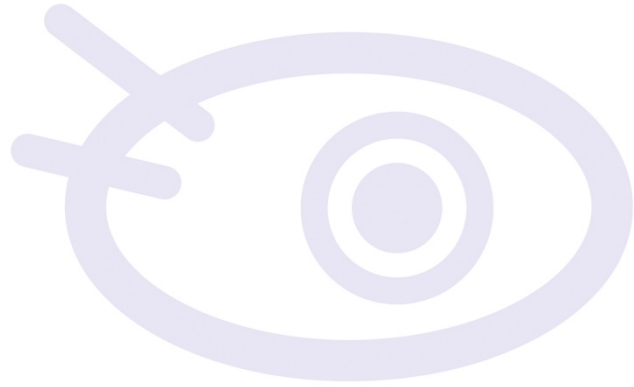
CONGRESS SECRETARIAT

Valör Kongre Organizasyonları

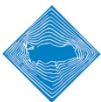
Turan Güneş Bulvarı Galip Erdem Caddesi 576. Sokak No.28 06550 Oran / Ankara
T. +90 312 491 88 88 F. +90 312 491 99 89 M. vebist@valor.com.tr



IDS
SS
2024CC



Honorably Invited Guest Person Presentation



TURKISH STATISTICAL ASSOCIATION

YAPAY ZEKA, NANOTEKNOLOJİ VE VERİ BİLİMİ DİNAMİKLERİ

Prof. Dr. Orhan Güvenen

Özet

Yapay zekâ, nanoteknoloji ve veri bilimi, bilim ve teknolojide derin etkiler yaratan üç temel disiplindir. **Yapay zekâ**, makine öğrenimi, derin öğrenme ve doğal dil işleme gibi alt dallarla büyük veri setlerini analiz ederek, genetik araştırmalar, klinik tanı ve bilimsel modellemelerde önemli çözümler üretir. Gelecekte duygusal zekâ, otonom sistemler ve karmaşık problemler için gelişmiş algoritmalarla daha da güçlenecek.

Nanoteknoloji, maddelerin atom ve molekül düzeyinde tasarımı, üretimi ve uygulamasını kapsayan bilim dalıdır. Hedefe yönelik ilaç dağıtımı, akıllı malzemeler ve enerji çözümleri gibi alanlarda devrimsel uygulamalara sahiptir. Özellikle sağlık ve enerji sektörlerinde nanopartiküller ve nano-elektronikler sayesinde yenilikçi çözümler üretiyor. **Veri bilimi**, büyük veri setlerini analiz etmek, anlamak ve bu verilerden bilgi üretmek amacıyla matematik, istatistik, bilgisayar bilimi ve alan uzmanlığı disiplinlerini kapsayan bilim dalıdır. Büyük veri analizi ve modelleme teknikleriyle bilimsel araştırmaları optimize eder, biyoinformatik, çevresel izleme ve sosyal dinamiklerin analizinde belirleyici etkiye sahiptir. Bu üç alanın birleşimi, sağlık, enerji ve çevre gibi önemli sektörlerde veriye dayalı, kişiselleştirilmiş ve daha etkin çözümler geliştirme olanağına sahiptir. Ancak bu gelişmeler, etik, güvenlik ve gizlilik gibi konuların sistem dinamiğinde değerlendirilmelerini gerektirir.

YAPAY ZEKA

Yapay zekâ (YZ), bilgisayarların ve yazılımların insan benzeri düşünme, öğrenme ve problem çözme yeteneklerine sahip olmasını sağlayan bir teknoloji alanıdır. Makine öğrenimi, derin öğrenme, doğal dil işleme gibi alt alanları kapsar ve verileri analiz ederek, kalıpları tanıyarak ve deneyimlerden öğrenerek çalışır (Himeur, 2022).

Bilime Katkıları

- Yapay zekâ, büyük veri setlerini işleyerek karmaşık sorunların çözümlenmesine yardımcı olur. Örnek, genetik araştırmalarında, gen dizilimlerini analiz ederek hastalıkların nedenlerini anlamaya katkı sağlar.
- Simülasyon ve Modelleme: Bilimsel simülasyonlar ve modellemelerde kullanılabilir. İklim değişikliği, biyolojik süreçler gibi karmaşık sistemleri modelleyerek tahminler yapılmasına olanak tanır.
- Keşif ve İnovasyon: Yeni ilaçların keşfi, malzeme bilimi alanlarda YZ potansiyel olarak yeni moleküllerin ve malzemelerin keşfine yardımcı olabilir.
- Otomasyon: Araştırma süreçlerini hızlandırarak tekrarlayan görevlerin otomasyonunu sağlar. Bu bilim insanlarının daha yaratıcı ve karmaşık sorunlara odaklanmasına olanak tanır.
- Klinik Tanı: Tıpta, hastalıkların tanısında ve tedavi planlamasına YZ destekli sistemler, doktorlara karar verme süreçlerinde yardımcı olur.
- Doğal Dil İşleme: Bilimsel literatürü taramak, özetlemek veya belirli bilgileri çıkarmak için YZ, metin analizi ve doğal dil işleme yeteneklerini kullanabilir.

Wang ve diğerlerinin belirttiği üzere yapay zeka (YZ), günümüzde bilimsel keşif sürecine giderek daha fazla entegre edilmekte ve araştırmaların daha hızlı ilerlemesine katkı sağlamaktadır. YZ, bilim insanlarının hipotezler geliştirmesine, deneyler tasarlamasına, büyük veri setlerini toplama ve analiz etmesine, geleneksel yöntemlerle ulaşılmaması zor olan yeni bakış açıları kazanmasına olanak tanımaktadır (Wang, 2023). Yapay zekâ, bilimsel araştırmaların hızlanmasına daha doğru sonuçlar elde edilmesine ve yenilikçi çözümlerin geliştirilmesine önemli katkılar sağlar. Ancak, etik ve güvenlik gibi konular da sistem dinamiğinde çok iyi değerlendirilmelidir.

Yapay Zekânın Geleceği

- Daha gelişmiş öğrenme yöntemleri: Mevcut makine öğrenimi yöntemlerinin ötesine geçerek daha az veriyle daha fazla öğrenme kapasitesine sahip sistemler geliştirilebilir.
- Kendi kendine öğrenme: Yapay Zekâ sistemleri deneyimlerden bağımsız olarak, kendi başlarına öğrenme yeteneklerini geliştirebilir.
- Duygusal ve sosyal zekâ: YZ, insan duygularını anlama ve buna uygun tepkiler verme yeteneği kazandıkça sosyal etkileşimlerde daha etkili olabilir.

- Karmaşık problem çözümü: Bilim ve mühendislik problemlerinin daha karmaşık hale gelmesiyle birlikte YZ bu tür sorunları çözmek için yeni yaklaşımlar geliştirebilir. Örnek, enerji, iklim değişikliği gibi küresel sorunların çözümünde etkili olabilir.
- Otonom sistemler: Otonom araçlar ve robotlar daha yaygın hale gelebilir. Bu hususta ulaşım ve sağlık hizmetleri gibi birçok sektörde çok önemli etki yapabilir.
- Kişiselleştirilmiş deneyimler: Eğitimden sağlık hizmetlerine kadar YZ'nin kişiselleştirilmiş çözümler sunma yeteneği artacaktır.
- Yaratıcılık ve sanat: YZ, sanat, müzik ve edebiyat gibi alanlarda daha etkili olabilir. Yeni sanat biçimlerinin gelişmesinde ve mevcut eserlerin yorumlanmasında yardımcı olabilir.

Bu beklenen ilerlemelerle etik, güvenlik, mahremiyet gibi konularda önem kazanmaktadır. Bu unsurların sistem dinamiğinde değerlendirilmesi, sağlıklı bir ilerleme için önemli olacaktır.

Yapay Zekânın Gelecekte İnsanlığa Etkileri

1. İş Gücü ve Ekonomi

- Otomasyon: Yapay zekâ, birçok sektörde görevleri otomatikleştirebilir. Bu, verimliliği artırırken bazı mesleklerin ortadan kalkmasına yol açabilir. Ancak, yeni iş alanları da ortaya çıkacak, örneğin yapay zekâ ve veri mühendisliği veya veri bilimi analizi gibi.
- İş Gücü Dönüşümü: İş gücünün yeni becerilere ihtiyaç duyması, eğitim sistemlerinin ve işgücü geliştirme programlarının dönüşümünü gerektirecek. Sürekli öğrenme ve adaptasyon, iş gücünün önemli bir özelliği haline gelecek.

2. Sağlık Hizmetleri

- Hızlı Teşhis: Yapay zekâ, büyük veri ve makine öğrenimi kullanarak hastalıkların erken teşhisinde önemli bir rol oynayabilir. Örneğin, görüntüleme verilerini analiz ederek kanser gibi hastalıkları daha hızlı tespit edebilir.
- Kişiselleştirilmiş Tedavi: Genetik veriler ve bireysel sağlık geçmişi kullanılarak kişiye özel tedavi planları oluşturulabilir. Bu, tedavi süreçlerinin etkinliğini artırır.
- Sağlık Yönetimi: YZ, hasta verilerini izleyerek sağlık durumlarını sürekli takip edebilir ve bu sayede hastalıkların önlenmesine yönelik erken uyarılar sağlayabilir.

3. Eğitim

- Kişiselleştirilmiş Öğrenme: Öğrencilerin öğrenme yöntemlerine ve hızına göre uyarlanan eğitim programları geliştirebilir. Bu, öğrenme sürecini daha etkili ve kapsayıcı hale getirir.
- Erişim Artışı: Online eğitim ve yapay zekâ destekli platformlar, eğitim materyallerine erişimi artırarak daha fazla insana ulaşabilir. Bu, eğitimde fırsat eşitliğini destekleyebilir.

4. Etik ve Güvenlik

- Veri Gizliliği: Yapay zekâ uygulamaları, büyük miktarda kişisel veri toplayabilir. Bu durum, veri güvenliği ve mahremiyet konularında endişelere yol açmaktadır.
- Karar Alma Süreçleri: YZ sistemlerinin karar alma süreçlerinin şeffaf olmaması, adalet ve eşitlik gibi etik sorunlar gündeme getirebilir. Bu nedenle, YZ'nin geliştirilmesi ve kullanılması sırasında etik kuralların belirlenmesi önemlidir.
- Manipülasyon ve Yanlış Bilgi: Yapay zekâ, sahte içerikler oluşturma kapasitesine sahip olduğu için, yanlış bilgi yayılmasına neden olabilir. Bu durum, toplumsal güveni sarsabilir.

Bainbridge'in belirttiği üzere Nanobilimdeki ilerlemeler, diğer teknolojilerle daha fazla bağlantı kurulmasına olanak sağladıkça, nanoteknoloji ile ilgili etik meseleler önümüzdeki yıllarda giderek daha karmaşık hale gelecektir (Bainbridge, 2005).

5. Sosyal Etkiler

- Toplumsal Değişim: Yapay zekâ, sosyal dinamikleri etkileyerek bireylerin yaşam tarzlarını ve alışkanlıklarını değiştirebilir. Örneğin, YZ destekli hizmetlerin artışı, insan etkileşimini azaltabilir.

- Eşitsizlik: Teknolojiye sahip olanlar ve kullanabilenlerle, sahip olamayanlar ve kullanamayanlar arasındaki fark dijital uçurumun derinleşmesine yol açar; bu durum, toplumlar arasında ekonomik ve sosyal eşitsizlikleri artırabilir.

Yapay Zekânın Gelecekte İnsanlığa Olumsuz Etkileri Olabileceği Alanlar

1. İşsizlik ve Ekonomik Eşitsizlik

- Otomasyonun Yaygınlaşması: Birçok sektörde yapay zekânın iş gücünü otomatikleştirmesi, düşük nitelikli işlerin kaybına yol açabilir. Bu, işsizlik oranlarını artırabilir ve ekonomik eşitsizliği derinleştirebilir.
- Yeni Becerilere İhtiyaç: İş gücünün yeni teknolojilere uyum sağlaması zor olabilir. Eğitim sistemlerinin yeterince hızlı dönüşmemesi, bazı bireylerin iş bulma şansını azaltabilir.

2. Veri Güvenliği ve Gizlilik

- Veri İhlalleri: Yapay zekâ uygulamaları, büyük miktarda kişisel veri topladığı için, veri güvenliği ihlalleri büyük sorunlar yaratabilir. Bu tür durumlar, bireylerin gizliliğini tehdit edebilir.
- Kişisel Bilgilerin Kötüye Kullanımı: Toplanan verilerin kötü niyetli kişilerce kullanılması, kimlik hırsızlığı gibi suçları artırabilir.

3. Etik Sorunlar

- Karar Alma Süreçleri: Yapay zekâ sistemlerinin karar alma süreçlerinin şeffaf olmaması, adaletsizlik ve ayrımcılık gibi etik sorunlara yol açabilir. Örneğin, algoritmaların taraflı verilere dayanması durumunda belirli gruplar dezavantajlı hale gelebilir.
- Manipülasyon: Yapay zekâ, bilgi manipülasyonu ve sahte içerik üretimi için kullanılabilir. Bu, kamuoyunu yanıltabilir ve toplumsal huzursuzluğa neden olabilir.

4. Güvenlik Sorunları

- Otonom Silahlar: Yapay zekâ destekli otonom silahların kullanımı, savaş ve çatışmalarda kontrolsüz güç kullanımına yol açabilir. Bu tür teknolojiler, uluslararası güvenliği tehdit edebilir.
- Siber Saldırıları: Yapay zekâ, siber saldırıların daha karmaşık ve etkili hale gelmesine olanak tanıyabilir. Bu durum, kritik altyapılara yönelik tehditleri artırır.

5. Toplumsal Yalnızlaşma

- İnsan Etkileşiminin Azalması: YZ tabanlı sistemlerin artışı, insan etkileşimini azaltabilir. Bu, sosyal ilişkilerin zayıflamasına ve yalnızlık gibi psikolojik sorunların artmasına neden olabilir.

6. Kontrol Kaybı

- Yapay zekânın insan kararlarını asacak şekilde otonom hale gelmesi, kontrol kaybı riskini artırabilir. Bu durum, çeşitli senaryolarda beklenmedik sonuçlara yol açabilir.

NANOTEKNOLOJİ

Nanoteknoloji, maddelerin atom ve molekül düzeyinde (yaklaşık 1 ile 100 nanometre arası) tasarımı, üretimi ve uygulamasını içeren bir bilim dalıdır. Bu alanda geliştirilen malzemeler, olağanüstü özellikler sergileyebilir; örneğin, daha güçlü, hafif, iletken veya reaktif hale gelebilirler (Güvenen, 2009, 2015).

Bilime Katkıları

- Malzeme Bilimi: Nanoteknoloji, daha hafif ve dayanıklı malzemelerin geliştirilmesine olanak tanır. Örneğin, karbon nanotüpleri ve grafen, yüksek dayanıklılık ve iletkenlik özellikleriyle yeni uygulama alanları sunar.
- Tıp ve Sağlık: Nanoteknoloji, hedefe yönelik ilaç dağıtımı, kanser tedavisi ve görüntüleme tekniklerinde devrim yaratmaktadır. Nanopartiküller, ilaçların doğrudan kanser hücrelerine ulaşmasını sağlayarak yan etkileri azaltabilir.
- Enerji: Yenilenebilir enerji kaynaklarının verimliliğini artırmak için nanoteknoloji kullanılabilir. Örneğin, güneş panellerinde daha etkili ışık emilimi sağlayan nano materyaller geliştirilmiştir.
- Elektronik: Nanoteknoloji, daha küçük ve daha güçlü elektronik bileşenlerin geliştirilmesine olanak tanır. Bu, bilgisayar ve iletişim teknolojilerinde verimliliği arttırabilir.

- Çevre: Nanoteknoloji, su arıtma, hava temizleme ve atık yönetimi gibi çevresel sorunların çözümünde kullanılabilir. Nano materyaller, kirleticileri etkili bir şekilde temizlemek için tasarlanabilir.
- Gıda ve Tarım: Tarımda, pestisitlerin ve gübrelerin daha verimli kullanılmasını sağlayarak tarım verimliliğini artırabilir. Ayrıca, gıda güvenliği ve izlenebilirliği için nanoteknolojik sensörler geliştirilmektedir.
- Teknoloji ve inovasyon: Nanoteknoloji, birçok sektörde yeni ürünlerin ve süreçlerin geliştirilmesine olanak tanır, bu da ekonomik büyümeyi teşvik edebilir.

Nanoteknoloji, birçok alanda yenilikçi çözümler sunarak bilime ve topluma önemli katkılarda bulunmaktadır. Ancak, bu alandaki gelişmelerin etik, sağlık ve çevresel etkileri de göz önünde bulundurulmalıdır.

Nanoteknolojide Gelişmeler

- Hedefe Yönelik İlaç Dağıtımı: Nanoteknolojinin sağlık alanındaki uygulamaları daha da yoğunlaşacak, özellikle kanser tedavisinde nanopartiküllerin belirli hücrelere yönlendirilerek daha etkili tedavi yöntemlerini mümkün kılabilir.
- Akıllı Malzemeler: Kendini onaran, çevresel değişikliklere tepki verebilen veya enerji depolama kapasitesini artıran akıllı malzemeler geliştirilebilir. Bu nitelikte malzemeler; inşaat, otomotiv, tekstil gibi birçok sektörde önemli etki sağlayabilir.
- Gelişmiş Enerji Çözümleri: Nanoteknoloji, güneş hücreleri, bataryalar ve süper kapasitörler gibi enerji depolama sistemlerinde daha verimli malzemelerin geliştirilmesini sağlayabilir. Bu husus yenilenebilir enerji kullanımını artırabilir.
- Çevre Dostu Uygulamalar: Nanoteknoloji yöntemleriyle su arıtımı, hava temizliği ve atık yönetimi gibi çevre sorunlarının çözümünde etkili sistemler geliştirilebilir.
- Gıda Güvenliği ve Tarım: Tarımda pestisitlerin ve gübrelerin daha etkili kullanılmasını nanoteknoloji sağlayabilir.

Nanoteknolojinin Gelecekte Yapabileceği İşler

- Kendi Kendine Onarılan Sistemler: Nano malzemeler hasar gördüğünde kendi kendine onarabilen sistemlerin geliştirilmesine olanak sağlayabilir, bu husus malzeme ömrünü uzatabilir.
- Yapay Zekâ ve Entegre Nanoteknoloji: Nanoteknolojinin yapay zekâ ile birleştirilmesiyle birlikte karar verme sistemleri geliştirilebilir. Örnek, tıbbi uygulamalarda gerçek zamanlı veri analizi yaparak anında müdahale sağlanabilir.
- Nano Elektronik ve Hesaplama: Nanoteknoloji, daha güçlü ve enerji verimli bilgisayar üretilmesine olanak tanıyabilir.
- Yeni Sensör Sistemler: Gıda güvenliği, çevre izleme ve sağlık uygulamalarında kullanılacak yeni sensörlerin geliştirilmesine olanak tanıyabilir. Bu sensörler belirli kirleticileri veya sağlık sorunlarını tespit edebilir.
- Kişiselleştirilmiş Tıp: Nanoteknoloji uygulamaları, bireylerin genetik ve biyolojik özelliklerine göre, özelleştirilmiş tedavi yöntemlerinin geliştirilmesine yardımcı olabilir.

Nanoteknolojinin bu gelişmelerle birlikte etik, güvenlik sağlık konularında da etkilerinin değerlendirilmeleri gerekli bir husustur.

Nanoteknolojinin gelecekte insanlığa etkisi ne olabilir?

- Sağlık ve Tıp: Nanoteknoloji, ilaçların hedefe yönelik taşınmasını ve hastalıkların daha hızlı teşhisini sağlayabilir. Örneğin, kanser tedavisinde nanoparçacıklar kullanılarak tümörlere doğrudan ilaç ulaştırabilir, tedavilerin hızlandırılması sağlanabilir.
- Malzeme Bilimi: Nanomalzemeler, yeni ve daha güçlü malzemelerin geliştirilmesine olanak tanır. Bu, inşaat, otomotiv ve elektronik gibi birçok sektörde yenilikçi ürünlerin ortaya çıkmasını sağlayabilir.
- Çevresel Koruma: Nanoteknoloji, çevre kirliliğini azaltmaya yönelik çözümler geliştirebilir. Su ve hava temizleme sistemleri, nanomalzeme kullanımıyla daha etkili hale getirilebilir.
- Eğitim ve Bilim: Nanoteknolojinin gelişimi, bilimsel araştırmaların daha derinlemesine yapılmasını sağlayarak, yeni keşiflerin önünü açabilir.

Bu etkiler, nanoteknolojinin toplum ve ekonomi dinamikleri değiştirebilecek olanağa sahip olduğunu göstermektedir.

VERİ BİLİMİ

Veri bilimi; büyük veri setlerini analiz etmek, anlamak ve bu verilerden bilgi çıkarmak amacıyla matematik, istatistik, bilgisayar bilimi ve alan uzmanlığı gibi disiplinleri birleştiren bir alandır. Veri bilimcileri; verileri toplar, işler, analiz eder ve sonuçlarını görselleştirerek bu bilgileri karar verme süreçlerinde kullanılabilir hale getirir (Sarker, 2021).

Bilime Katkıları

- **Veri Analizi ve Modelleme:** Veri bilimi, karmaşık veri setlerini analiz ederek bilimsel hipotezlerin test edilmesine yardımcı olur. Bu husus, özellikle sosyal bilimler, sağlık bilimleri ve çevre bilimlerinde önemli bir etkiye sahiptir.
- **Öngörü Analizi:** Veri bilimiyle geliştirilen öngörücü modeller, iklim değişikliği, hava durumu tahminleri veya hastalıkların yayılma olasılıkları gibi konularda kullanılabilir.
- **Biyoinformatik:** Genetik verilerin analizi ve biyolojik süreçlerin anlaşılması için, veri bilimi teknikleri kullanılmaktadır. Bu husus, hastalıkların anlaşılmasına ve tedavi yöntemlerinin geliştirilmesine katkı sağlar.
- **Deney Tasarımı:** Bilim deneylerinin tasarımında, veri bilimi, hangi değişkenlerin kontrol edileceği ve hangi verilen toplanacağı noktasında rehberlik eder.
- **Veri Görselleştirme:** Bilimsel verilerin daha anlaşılır bir hale gelmesi için görselleştirme teknikleri kullanılır.
- **Çevresel İzleme:** Veri bilimi, ekosistemlerin izlenmesi ve çevresel değişimlerin değerlendirilmesinde önemli bir rol oynar. Uzaktan algılama ve coğrafi bilgi sistemleri gibi araçlarla çevre verilerinin analizi yapılabilir.
- **Sosyal Bilimler ve Davranış Analizi:** İnsan davranışlarının, sosyal dinamiklerin ve ekonomik eğilimlerin analizi, veri biliminin önemli bir uygulama alanıdır. Politika geliştirme ve sosyal müdahale programlarının oluşturulmasında yardımcı olur.
- **Tıp ve Sağlık:** Damiani ve diğerlerinin belirttiğine göre, tıpta araştırmalar, modern teknolojilerin sağladığı büyük miktardaki hasta verisini değerlendirmek durumundadır. Bu veriler, istatistiksel çalışmaları desteklemek ve nedensel ilişkileri belirlemek için kullanılabilir. Ancak, hastaneler arasında dağınık halde bulunan bu verilerin etkin bir şekilde birleştirilebilmesi ve bu hassas bilgilere dair uygun politikaların uygulanması büyük bir gereklilik olarak öne çıkmaktadır (Damiani, 2015).

Veri bilimi; bilimde, endüstride ve toplum genelinde önemli katkılar sağlar. Verilerin etkili bir şekilde analiz edilmesi, daha bilinçli kararlar alınmasını ve yenilikçi çözümler geliştirilmesini sağlar.

Veri Biliminde Öngörülen Gelişmeler ve Karar Verme Sistemleri

- **Otonom Veri Analizi:** Veri bilimi, daha otomatik hale gelecek, otonom sistemleri insan müdahalesini azaltarak analiz süreçlerini azaltma eğiliminde olacaktır.
- **Gelişmiş Makine Öğrenimi ve Yapay Zekâ Teknikleri:** Makine öğrenimi ve yapay zekâ teknikleri, veri bilimi uygulamalarında daha da yaygınlaşacak. Bu husus, daha karmaşık veri setlerinin daha doğru bir şekilde analiz edilmesine olanak tanıyacak.
- **Gerçek zamanlı Veri Analizi:** Gerçek zamanlı veri işleme ve analiz özellikle nesnelerin interneti (IoT) uygulamalarıyla birlikte büyük bir öneme sahip olacak. Bu husus karar verme süreçlerini hızlandırıcı nitelikte olacaktır.
- **Kişiselleştirilmiş Veri Analizi:** Veri bilimi, bireylerin ve kullanıcıların ihtiyaçlarına göre özelleştirilmiş çözümler sunma kapasitesini artırma niteliğindedir.
- **Veri Güvenliği ve Etik:** Veri güvenliği ve etik konuları, veri biliminin önemli bir parçası haline gelecek. Kullanıcı verilerinin korunması, gizlilik ve şeffaflık sağlamak için daha etkili yöntemler geliştirecektir.
- **Karmaşık sorunları çözme yeteneği:** Veri bilimi, iklim değişikliği, sağlık krizleri gibi karmaşık sorunların çözümünde daha etkili rol alabilir.
- **Sosyal Dinamiklerin Analizi:** Veri bilimi, sosyal medya ve diğer platformlardaki davranışların analizini yaparak toplumsal eğilimleri daha iyi anlamamıza yardımcı olabilir.
- **Veri Biliminin Gelecekte İnsanlığa Etkileri:** Veri bilimi, sağlık alanında hastalıkların yayılma trendlerini anlamamızda, önemli katkı sağlayabilir. Eğitimde daha etkili ve kapsayıcı çözümler sunabilir. Ekonomik

verimliliği artırabilir, bunların yanında toplumsal sorunların çözümünde sosyal problemleri anlamak ve çözüm önerileri geliştirmek için kullanılabilir.

Veri bilimi hızla değişen bir alandır ve birçok sektörde dönüşüme yol açma potansiyeline sahiptir. Ancak, bu gelişmelerde etik, güvenlik ve gizlilik gibi konuların sistem dinamiğinde çok iyi değerlendirilmeleri gerekli bir koşuldur.

YAPAY ZEKÂ, NANOTEKNOLOJİ VE VERİ BİLİMİ ARASINDAKİ ETKİLEŞİMLER

Yapay zekâ, nanoteknoloji ve veri biliminin birlikte değerlendirilmesi, bilimsel araştırma ve endüstriyel uygulamalarda kapsamlı olanaklar sağlar. Bu alanların birleşimi, her birinin güçlü yönlerini bir araya getirerek yenilikçi çözümler geliştirilmesine olanak sağlar. Bu üç disiplinin sinerjisini değerlendirdiğimizde şu imkanlar ortaya çıkar:

Gelişmiş Malzeme Tasarımı

- **Nanoteknoloji**, atom ve molekül seviyesinde yeni malzemeler tasarlamayı sağlar. YZ algoritmaları, bu malzemelerin özelliklerini optimize etmek ve yeni malzemeler geliştirmek için kullanılabilir.
- **Veri Bilimi**, deneysel verileri analiz ederek, nanomalzemelerin performansını daha iyi tahmin etmek ve süreçleri hızlandırmak için gerekli modelleri oluşturur.
- Çok daha hafif, dayanıklı veya enerji verimli malzemeler üretilebilir.
- Grafen, karbon atomlarının tek bir tabakasından oluşan bir nanomalzemedir ve son derece güçlü, hafif ve iletken özelliklere sahiptir. Ancak, grafeni süper kapasitörler veya giyilebilir elektronikler gibi endüstrilere uygulamanın en iyi yolunu bulmak büyük veri analizi gerektirir.

Tıp ve Biyoteknoloji Uygulamaları

- **Nanoteknoloji**, ilaç taşıyıcı sistemler ve biyosensörler gibi tıbbi uygulamalarda büyük bir rol oynar. YZ, bu sistemlerin optimize edilmesi ve kişiselleştirilmiş tedavi seçenekleri sunulması için kullanılabilir.
- **Veri Bilimi** ile hastalıklara dair büyük veri kümeleri analiz edilerek, daha iyi tanı ve tedavi yöntemleri geliştirilebilir. Nanoteknolojik araçların, özellikle kanser tedavisi gibi alanlarda, hastanın genetik yapısına göre optimize edilmesi YZ tarafından kolaylaştırılabilir.

Enerji Verimliliği ve Çevresel Uygulamalar

- Yapay zekâ, sensörlerden gelen verileri kullanarak akıllı şehirlerin kaynaklarını enerji ve trafik yönetimlerinin daha verimli olmasını sağlar.
- Nanoteknolojik yenilikler, daha verimli güneş pilleri, bataryalar ve diğer enerji çözümlerini mümkün kılar. YZ, enerji üretimi ve depolama süreçlerinin optimizasyonunda kullanılabilir.
- **Veri Bilimi** ile enerji tüketim verilerinin analiz edilmesi, enerji yönetiminde önemli kazançlar sağlar. Ayrıca çevresel izleme ve sürdürülebilir enerji kaynaklarının optimizasyonu sağlanabilir.
- Yapay zekâ ile birleştiğinde nano sensörler çevreleri sürekli izleyebilir ve hava kirliliği, su kalitesi ve diğer çevresel faktörlerdeki eğilimleri tahmin edebilir. Yapay zekâ algoritmaları, şehir genelindeki binlerce nano sensörden gelen verileri analiz eder. Bu sensörler, kirlilik seviyelerini, radyasyonu ve bakterileri tespit edebilir ve yapay zekâ sistemi potansiyel tehlikeleri tahmin edebilir.

Üretim ve Otomasyon

- **Nanoteknoloji**, yüksek hassasiyetle mikro ve nano ölçekli cihazlar üretmeyi sağlar. Bu süreçte, YZ üretim süreçlerini otomatikleştirip iyileştirirken, veri bilimi de süreçteki hataları ve aksaklıkları analiz eder, kaliteyi artırır.
- Akıllı üretim süreçlerinde, nanoteknolojinin sunduğu ayrıntılı bilgiler, YZ ve veri bilimi ile birleşerek daha hızlı, daha esnek ve düşük maliyetli üretim çözümleri ortaya koyar.

Sağlık ve Yaşam Bilimleri

- Nanosensörler ve cihazlar, biyoloji süreçlerini hücre düzeyinde izleyebilir ve çok daha doğru tanımlar koyabilir. YZ bu verileri analiz ederek tedavi planlamalarında büyük bir hız ve doğruluk sağlar.
- **Veri Bilimi** sayesinde biyomedikal veriler analiz edilerek, klinik karar destek sistemleri oluşturulur ve bireysel sağlık verileri nanoteknolojik cihazlarla toplanıp değerlendirilebilir.

Yapay Zekâ ile Nanorobotlar

- **Nanoteknoloji** ile geliştirilen nanorobotlar; tıp, tarım, çevre bilimleri gibi çeşitli alanlarda büyük potansiyele sahiptir. Bu nanorobotların karmaşık görevleri yerine getirmesi için YZ algoritmaları ile yönlendirilmesi mümkündür.
- **Veri Bilimi**, nanorobotların çalışma kapasitelerini analiz ederek, gerçek zamanlı optimizasyon ve görev dağılımı sağlar.
- Çevresel verileri toplamak, uzay hava durumunu izlemek ve telekomünikasyon konusunda yardımcı olabilmek için nano boyutta sensörleri yörüngede taşıyan küçük, maliyeti fazla olmayan uydulardır.

Bilimsel Araştırmaların Hızlanması

- Nanoteknoloji ve veri bilimi, bilimsel deneylerin daha hızlı, daha hassas ve düşük maliyetli olmasını sağlar. YZ algoritmaları ise deneylerden elde edilen verilerin analizini otomatik hale getirir, bu da bilimsel keşifleri hızlandırır.
- Özellikle karmaşık biyolojik süreçler ya da malzeme bilimlerinde kullanılan simülasyonlar, YZ ve veri bilimi sayesinde hızla analiz edilip optimize edilebilir.
- Yapay zekâ, nanoteknoloji ve veri biliminin etkileşimi, gelişmiş malzemeler ve yapay zekâ destekli nanosensörler, uzay görevlerinin daha verimli ve etkin olmasına yardımcı olur.

Güvenlik ve Savunma Teknolojileri

- **Nanoteknoloji**, güvenlik ve savunma alanında yeni nesil malzemeler ve sensörler geliştirmeye olanak tanır. Bu sistemler, YZ ile entegre edilerek akıllı karar verme mekanizmalarına sahip olabilir.
- **Veri Bilimi** ile büyük veri kümeleri analiz edilerek, tehdit algılama, izleme ve savunma sistemleri optimize edilir.

Büyük Veri ve Modelleme

- **Veri Bilimi**, nanoteknolojik ve biyolojik sistemlerin karmaşık veri setlerini analiz etmek için kullanılır. YZ, bu verilerden öğrenerek modeller geliştirir ve simülasyonlar oluşturur.
- Nanoteknolojik cihazların üretim süreçlerinde elde edilen veriler, YZ ve veri bilimi ile sürekli optimize edilir, böylece maliyetler düşer ve süreçlerin verimliliği artar.

Yapay Zekâ, Nanoteknoloji ve Veri Biliminin Bilim Metodolojisine Etkileri

Bu teknolojiler yalnız belirli endüstrileri uygulamaları dönüştürmekle kalmıyor aynı zamanda bilim metodolojisini önemli düzeyde etkiliyor; yeni düşünme, deney tasarlama, veri yorumlama olanakları getiriyor. Geleneksel bilimsel araştırma, hipotez oluşturma ve ardından veri toplamaya dayanıyordu. Ancak **yapay zekâ, nanoteknoloji ve Veri Bilimi** ile bu süreç daha dinamik hale geliyor ve bilim insanlarının veri toplamayı otomatikleştirip optimize etmelerine olanak tanıyor. Deney tasarımında yapay zekâ deneylerin tasarlanma şeklini değiştiriyor ve hipotez üretimini otomatikleştiriyor. İnsanların hipotez belirlemesi yerine yapay zekâ algoritmaları büyük veri kümelerini inceleyerek gözlemlenen kalıplara dayalı birden fazla hipotez öneriyor. Nanosensörler ve cihazlar gerçek zamanlı veri toplayarak biyolojik sistemler, kimyasal reaksiyonlar veya çevresel süreçler hakkında daha fazla ayrıntı sağlıyor.

Bazı durumlarda yapay zekâ sistemleri toplanan verilere dayalı olarak deneyleri gerçek zamanlı ayarlayabilirler. Klinik deneylerden ve moleküler simülasyonlardan elde edilen verilerden yararlanarak yeni ilaçların keşfi mümkün olmaktadır.

Bilim, genellikle “teori odaklı” olmuş deneyler de bu teorileri doğrulamak veya yanlışlamak üzere tasarlanmıştır. Ancak nanoteknoloji tarafından üretilen büyük veriler ve yapay zekâ tarafından analiz edilen bu verilerle, bilim daha fazla “veri odaklı” olmaktadır.

Sonuç olarak; **yapay zekâ**, **nanoteknoloji** ve **Veri Bilimi** birbirlerini destekleyen disiplinlerdir. Bu alanların kesişiminde, tıptan enerjiye, çevre bilincinden sanayiye kadar geniş bir yelpazede, bilimler arası metodoloji kapsamında, yenilikçi çözümler ve büyük atılımlar mümkündür (Güvenen, 2013, 2016). Bu teknolojiler, bilimsel keşifleri hızlandırırken, günlük yaşamda da daha etkin ve verimli uygulamalar sunar.

Yapay zekâ ve nanoteknolojinin kimya ve fizik bilimlerinde devrim niteliğinde etkisi olmuştur. Yapay zekâ ve veri bilimi, bilim insanlarının malzeme yapılarının milyonlarca potansiyel kombinasyonunu simüle etmelerine ve benzersiz özelliklere (Örneğin: süper iletkenlik, ısı dayanaklılığı) sahip olanları aramalarına olanak tanır. Veri bilimi araçları, araştırmacıların laboratuvarında sentezlenmeden önce malzemelerin özelliklerini tahmin etmelerine olanak tanır; bu da malzeme araştırmalarını köklü bir şekilde değiştirir, laboratuvar testlerine ihtiyaç duymadan keşifleri hızlandırır.

Bilim metodoloji, teorik modeller ve ampirik deneyler arasındaki döngüsel bir süreçti. Yapay zekâ, nanoteknoloji ve veri bilimi sayesinde simülasyon ve deney yakın bir entegrasyona kavuşuyor.

Nanoteknoloji, atom düzeyinde süreçlerin ayrıntılı simülasyonlarını mümkün kılar ve bu simülasyonlardan elde edilen ampirik veriler, yapay zekânın simülasyonlarını geliştirmesine yardımcı olur.

Nanosensörlerden ve laboratuvar verilerinden elde edilen yapay zekâ modelleri, gerçek dünyadaki davranışları yansıtan simülasyonlar olabilir, bu da simülasyon ile deney arasında daha sıkı bir geri bildirim oluşturur.

Yapay zekâ destekli otonom sistemlerinin devreye girmesiyle birlikte bilim kendi kendini öğrenen modellerin ve otonom makinelerin, insan müdahalesi olmadan bilimsel keşifler yapabileceği bir paradigmaya doğru evriliyor.

Yapay zekâ sistemleri kendi deneylerini tasarlayıp uygulayabilir ve minimum insan gözetimiyle çalışabilir. Nanosensörlerle donatılmış yapay zekâ destekli robotlar derin deniz hendekleri veya diğer gezegenler gibi farklı çevrelerde keşif yaparak veri toplayabilir ve araştırma stratejilerine gerçek zamanlı olarak optimize edebilir. NASA'nın marsta kullandığı yapay zekâ destekli geziciler (Rovers) otonom keşif yönünde örnek teşkil ediyor, burada makineler, gerçek zamanlı verilere dayalı kararlar alıyor.

Yapay zekâ, nanoteknoloji ve Veri Biliminin entegrasyonu bilim metodolojisinde köklü bir değişime yol açıyor: bilim, hipotez odaklı yaklaşımlar yerine daha fazla veri odaklı çalışmalara dayanıyor bu da yeni olguların keşfedilmesini sağlıyor.

“Bilim, matematik, istatistik, veri bilimi uygulamaları, yapay zeka, bilgi sistemleri, karar sistemleri Türkiye optimali ve insanlık optimali arayışında, insanlık refahına, barışına katkı sağlayabildikleri süre, en önemli katma değerlerini sağlamış olurlar. Bilim üretmek, istatistik üretmek, yorumlamak; karanlığa ışık götürmek kadar önemli bir görevdir.” (Güvenen, 2011)

2024 Fizik ve Kimya Nobel Ödülleri yapay zekâ kapsamında yapılan araştırmalara verildi, konunun önemini belirten bir gösterge olarak değerlendirilebilir.

Nanoteknoloji, yapay zekânın daha etkin çalışması ve kullanımı için, çok hızlı bilgisayar ve bellek yapımında daha etkin olacaktır.

Yapay zekâ ve nanoteknoloji ile geliştirilecek yeni nesil bataryalar mobil cihazlarda YZ'nin daha etkin kullanımını sağlayabilecek.

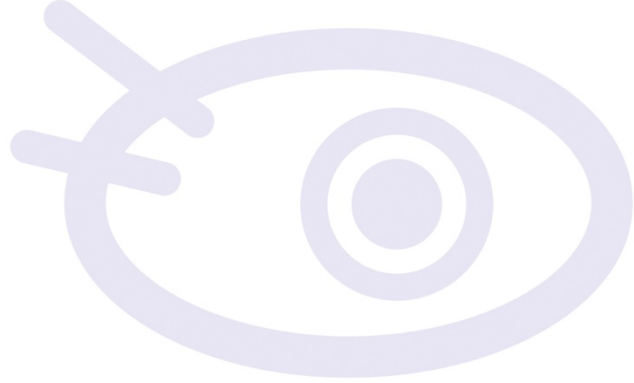
Yapay zekânın veri bilimiyle gelecekte rasyonel düşünme modelleri, sosyoekonomik siyaset ve finans analizleri çok etkili olacağı beklenmektedir. “*İnsanların geleceği için en ciddi sorun teknolojinin gelişme hızı ile insanların uyum sağlama hızı arasındaki fark olacaktır.*”

KAYNAKÇA

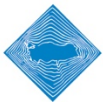
- Bainbridge, W. S. (2015). Ethical Considerations in the Advance of Nanotechnology, *Nanotechnology: Science, Innovation, and Opportunity*. p. 238, Prentice Hall.
- Damiani, A., Vallati, M., Gatta, R., Dinapoli, N., Jochems, A., Deist, T., Van Soest, J., Dekker, A., & Valentini, V. (2015). Distributed learning to protect privacy in multi-centric clinical studies. In *Lecture notes in computer science* (pp. 65–75). https://doi.org/10.1007/978-3-31919551-3_8
- Güvenen O. (2009). “*Impacts of Nanoscience and Nanotechnology on SocioEconomic Phenomena*”. *Bilkent University Graduate Lectures, 20i9*.

- Güvenen O. (2011). *İstatistik Bilimi, Etik, Dünya Dinamikleri, Bilgi Tahrifatı ve Karar Sistemlerine Etkileri*. *İstatistik Araştırma Dergisi*, Cilt 8 Sayı 2, s. 1-12, Türk İstatistik Kurumu, Ankara, ISSN 1303-6319
- Güvenen O. (2013). "Bilim, Bilimlerarası Metodoloji, Kompleks Sistemler, Nanoteknoloji ve Küresel Dinamikler: 2013 Akademik Yılı Açılış Dersi". *İstanbul Gedik University*
- Güvenen O. (2015). "Nanoteknolojilerin Toplum ve Ekonomi3 Yapılarına Etkileri". *Mülkiyeliler Birliği Yayını 1 (2015): 187-205*.
- Güvenen O. (2016). "Transdisciplinary Science Methodology as a Necessary Condition in Research and Education". *Transdisciplinary Journal of Engineering & Science* 7, s.1, <https://doi.org/0.22545/2016/0080>.
- Himeur, Y., Elnour, M., Fadli, F., Meskin, N., Petri, I., Rezgui, Y., Bensaali, F., & Amira, A. (2022). AI-big data analytics for building automation and management systems: a survey, actual challenges and future perspectives. *Artificial Intelligence Review*, 56(6), 4929–5i21. <https://doi.org/10.1007/s10462-022-10286-2>
- Sarker, I. H. (2021). Data Science and Analytics: An Overview from DataDriven Smart Computing, Decision-Making and Applications Perspective. *SN Computer Science*, 2(5). <https://doi.org/10.1007/s42979-021-00765-8>
- Wang, H. et al. (2023). Scientific discovery in the age of artificial intelligence. *Nature*, 620(7972), 47–60. <https://doi.org/10.1038/s41586-023-06221-2>

IDS
SS
2024CC



Proceedings



TURKISH STATISTICAL ASSOCIATION

CONTENTS

Yerel Yönetimlerde Uygulanmış Bir Araştırma Modeli: Ankara İli Örneği	4
İlaç Sektöründe Gelişmiş Talep Tahmini: Entegre Makine Öğrenmesi ve Derin Öğrenme Yaklaşımları	16
Comparative Analysis of LLM-Based and Conventional NLP-Based Methods for MultiClass Text Classification in Turkish.....	22
An Overall Equipment Efficiency Predictive Analysis of a Hydraulic Press System by Time Series Forecasting with Topological Features	27
Sigortacılıkta Araç Sigortası Talep Sınıflandırmasında Gelişmiş Makine ve Derin Öğrenme Yöntemleri: Yenilikçi Yaklaşımlar ve Performans Değerlendirmeleri.....	34
IMDb Score Estimation using Movie Dialogues: A Text Mining and Machine Learning Hybrid Approach	44
Artificial Neural Networks in Marketing: A Comprehensive Review of Consumer Behavior Applications	50
Comparative Analysis of Monthly Average Rainfall in Amasya Province with Exponential Smoothing Methods	71
SARIMA-YSA Hibrit Model Yaklaşımı: Türkiye’de Turizm Üzerine Bir Uygulama	78
Analyzing and Improving the Impact of Software Changes on IT Operations: A Process Mining and Forecasting Approach.....	88
Zaman Serisi Tahmin Kombinasyonları İçin Çok Kriterli Karar Verme Yöntemine Dayalı Yeni Bir Yaklaşım	95
Nesne Tanıma ve Sınıflandırmada Kullanılan Derin Öğrenme Yöntemleri	104
Classification of Brand Perception Using Random Forest: Brand Preference, Brand Loyalty and Brand Trust.....	113
Yapay Sinir Ağları ile İstanbul Elektrik Talep Tahmini	120
Integrating RFM Analysis with Power BI for Enhanced Customer Segmentation and CLV Insights	127
Avrupa İnsan Hakları Mahkemesi Yargı Kararlarının Metin Madenciliği Yöntemleri İle Tahmin Edilmesi.....	135
Birliktelik Kuralı Analizi ile Ürün Kod Atama Uygulaması	143
MEDOS: Kısa Dönemli İstatistiklerin Mevsimsellikten Arındırılması için Bir R-Shiny Uygulaması	151
Yapay Zeka Destekli Veri Sınıflandırma ve Yönetim Modeli.....	160
Covid-19 ve Deprem Süreçlerinde Uzaktan Eğitim: Öğrenci Beklentileri ve Deneyimlerinin Makine Öğrenimi ile Karşılaştırılması.....	167
Enhancing Time Series Decomposition with Artificial Intelligence: Improving Pattern Identification and Interpretation in Economic Data.....	176

Borsa İstanbul Bist100 Endeksi için Kurulan Modellerin Karşılaştırılması	183
Yapay Zekanın Kurumsal Dönüşümdeki Rolü	190
Airline Business Model Evolution: A K-Means Clustering Analysis.....	197
TCMB Açık Veri Platformu: EVDS Geçmişi ve Geleceği.....	205

Yerel Yönetimlerde Uygulanmış Bir Araştırma Modeli: Ankara İli Örneği

Dr. Filiz Aydın^{1*}

¹Siyaset Bilimi ve Kamu Yönetimi ABD, Kamu Yönetimi Bölümü / Hacettepe Üniversitesi, Ankara, Türkiye

Özet

Bu bildiriye sosyal bilimlerin alt bilim dalı olan yerel yönetimler alanında nitel, nicel ve karma araştırma yöntemlerinin uygulaması ve bir araştırma modelinin nasıl kurulması gerektiği “Yeni Büyükşehir Yönetim Modelinde İlçe Belediyelerinin Hizmet Sunumlarına İlişkin Vatandaş Memnuniyet Düzeyini Ölçmeye Yönelik Bir Araştırma: Ankara Örneği” adlı kabul edilmiş ve yayınlanmış doktora tezi örneği üzerinden ele alınacaktır. Ayrıca doktora tezi kapsamında incelenen ve araştırma modeli kurulan benzer çalışmaların (makale ve lisansüstü tezlerin) araştırma modelleri de incelenerek diğer araştırmacıların yöntem ve model yaklaşımları değerlendirilecektir.

Nihai olarak yerel yönetimler ana bilim dalında ve daha genel olarak sosyal bilimlerde amaca uygun, gerçekçi ve doğru sonuçlar elde edebilmek için nitel ve nicel araştırma modellerinin nasıl kurulması gerektiği, örneklem dağılımları hazırlanırken dikkat edilmesi gereken hususlar ve yöntem seçimlerine ilişkin genel tespit ve önerilerde bulunulacaktır. Tüm bu tespitler büyük ölçüde söz konusu doktora tezinde kullanılan yöntem üzerinden ele alınarak, benzer çalışmalarda kullanılan yöntemler de değerlendirilecektir.

Bildiriyi konu alan araştırmasında nicel araştırma ve nitel araştırma olmak karma araştırma yöntemi kullanılmıştır. Örnekleme seçilen ilçeler 2012 yılından önce kırsal alanı bulunanlar arasından; ‘ilçe nüfusu, kırsal alan nüfusu, ilçelerin alanı (km²), ilçelerde km² ye düşen nüfus, Ankara şehir merkezine mesafe (km), 2022 sosyo- ekonomik gelişmişlik sıralaması skoruna (SEGE)’ göre 3 gruba ayrılmış ve bu 3 grup içinden rastgele 9 ilçe seçilmiştir. Nicel araştırma kapsamında 380 kişi ile yüz yüze anket, nitel araştırma kısmında ise 19 mülakat yüz yüze ve telefon yolu ile görüşülerek karma yöntem kullanılarak yapılmıştır.

Genel olarak benzer çalışmalarda kullanılan araştırma yöntemleri değerlendirildiğinde örneklem büyüklüklerine ve görüşülen kişi özelliklerine bakıldığında en uygun yöntemin nitel araştırma mı nicel araştırma mı olduğu konusunda belirsizlik yaşandığı söylenebilir. Farklı özelliğe sahip katılımcıların gruplandırılması yoluna gidilmediği görülmüş olup analizler de farklı özelliğe sahip oldukları halde toplu olarak yapılmıştır. Benzer çalışmalarda karma yöntemin hiç kullanılmadığı görülmüş olup bu yöntemin daha yaygınlaştırılması gerektiği önerilmektedir. Doğru sonuçlar elde edebilmek için araştırma modelinin çok önemli olduğu bir kere daha vurgulanmak istenmektedir.

Anahtar Kelimeler: Yerel Yönetimler, Nitel, Nicel ve Karma Araştırma, Örneklem Dağılımı, İstatistik, Ankara.

GİRİŞ

Sosyal bilimler, toplumsal olayların ve olguların nasıl ve neden ortaya çıktığını ortaya çıkarmak için insan davranışlarının ne biçimde değiştiğini ortaya koymaya çalışırken teknolojik değişimle birlikte dönüşüm sürecine girmiştir. Bu değişim ve dönüşüm sürecinde özellikle toplumun yönetiminde kullanmak üzere daha fazla kesin bilgiye ve veriye ihtiyaç duyulmaya başlanmıştır.

Sosyo-ekonomik, yönetsel, toplumsal ve davranışsal değişime bağlı olarak da tutumların, davranışların, inançların vd. değişimini daha net görebilmek ve analiz edebilmek için sayısal verilere olan ihtiyaç nicel araştırma geleneğine olan talebi daha da arttırmıştır. Buradaki en genel amaç ise kısıtlı olan doğal ve beşeri kaynakların verimli, etkin ve etkili kullanılması ile tüm süreçlerin bilimsel yöntemlere, planlamaya dayalı olarak yürütülmesi gerekliliğidir.

Sosyal bilimlerde özellikle lisansüstü eğitimde bilimsel araştırma yöntemlerinin kullanarak kanıta dayalı eser yazmak birçok akademisyen ve üniversite tarafından öğrencilere tavsiye edilmektedir. Geçmişte nicel araştırma yöntemleri daha fazla tercih edilirken artık nitel araştırma yöntemleri de son yıllarda daha çok tercih edilir durumdadır (Karataş, 2015: 1). Nicel ve nitel araştırma yöntemlerinin bir arada kullanıldığı “karma yöntem” de araştırmacılara elde ettikleri sonuçlar açısından zenginlik katarak elde edilen sonuçların farklı araştırma yöntemleriyle doğrulanma fırsatı da vermektedir. Karma araştırma yöntemi; tek bir araştırmada ya da araştırmalar içerisinde nicel ve nitel araştırma yöntemleriyle verilerin toplanarak, istatistiksel açıdan analiz edilmesini ve sonuçların da yorumlanmasını içermektedir (Leech ve Onwuegbuzie, 2009, s.266).

Tek bir yöntemle bütünün yakalanamayacağı görüşü sosyal bilimlerde de yaygınlaşmaya başlanmış olup Yıldırım ve Şimşek’ göre de “Herhangi bir olguya ilişkin bütüncül anlayış ancak çoklu bakış açıları yoluyla elde edilebilir.” görüşü dile getirilmektedir (Yıldırım ve Şimşek, 2008: 28). Karma yöntemin en önemli özelliği sayısal verilerle, sözel verileri birlikte kullanılabilmesidir ki bu yöntemle araştırma konusu daha geniş açıyla ele alınırken derinlemesine de irdelenmesini sağlamaktadır.

Nicel arařtırmalarda daha byk rneklemlerle sayısal veriler toplanırken genel tahminler elde etmek amalanmaktadır. Nitel arařtırmalarda ise daha az sayıda denekle derinlemesine detaylı grřmeler yapılarak katılımcıların deneyimleri ve dřnceleri daha iyi anlařılmaya alıřılmaktadır. Karma yntemde rneğın daha geniř bir gruba anket yapılırken, daha kk gruba da derinlemesine grřmeler yapılabilmektedir.

Karma ynteminin arařtırmacı iin ok nemli faydaları vardır. Bunların en bařında arařtırma konusu hakkında kapsamlı bilgi saėlanırken konuyu farklı aılardan ele alınmasını saėlamaktadır ve en nemlisi de elde edilen sonuların gvenilirliėi arttırmaktadır. Ancak karma yntemde nicel ve nitel verileri birleřtirmek ok kolay olmadığı gibi hatta bazı arařtırmacılara karmařık da gelebilmektedir. Farklı yntemler kullanmak sonular aısından da zaman zaman eliřki yaratıyor gibi grnse de bu durum zenginlik ve tamamlayıcılık olarak deėerlendirilmez (Tunalı, Gz ve zen:2016).

Nitel arařtırmada en ok kullanılan yntemler olan grřme, odak toplantı, gzlem ve dokman analizi ile algıların, durumların ve olayların doėal ortamlarda btnel olarak ortaya konulması hedeflenmektedir. Bu yntemde yorumlayıcı yaklařım esas olup ve olaylar yklenen anlamlar aından ele alınmaktadır (Altunıřık ve diėerleri, 2010: 302). Nitel arařtırmanın en belirgin zelliėi derinlemesine bilge elde etmenin yolunu sunması aısından arařtırmacıyı keřfe ynelerek znel bakıř aısı imkanı sunmasıdır (zdemir, 2010: 326). Nitel arařtırmalarda neden sonu iliřkisi kullanılmadıėı gibi istatistiklere de minimal yer vermektedir. evre, sre ve algılara iliřkin veri toplama imkanı sunan nitel arařtırmalarda tek bir grř bile arařtırmacı aısından deėerli bulunup nemli sonu olarak kabul edilebilmektedir. Nitel arařtırmalar arařtırma tasarımı aısından nicel arařtırmaya gre daha fazla esneklik saėlarken keřfedici zelliėi ile derinlemesine analize deėer konular iin olduka faydalı bir yntemdir (Neuman, 2012: 228). Arařtırmacı denekle empati kurarak farklı bakıř aılarını, rntleri doėal ortamda ortaya ıkarabilir ve arařtırmacıya nicel arařtırmadan farklı olarak “ka tane”, “hangi sıklıkta” sorularından farklı olarak “neden” sorusunu sorma imkanı tanır. Nitel arařtırma yntemi yorumlayıcıyken, nicel arařtırma tanımlayıcı zelliėe sahiptir. Nitel arařtırmada rnekleme sayısını arařtırmacı belirlemektedir. Nicel arařtırmaya gre ok daha az katılımcı ile arařtırma yapılırken sonuları genellenmesi kaygısı bulunmamaktadır.

Nicel arařtırma ise mevcut durumu tespit etmek ve olgulara dair bilgi sahibi olmak amacıyla sayısal deėerlerin sistematik, yansız bir řekilde lldė ve bu yapılan lmlerin tekrar edilebildiėi srecin tamamını kapsamaktadır (Burns ve Grove, 1993). Nicel arařtırmalarda evrenden olasılıklı rnekleme yapılarak elde edilen sonular ve tahminler istatistiklerle aıklanmaktadır. Nicel arařtırmada hipotezler kesin olarak belirlenirken arařtırma sonuları da kesin olarak aıklanmaktadır (Fraenkel ve Wallen, 2006). Evrenden alınan rneklemin bir alt kme olarak evreni temsil etmesi gerekmektedir (ıngı, 1994).

Tezin konusuna kısaca deėinilecek olursa; artan kentleřme, eřitli sorunlarla birlikte, kent alanlarının belirlenmesi ve buna gre ynetim yapısının oluřturulmasına dair lek byklė arayıřlarını beraberinde getirmiřtir. Trkiye’de de 1982 Anayasasında yapılan deėiřiklikle 1984 yılında 3030 sayılı, 2004 tarihli 5216 sayılı, 2008 yılında 5747 sayılı ve 2012 tarihinde 6360 sayılı Bykřehir Belediye Kanunları ıkarılarak belediye sınırlarında deėiřikliėe gidilmiřtir. 6360 sayılı Kanunla yapılan son dzenlemede ise 14 bykřehir 30’a ıkarılarak, 30 bykřehirde bulunan il zel idareleri kapatılmıř ve ky ile belde belediyeleri de mahalle ynetim birimine dnřtrlmřtir. Bykřehirlerdeki ile belediyeleri ynetim alanı ise ile mlki sınırlarına kadar geniřletilmiřtir. Bu kanunlarla zaman iinde yerelde merkezileřme artmıřtır. Merkezileřmenin artması, hizmete eriřimde glk, maliyetlerin ykselmesi ve demokratik katılımın azalması gibi sorunları da beraberinde getirmiřtir.

6360 sayılı Kanun zellikle kırsal alanın ynetiminde, bykřehir ve ile belediyelerin ynetimi aısından kapsamlı ve nemli etkiler yaratmıřtır. Tzel kiřilik aısından kırsal alanın siyasal ve ynetsel temsil gc zayıflamıř, hizmet sunumu ise kırsal alan aleyhine deėiřiklik gstermiřtir. Tez de bu tr eleřtirilerin ne lde geerli olup olmadığı ortaya koymak iin hipotezler geliřtirilmiřtir. Tez kapsamında iki hipotez test edilmiřtir ve arařtırma sonucunda iki hipotez de kabul edilmiřtir.

H₁: 6360 sayılı Kanunla deėiřen bykřehir ynetim modelinde ile belediyelerinin hizmet sunumu ile vatandař memnuniyeti arasında bir iliřki vardır.

H₂: 6360 sayılı Kanunla deėiřen bykřehir ynetim modelinde yerel ynetimlerin demokratiklik (yerel katılım) dzeyi ile ile belediyelerinin hizmet sunumu hakkındaki vatandař memnuniyeti arasında bir iliřki vardır.

Bu kapsamda, 6360 sayılı Kanun ile yrrlėe giren yeni bykřehir belediye ynetim modeli, eski model ile karřılařtırmalı bir řekilde ele alınmıř, ile belediyelerinin hizmet sunumları hakkındaki vatandař memnuniyet dzeyleri llmřtir. Ayrıca bu deėiřimin demokratik katılıma etkileriyle birlikte vatandař memnuniyet dzeyi de deėerlendirilmiřtir (Aydın, 2024)

1. DOKTORA TEZİ KAPSAMINDA KULLANILAN ARAŞTIRMA YÖNTEMİ

Ankara ilinin kırsal alanını kapsayan söz konusu doktora tezinin alan araştırmasında nicel (yerel halk) ve nitel (muhtarlar) olmak üzere (karma yöntem) iki farklı araştırma yöntemi kullanılmıştır. Önce il belirlenmiş, sonra ilçeler seçilmiş, daha sonra da ilçelerden mahalleler seçilerek, nicel araştırma kapsamında örneklem büyüklüğü oransal olarak aşama aşama olasılıklı olarak dağıtılmıştır.

1.1 Örneklem Büyüklüğünün Belirlenmesi

Araştırma sorularının belirlenmesini takiben alan araştırması yönteminin ilk adımı örneklem büyüklüğünün belirlenmesi olmuştur. Bunun için öncelikle kapsam içine alınacak ilçeler belirlenmiştir. Bu ilçeler tezin konusu ile ilgili olarak 6360 sayılı Kanunun yasalaştığı yıl olan 2012 itibariyle kırsal alana sahip ilçeler arasından seçilmiştir. Örnekleme alınan ilçelerin “kırsal alan nüfusu”, 6360 sayılı Kanun yasalaşmadan önce yapılmış olan 2009 yılı Mahalli İdareler Genel Seçimindeki “köy ve belde seçmen sayılarının”, genel seçmen büyüklüğü içindeki oranları dikkate alınarak belirlenmiştir. Bu oran Ankara ili için %3,7 olarak hesaplanmıştır. Kırsal alan nüfusu olmayan 8 ilçe (Altındağ, Çankaya, Etimesgut, Keçiören, Mamak, Pirsaklar, Sincan ve Yenimahalle) kapsam dışı bırakılmıştır (Tablo 2).

Kalan 17 ilçenin genel özellikleri incelendiğinde üç grupta yoğunlaştığı tespit edilmiştir ki genel özelliklere bir sonraki alt bölümde değinilecektir. Örneklem büyüklüğü 3 ayrı grupta beklenen memnuniyet oranlarının %40, %50 ve %60 olması halindeki 2 serbestlik dereceli ki-kare testi için W (effect size) =0,163'lük bir etki büyüklüğünü %80 güç ve %95 güven düzeyinde incelemek için gereken örneklem sayısı en az 363 olarak hesaplanmıştır. Alanda toplam 380 anket yapılmıştır.

Nitel araştırma kapsamında ise seçilen 9 ilçede ise muhtarlarla yarı yapılandırılmış görüşme formu eşliğinde her ilçede 2 olmak üzere toplam 18 mülakat yapılması planlanmış olsa da toplamda 19 mülakat yapılmıştır (Aydın, 2024:214).

1.2 İlçelerin Seçimi ve Gruplandırılması

2012 yılı öncesinde kırsal alana sahip olan ilçeler; “ilçe nüfusu”, “kırsal alan nüfusu”, “ilçelerin alanı (km²)”, “ilçelerde km² ye düşen nüfus”, “Ankara şehir merkezine uzaklık (km)”, “2022 yılı sosyo-ekonomik gelişmişlik sıralaması¹ skoruna” göre 3 gruba ayrılmış ve bu 3 gruptan benzer ilçeler içinden rastgele 9 tanesi seçilmiştir.

Tablo 1. Örnekleme Seçilen İlçelerin Değişken Verileri- Hedeflene ve Gerçekleşen Sayılar

SN	Seçilen İlçeler	2020 Nüfusu	İl Nüfusu içindeki Payı (%) [*]	2009 Mahalli İdareler Genel Seçimleri	İlçelerin		Şehir Merkezine Mesafesi (Km)	2022 Yılı Sosyo-Ekonomik Gelişmişlik (SEGE)		2020 Kırsal Alan Nüfusu Tahmini	Örneklem Oranı (0,0036) (Planlanan Sayı)	Gerçekleşen Sayılar
					Alanı (Km ²)	Km ² 'ye Düşen Nüfus		Sıralama Skoru	Kademesi			
1	Akyurt	37.456	0,66	3,0	369	102	29	0,701	2	1135	4	10
2	Kazan	12.941	0,23	6,2	547	24	34	1,367	2	796	3	12
3	Polatlı	126.623	2,24	22,0	3.618	35	72	0,86	2	27.811	99	88
4	Ayaş	13.686	0,24	30,9	1.041	13	46	0,525	3	4.230	15	37
5	Çubuk	91.142	1,61	8,4	1.198	76	37	0,373	3	7.678	28	46
6	Nallıhan	27.434	0,48	60,9	2.079	13	132	0,243	3	16.716	60	52
7	Kalecik	56.736	1,00	39,4	1.110	51	51	0,083	4	22.371	81	62
8	Çamlıdere	8.883	0,16	56,7	782	11	70	-0,232	5	5.038	18	30
9	Bala	25.780	0,46	59,0	1851	14	48	-0,351	5	15200	55	43
Toplam		400.681								100.975	363	380

¹ T.C. Sanayi ve Teknoloji Bakanlığı Kalkınma Ajansları Genel Müdürlüğü (2022).

² <https://www.ysk.gov.tr/doc/dosyalar/docs/2009MahalliIdareler/SecmenSayilari/Belde.htm>

<https://www.ysk.gov.tr/doc/dosyalar/docs/2009MahalliIdareler/SecmenSayilari/Ilce.htm>

<https://www.ysk.gov.tr/doc/dosyalar/docs/2009MahalliIdareler/SecmenSayilari/Koy.htm>

Erişim Tarihi: (01.10.2022).

Yukarıdaki tabloda örnekleme seçilmiş ilçelerin hangi değişkenler dikkate alınarak seçildiğine yer verilmiştir. Seçimde ilçelerin 2020 yılı nüfusu, 2009 Mahalli İdareler Genel Seçimleri seçmen sayısı³ göre köy ve belde payı, alanı, il merkezine uzaklığı, ilçelerin alanı ve Km²'ye düşen nüfus, 2022 SEGE ilçe gelişmişlik skoru, 2020 kırsal alan nüfusu dikkate alınmıştır. Ayrıca ilçe örneklem büyüklükleri yer almaktadır. Görülmektedir ki ilçeler 3 grupta toplanmaktadır.

Tablo 2. Ankara İlinin 2009 Mahalli İdareler Genel Seçimleri Sonuçlarına Göre Seçmen Verileri

SN	İlçeler	2009 Yılı Mahalli İdareler Genel Seçimleri Sonuçlarına Göre;										Gruplar
		Gelişmişlik Kademesi	Toplam Seçmen Sayısı	Toplam Belde Belediye esi Sayısı	Kırsal Alanda ki Belde Belediye esi Sayısı	Köy ve Belde Belediyesi Toplam Seçmen Sayısı	Belde Belediye esi Seçmen Sayısı	Köy Seçmen Sayısı	Belde Seçmen Oranı	Köy Nüfus Oranı		
1	Altındağ	1	249.620	1	0	0	0	0	0,0%	0,0%	Kapsam Dışı	
2	Çankaya	1	602.249	0	0	0	0	0	0,0%	0,0%		
3	Etimesgut	1	213.577	0	0	0	0	0	0,0%	0,0%		
4	Keçiören	2	548.833	2	0	0	0	0	0,0%	0,0%		
5	Mamak	2	355.858	1	0	0	0	0	0,0%	0,0%		
6	Pursaklar	2	60.284	0	0	0	0	0	0,0%	0,0%		
7	Sincan	2	289.497	1	0	0	0	0	0,0%	0,0%		
8	Yenimahalle	1	453.069	0	0	0	0	0	0,0%	0,0%		
9	Gölbaşı	1	57.879	3	0	1875	0	1875	0,0%	3,2%	Birinci Grup	
10	Akyurt*	2	16.208	0	0	491	0	491	0,0%	3,0%		
11	Elmadağ	2	28.859	3	0	956	0	956	0,0%	3,3%		
12	Kahramankazan*	2	24.958	0	0	1536	0	1536	0,0%	6,2%		
13	Polatlı*	2	74.563	1	0	16377	0	16377	0,0%	22,0%	İkinci Grup	
14	Ayaş*	3	11.150	3	1	3446	1526	1920	13,7%	17,2%		
15	Şereflikoçhisar	3	24.850	4	4	5971	1774	4197	7,1%	16,9%		
16	Çubuk*	3	52.599	3	0	4431	0	4431	0,0%	8,4%		
17	Beypazarı	3	34.296	3	3	9786	2995	6791	8,7%	19,8%	Üçüncü Grup	
18	Kızılcahamam	3	19.636	1	1	7748	624	7124	3,2%	36,3%		
19	Nallıhan*	3	24.163	2	2	14723	6665	8058	27,6%	33,3%		
20	Evren	4	3.405	0	0	1206	0	1206	0,0%	35,4%		
21	Güdül	4	7.938	3	3	5619	2393	3226	30,1%	40,6%		
22	Kalecik*	4	11.282	0	0	4388	0	4388	0,0%	38,9%		
23	Çamlıdere*	5	7.688	0	0	4360	0	4360	0,0%	56,7%		
24	Bala*	5	23.392	3	1	13792	2237	11555	9,6%	49,4%		
25	Haymana	5	28.943	6	5	22284	8896	13388	30,7%	46,3%		
	Toplam		3.224.796	40	20	118989	27110	91879	0,8%	2,8%		

Kaynak: <http://www.ankara.gov.tr/ilcelerimiz>

<http://kutuphane.ankaraka.org.tr/upload/dokumandosya/istatistiklerle-ankara-2010.pdf>

Yukarıdaki tabloda kapsam dışı olan ilçeler, birinci, ikinci ve üçüncü grupta yer alan ilçeler ve seçilen ilçeler gösterilmektedir. Seçilen ilçeler (*) ile gösterilmiştir. Ayrıca ilçelerdeki belde belediyeleri sayılarına da yer verilmiştir. 2004 yılında yasalaşan 5216 sayılı Kanun ile Ankara'da 40 olan belde belediyesi sayısı 20'ye düşürülmüş olup 6360 sayılı Kanunla da bu belde belediyeleri mahalleye dönüştürülmüştür. Bu tablo değerlerine göre görülecek belde belediyesi sayıları ilçe gruplarına göre hesaplanmıştır.

İlçeler aşağıdaki tabloya göre gruplandırılmıştır (Tablo-3).

- ✓ Birinci grup: Merkeze yakın, ilçe gelişmişlik kademesi '2' olanlar, km²'ye düşen nüfusu yüksek olanlar ve kırsal alanı düşük olanlar,

³ <https://www.ysk.gov.tr/doc/dosyalar/docs/2009MahalliIdareler/SecmenSayilari/Belde.htm>
<https://www.ysk.gov.tr/doc/dosyalar/docs/2009MahalliIdareler/SecmenSayilari/Ilce.htm>
<https://www.ysk.gov.tr/doc/dosyalar/docs/2009MahalliIdareler/SecmenSayilari/Koy.htm>
(Erişim Tarihi: 01.10.2022).

- ✓ İkinci grup: Merkeze orta yakınlıkta, ilçe gelişmişlik kademesi ‘3’ olanlar, km²'ye düşen nüfusu orta büyüklükte olanlar ve kırsal alanı orta büyüklükte olanlar,
- ✓ Üçüncü grup: Merkeze en uzak, ilçe gelişmişlik kademesi ‘4’ ve ‘5’ olanlar, km²'ye düşen nüfusu en düşük olanlar ve kırsal alanı en yüksek olanlardan oluşmuştur.

Tablo 3. Örnekleme Seçilen İlçelerin Gruplara Göre Dağılımı

Gruplar	Köy ve Beldelerde Yaşayan Seçmen Oranı Aralığı (%) (2009 Mahalli İdareler Genel Seçimleri)	İlçelerin Alan-Arahlık (Km ²)	İlçelerde Kişiye Düşen Nüfus Aralığı (Kişi Sayısı)	Ankara Şehir Merkezine Mesafeye Aralığı (Km)	2022 SEGE Skoru	Grupların Örneklem Büyüklüğü (Kişi Sayısı)	Mülakat Sayıları (Kişi Sayısı)
Birinci Grup İlçeler (Akyurt, Kazan, Polatlı)	%3-22	369-3618	35-343	29-72	2	106	6
İkinci Grup İlçeler (Ayaş, Çubuk, Nallıhan)	%8,4-60,9	1041-2079	37-247	37-132	3	103	6
Üçüncü Grup İlçeler (Kalecik, Çamlıdere, Bala)	%39,4-59	782-1851	24-154	48-70	4-5	154	6
Toplam						363	18

1.3 Örneklem Büyüklüğünün Yerel Yönetim Birimlerine Dağılımı

Aşağıdaki tabloda toplam örneklem büyüklüğünün köy ve belde belediyesi yerel yönetim birimlerine göre dağılımı yer almaktadır. Alan araştırması bu dağılıma uygun olarak gerçekleştirilmiştir.

Tablo 4. Ankara İlinin 6360 Sayılı Kanun Öncesi Yerleşim Yerlerinin Dağılımı

Açıklama	Orman Köyü Sayısı (Orman İçi ve Bitişik)	Diğer Köy Sayısı	Toplam Köy Sayısı (Orman Köyü+Diğer Köy)	Belde Belediyesi Sayısı	Toplam
Ankara	397	287	684	42*	726
Köy/Belde Bel. Dağılımı (%)	54,7%	39,5%	94,2%	5,8%	100%
Örneklem Sayıları	199	143	342	21	363

Kaynak: (Adıgüzel, 2012, s. 161)

1.4 Mahallelerin Seçimi

Aşağıdaki tabloda görüldüğü üzere 6360 sayılı Kanunun yasalaşmasıyla Ankara'daki mahalle sayısı %90, artarak 1531'e yükselmiştir.

Tablo 5. 6360 Sayılı Kanun İle Ankara İlinin Artan Mahalle Sayısı

İl	6360 Kanun Öncesi Mahalle Sayısı	6360 Kanun Sonrası Mahalle Sayısı	Toplam Mahalle Sayısı	Artış Yüzdesi (%)
Ankara	805*	726	1531	90,2

İlçeler seçilmesini takiben mahalleler de aşağıda belirtilen üç farklı değişkene göre seçilmiştir.

Değişken-1: Mahallenin bağlı olduğu ilçe merkezine uzaklığı,

Değişken-2: Eskiden belde belediyesi veya köy statüsünde olma,

Değişken-3: Orman içi ve bitişik mahallesi veya mahalle olma

İlçelerde uygulanacak anket sayıları her birime yaklaşık eşit birimde (her mahalleye ortalama 10 anket olacak şekilde) dağıtılmıştır.

Kapatılan belde belediyeleri kapsamında Ayaş'ın Oltan, Nallıhan'ın Çayırhan, Bala'nın Afşar belde belediyelerinde anket ve mülakatlar yapılmıştır. Üç gruptan bir belde belediyesi seçilmiştir.

1.5 Alan Araştırmasının Gerçekleşmesi

Alan Araştırması 2022 yılı Temmuz ayında Ankara İlının 9 ilçesinde (Akyurt, Kazan, Polatlı, Ayaş, Çubuk, Nallıhan, Kalecik, Çamlıdere, Bala) 380 kişi ile yüz yüze anket yapılarak gerçekleştirilmiştir. Alan araştırması sonucunda anket ve mülakat görüşmeler, ilçelerin grup içindeki büyüklüklerine, mahalle statü dağılımlarına, Ankara il geneli yaş ve cinsiyet oranlarına uyumlu olarak gerçekleştirilmiştir. Toplam 32 yerleşim yerinde (3 belde belediyesinde -Çayırhan, Oltan, Afşar- ve 29 mahallede) 380 anket yüz yüze ve 19 mülakat yapılmıştır. Aşağıdaki tabloda alan çalışması sonucunda yapılan görüşmelerin ilçelere göre cinsiyet, yaş aralığı ve mahalle statüsüne göre dağılımı yer almaktadır.

Tablo 6. Alan Araştırması Sonucunda İlçelere Göre Gerçekleşen Sayılar

Cinsiyet	13 Tem-23 Tem	14.Tem	15.Tem	16.Tem	24.Tem	24.Tem	26.Tem	27.Tem	28.Tem	Toplam	Yüzde (%)
	Polatlı	Ayaş	Çubuk	Çamlıdere	Akyurt	Kazan	Bala	Nallıhan	Kalecik		
Kadın	37	11	22	14	3	4	11	19	32	153	40,3%
Erkek	51	26	24	16	7	8	32	33	30	227	59,7%
Toplam	88	37	46	30	10	12	43	52	62	380	100,0%
Yaş aralığı											
18-25	6	1	5	4	0	0	0	4	4	24	6,3%
26-33	3	2	3	3	1	1	2	1	0	16	4,2%
34-41	4	2	4	1	1	1	3	2	6	24	6,3%
42-49	10	3	7	8	2	0	4	7	5	46	12,1%
50-57	23	9	6	1	2	4	9	8	15	77	20,3%
58-65	18	7	9	9	1	4	12	14	13	87	22,9%
66-75	22	9	8	4	3	2	12	11	15	86	22,6%
76 ve üzeri	2	4	4	0	0	0	1	5	4	20	5,3%
Toplam	88	37	46	30	10	12	43	52	62	380	100,0%
Mahalle Statüsü											
Köy	45	11	20	8	0	0	15	16	22	137	36,1%
Orman İçi-Bitişiği Köy	43	14	26	22	10	12	15	24	40	206	54,2%
Belde Belediyesi	0	12	0	0	0	0	13	12	0	37	9,7%
Toplam	88	37	46	30	10	12	43	52	62	380	100,0%

Planlanan tabaka büyüklükleri 3 grup için sırasıyla 107, 123, 133 iken 110, 135, 135 olarak gerçekleştirilmiştir. Örneklem planlamasında %54,7 orman köyü, %39,5 diğer köy ve %5,8 belde belediyesi iken alan araştırması sonucunda bu oranlar sırasıyla %54,2, %36,1 ve %9,7 olarak gerçekleştirilmiştir.

Tablo 7. Muhtar Mülakatlarının İlçelere Göre Dağılımı (Nitel)

Sayı	Polatlı	Ayaş	Çubuk	Çamlıdere	Akyurt	Kazan	Bala	Nallıhan	Kalecik	Toplam
Planlanan	2	2	2	2	2	2	2	2	2	18
Gerçekleşen	3	3	2	2	1	1	2	3	2	19

Mülakatlar kapsamında 11 muhtar ile yüz yüze 9 muhtarla telefonla görüşerek, ses kaydı alınarak yapılmıştır. Telefon mülakatları, mahalle ziyaretlerinin gerçekleştirildiği tarihlerde, muhtara çeşitli nedenlerle (Sağlık sorunları, tarlada vb. olması, şehir merkezinde olması vd.) erişilememesi nedeniyle yapılmıştır.

1.6 Karşılaştırmalı Analizler

Toplanan nicel veriler SPSS programında analiz edilmiş olup, nitel veriler ise içerik analizine tabi tutulmuştur. Tezin hipotezlerde belirtilen ilişkilerin var olup olmadığı ve varsa ilişkilerin yönü (olumlu ya da olumsuz) tespit edilmiştir.

Ayrıca nicel veriler “ilçe belediyelerine”, “coğrafi alan büyüklüğüne”, “kırsal alan büyüklüğüne”, “nüfusuna”, “gelişmişlik düzeyine”, “ilçe merkeze uzaklığına”, “yönetim birimi türlerine” ve “demografik değişkenlere” göre karşılaştırmalı biçimde analiz edilmiştir. Araştırmada temel karşılaştırmalar “ilçe merkezine olan uzaklık”, “mahallenin daha önceki statüsü (köy, orman içi, bitişiği köy, belde belediyesi)”, “yerel halkın yerleşim yerlerindeki yaşama durumları (sürekli, mevsimlik, hafta sonu vd.)”, “ilçe gelişmişlik seviyeleri” ve “demografik özelliklere” göre yapılmıştır.

2. BENZER ÇALIŞMALARDA KULLANILAN ARAŞTIRMA YÖNTEMLERİ

İller	İlçeler	Yöntem	Araştırma Tekniği	Görüşme Yapılan Denek Tipleri	Kitle/Evren	Örneklem Sayısı	Değerlendirme	
MAKALE 1: Ökmen (2016), "Köylerin Mahalleye Dönüşüm Kararının Etkilerini Belirlemeye Yönelik Bir Araştırma: Manisa Büyükşehir Belediyesi" adlı makale								
1	Manisa	Manisada'ki 17 ilçenin 15'i (Ahmetli, Alaşehir, Demirci, Gölmarı, Kırkçağ, Köprübaşı, Kula, Salihli, Sarıgöl, Saruhanlı, Selendi, Soma, Şehzadeler, Turgutlu, Yunusemre)	Nicel	Anket (Yüz yüze?)	Vatandaş	Evren belirtilmemiş	257	Açıklama: Veri analiz yapılmamış Microsoft Excel 2010 programı ile yüzdesel dağılımlar belirlenmiş ve buna bağlı olarak grafikler oluşturulmuştur. Değerlendirme: 17 ilçenin 2 tanesi neden seçilmemiş? Niye 15'i seçilmiş? Benzer ilçe grupları arasında daha az ilçe seçilebilirdi. 257 kişi nasıl hesaplanmış? 257 kişinin ilçelere dağılımı nasıl belirlenmiş?
MAKALE 2: Göküş (2016), "6360 Sayılı Yasa İle Köy Yönetiminden Mahalle Yönetimine Geçiş Sonrası Etkin Hizmet Sunumuna İlişkin Muhtarların Görüşü: Silifke Köyleri Örneği" adlı Makale								
2	Mersin	Silifke	Nitel	Görüşme/ Mülakat	Muhtarlar	Araştırmanın evrenini Mersin ili, örneklemini ise Silifke ilçesine bağlı 6360 sayılı Yasa ile köy yönetiminden mahalle yönetimine dönüşmüş mahalleler oluşturmuştur.	66 muhtarın 36'sı ile görüşme yapılmış	Açıklama: Ulaşım zorluğu, zaman darlığı ve maliyetler nedeni ile 6 muhtarla yüz yüze görüşme yapma imkânı olmuş, diğerleri ise telefonla aranarak iletişim kurulmuştur. 38 mahalle muhtarının 2'si ile görüşme şansı bulunamamış. Değerlendirme: "Silifke ilçesinde 66 köy mahalleye dönüştürülmüştür" araştırmanın evreni budur. Nitel araştırma sonuçları nadiren de olsa nicel araştırma gibi verilen cevaplar % oransal olarak tablo biçiminde verilmiştir. 36 görüşmenin 6'sı ile yüz yüze görüşülmüş ise araştırma tekniği telefonla görüşmedir. (yüz yüzyüze+telefon yarı yapılandırılmış/tam yapılandırılmış?)
DOKTORA TEZİ 1: Avsar (2020), "Türkiye'de 6360 Sayılı Yasa Kapsamında Mahalle ve Köy Yönetiminde Yaşanan Değişim ve Dönüşüm" başlıklı doktora tezi								
3	7 bölgede 8 il (Kayseri, Erzurum, Samsun, Mardin, Tekirdağ, İzmir, Adana ve Kahramanmaraş)	-	Nicel	Anket	Muhtarlar, valiler, kaymakamlar, büyükşehir belediye başkanlarına ve ilçe belediye başkanlarına	Vali ve kaymakamların, Büyükşehir ve Büyükşehir ilçe belediye başkanlarının toplamı 268 kişi: muhtarların toplamı 4688 kişi	Birinci Grup: 159 İkinci grup: 340	Mülki amirler ve muhtarlar ayrı ayrı analiz edilmiş ve nicel tablolara yer verilmiştir. Her iki grup için 2 ayrı soru formu hazırlanmıştır. 7 Bölge İller? Birinci grup ayrılabilir miydi? İlçeler nasıl seçildi?
Y.LİSANS TEZİ 1: Şahin (2019), "6360 Sayılı Kanunla Birlikte Köy Tüzel Kişiliğinin Kaldırılmasıyla Oluşan Mahallelere Verilen Kamu Hizmetlerinin Değerlendirilmesi: Tekirdağ Örneği" adlı Y.Lisans Tezi								
4	Tekirdağ	Süleymanpaşa, Çorlu, Şarköy ve Marmaraeğlisi	Nicel	Yüz yüze, mail ve telefon aracılığıyla anket çalışması	Muhtarlar	Evren belirtilmemiş	55	11 ilçeden 4ü hangi nedenle seçilmiş? İlçeler arası karşılaştırma yapılmamış. 55 muhtar nicel çalışma için yeterli midir? 55 sayısı neye göre hesaplanmış?

İller	İlçeler	Yöntem	Araştırma Tekniği	Görüşme Yapılan Denek Tipleri	Kitle/Evren	Örneklem Sayısı	Değerlendirme	
<u>Y.LİSANS TEZİ 2:</u> Karpuz (2019), “6360 Sayılı Kanun Sonrası Mahalleye Dönüşen Beldeler Açısından Bir Değerlendirme: Akhisar Örneği” Yüksek Lisans Tezi								
5	Manisa	Akhisar (Akhisar özelinde saha çalışması yapılmasının nedeni ise hem Manisa'nın en büyük ilçesi olması hem de kapatılan belde sayılarının sayıca fazla olmasıdır)	Nicel	Anket	Eski belediye başkanları, muhtarları ve mahalle halkından	Araştırmanın evreni 6360 Sayılı Kanun'la Büyükşehir İlçe Belediyesi olan Akhisar'ın mahalleye dönüşen 9 yerleşim yerinin eski belediye başkanları, muhtarları ve halkından oluşmaktadır.	Örneklem ise farklı meslek ve yaş gruplarından oluşan 180 kişiden oluşmaktadır.	Açıklama: Mahalleye dönüşen 9 yerleşim yerinin 9 mahalle seçilmiş Değerlendirme: Belediye başkanı, muhtar ve vatandaşa aynı soru seti uygulanmamış ve bu 3 grubun yanıtları karşılaştırmalı ele alınmadığı gibi 180 kişi cevapları birlikte değerlendirilmiştir. Dolayısıyla 3 grupta araştırma yapıldığı söylenemez
<u>MAKALE 3 :</u> Çavuşoğlu (2020), “6360 Sayılı Kanun’un Mahalleye Dönüşen Birimler Üzerine Etkisi: Antalya İli Örneği” adlı makale								
6	Antalya	18 ilçe	Nicel	Anket	Muhtarlar	Araştırma çalışmasının evrenini, Antalya ilinde yer alan 19 ilçede bulunan, 6360 sayılı Kanun ile kapatılan 543 köy/yeni mahalle 74 belde belediyesi/yeni mahalle oluşturmaktadır	133	Açıklama: kapatılan 543 köy içerisinde nüfusa ve toplam kapatılan köy sayısı temel kriter alınarak seçilmiş 133 eski köy/yeni mahalle muhtarıyla Değerlendirme: 20 Belde belediyesi ve 113 köy muhtarı ile görüşme yapılmış ama tümü aynı kefedede değerlendirilmiştir. 18 ilçeye gitmek yerine birbirine benzer ilçeler bulunarak daha az sayıda ilçeye gidilebilirdi. Gidilmeye 1 ilçeye neden gidilmedi?
<u>MAKALE 4:</u> Bölük ve Ömürgönülşen (2020), “6360 Sayılı Kanun ile Büyükşehir Belediyeleri Sınırları İçerisinde Kapatılan ve Mahalle Statüsüne Dönüştürülen Belde Belediyeleri ve Köylerin Durumu: Bodrum Örneği”, adlı Makale								
7	Muğla	Bodrum	Nitel	Yüz yüze ve yarı-yapılandırılmış mülakat	Belde belediye başkanı ve muhtarlar	Kanun gereğince; kamu tüzel kişiliği sona eren Bodrum'daki 10 belde belediyesi ve 20 köy	22	Açıklama: 10 belde belediyesinden 4 beldenin belediye başkanı, 18 eski köy muhtarı ve/veya hâlihazırdaki mahalle muhtarı Değerlendirme: Bunlardan Muğla Büyükşehir Belediye Başkanı, 4 eski beldenin belediye başkanı, 18 eski köy muhtarı ve/veya hâlihazırdaki mahalle muhtarıyla görüşülmüştür. Analizler her iki grup için ayrı ayrı değil toplu yapılmıştır.
<u>MAKALE 5:</u> Gültekin (2021), “6360 Sayılı Kanun İle Mahalleye Dönüştürülen Köylerde Tarımsal Üretim: Kahramanmaraş İli Örneği” adlı makale								
8	Kahramanmaraş	Onikişubat ve Dulkadiroğlu	Nicel	Anket	Tarımsal Üreticiler	5034 üretici	128 üretici	Değerlendirme: Neden bu iki ilçe? Anketlerin %45,3'ü Dulkadiroğlu ve %54,7'si Onikişubat ilçelerinde ikamet eden Üreticilerle yapılmıştır. Denmiş ama 2 ilçedeki üretici oranı verilmemiş. Ve en önemlisi bu iki ilçe sonuçları karşılaştırılmamış.

Benzer çalışmalardan elde edilen sonuçlar aşağıdaki biçimdedir:

- 1- Örneklem çerçevesi (il, ilçe, mahalle seçimleri, sayıları nedenleri) ile ilgili açıklamalar yeterli değil ya da hiç bulunmamaktadır.
- 2- Evren/kitle bilgileri, çerçevesi, büyüklüğü birçok eserde belirtilmemektedir.
- 3- Örneklem büyüklüğünün nicel araştırmalar için nasıl hesaplandığı belirtilmemektedir.
- 4- Bir ilin bütün ya da çoğu ilçelerin seçilmesi yaklaşımı vardır ki buna gerek yoktur benzer ilçeler içinden daha az sayıda ilçe seçilerek araştırma odaklanmalıdır.
- 5- Belirlenen örneklem büyüklüğünün ilçelere nasıl dağıldığına ilişkin bir yaklaşım yoktur. Oysaki nüfus değişkeni doğrudan bu konuda belirleyicidir.
- 6- Nicel araştırmalarda demografik verilerin dağılımına dikkat edilmediği gözlenmiştir. Bu konu özelinde 18 yaş ve üzeri kişilerle anket yapılmalıdır.
- 7- Araştırma tekniği belirtilmesi kısmı da birçok eserde yeterli değildir. Anket tekniği, anketin yüz yüze mi, telefonla mı yapıldığı, bu tekniklerin neden tercih edildiği belirtilmelidir. Mülakat tekniğinde soru formunun yapılandırılmış mı, yarı yapılandırılmış mı olduğu belirtilmelidir. Hangi yöntem daha fazla kullanılmış ise bu yöntem öncelikli yazılarak ana yöntem olarak belirtilmelidir.
- 8- Mecbur kalınmadıkça aynı grup denek için birden fazla görüşme tekniği (yüz yüze, telefon, çevrimiçi gibi) kullanılmamalıdır. En doğru yöntem tek bir görüşme tekniği kullanılmalıdır. Zorunlu kullanılması halinde gerekçesi belirtilmelidir.
- 9- Araştırma sorusuna göre nitel araştırma yöntemi mi nicel araştırma yöntemi mi kullanılacağı konusunda bir kararsızlık/karışıklık görülmektedir. Burada en önemli hata farklı statüye sahip deneklerin tümünün aynı kategoride değerlendirilerek, aynı soru formu uygulanarak sonuçların analiz edilmesidir. Örneğin belediye başkanı, muhtar ve vatandaş. Buradaki doğru yaklaşım eğer bir belediye başkanları araştırması yapılmıyor ise belediye başkanı (az sayıda bu özellikte kişi olması nedeniyle) ile mülakat yapılması ve genel araştırma sonuçlarına dahil edilmemesidir. Ayrıca bu üç grup için de farklı soru setleri hazırlanmalıdır.
- 10- Nitel araştırma sonuçları oransal % olarak verilmemelidir. Tek bir denegin bile söylediği bazen araştırmacı için çok doğru ve geçerli olabilecektir (Makale 2).

Görüşme yapılan muhtarlardan bir kısmı hizmetlerin etkin bir şekilde sunulduğunu ve 6360 sayılı yasanın uygulamaya geçmesinden memnun olduğunu belirtirken, muhtarların bir kısmı hem yasayı eleştirmiş hem de etkin bir hizmet sunumu olmadığını belirtmiştir. Aşağıdaki tabloda ilçe belediyesinin ve büyükşehir belediyesinin etkin hizmet sunumuna ilişkin muhtarların vermiş olduğu cevaplardan elde edilmiş veriler yer almaktadır.

Tablo 3: İlçe ve Büyükşehir Belediyesi Etkinlik Tablosu¹¹

	İlçe Belediyesi		Büyükşehir Belediyesi	
	N	Yüzde	N	Yüzde
Etkin	1	2,8	11	30,6
Etkin Değil	35	97,2	25	69,4

- 11- Evren ve örneklem sayısına göre nicel ya da nitel araştırma yapılmasına karar verilmelidir.
- 12- Örnekleme birden fazla ilçe alınması gerekçesiyle birlikte belirtildikten sonra gruplandırma yapılmasa bile ilçeler arası bir karşılaştırma yapılması araştırma sonuçlarına zenginlik katacaktır. Araştırmalarda karşılaştırmalı analizlere yer verilmesi çok önemlidir.
- 13- Bir ilin bütün ilçelerine gitmek yerine birkaç değişkene (Örn: Gelişmişlik, nüfus, kırsal alan büyüklüğü vb.) bakılarak ilçeler gruplandırılmalı ve daha az sayıda ilçeye odaklanılmalıdır. Bu yöntem istatistik bilimini kullanmakla birlikte zaman, maliyet, planlama, doğru sonuca ulaşma gibi birçok konuda fayda sağlayacaktır.
- 14- Farklı statüdeki deneklerle (belediye başkanı, muhtar) yapılan araştırmalarda elde edilen sonuçlar aynı araştırma yöntemi (nicel/nitel) kullanılsa bile her grup için ayrı ayrı yapılmalıdır. Ayrıca örneğin aşağıda tabloda olduğu gibi farklı özelliklere sahip grupların frekansları her grup için ayrı ayrı verilmelidir. Araştırma eski belediye başkanları, muhtarları ve mahalle halkı ile anket yapılmıştır (Y. Lisans Tezi 2).

Cinsiyet Dağılımı:

cinsiyetiniz?		Frequency	Percent	Valid Percent	Cumulative Percent
Erkek		167	92,8	92,8	92,8
Valid Kadın		13	7,2	7,2	100,0
Total		180	100,0	100,0	

Ankete katılanlardan 167'si erkek, 13'ü kadından oluşmaktadır. Bunlardan %92.8'i erkekleri, %7'si kadınları temsil etmektedir. Katılımcıların büyük çoğunluğunu erkekler oluşturmaktadır. Bunun nedeni; kadınların ankete katılma konusunda çok istekli olmamalarıdır. Kırsalda yaşayan erkekler daha çok kıraathanelerde bulunmaktadır. Bunun üzerine kıraathanelere gidilerek katılımcılara ulaşılmıştır.

- 15- Araştırmada evren büyüklüğü belirlendiği halde bu büyüklüğün seçilen ilçelere dağılımının hesaplanmaması, toplamda hedeflenen görüşme sayısının (örneklem) ilçelere dağılımının hesaplanmaması ve bu ilçe sonuçlarının karşılaştırmalı analiz edilmemesinin araştırmalar açısından önemli bir eksiklik olduğu düşünülmektedir.

GENEL DEĞERLENDİRME

Araştırmacıların örneklem çerçevesi ile ilgili açıklamalarının “neden” soruna yanıt verecek biçimde yeterli olması, evren ait bilgiye sahip olmaları gerekmektedir. Örneklem büyüklüğünün de nicel çalışmalara da mutlaka bir formüle dayanması gereklidir. Örneklem dağılımının da alt tabakalara olasılıksal olarak dağılımının sağlanması ve gerçekleşme rakamlarının da bu oranlarla uyumlu olması beklenmektedir. Her hangi bir araştırma tekniğinin araştırma konusuna uygun olup olmadığına karar verilirken öncelikle araştırma sorusunun cevabını elde etmek için kullanılacak teknikle cevap alınıp alınmayacağı çok iyi analiz edilmelidir. Bu aşamada kuramsal bilgi önem taşımaktadır. Farklı statüdeki deneklerin tek bir soru formuna ya da tek bir araştırma tekniğine tabi tutulması konusunda araştırmacılar çok dikkatli olmalıdır. Bunun yerine karma yöntemle çok daha gerçekçi sonuçlar elde edilebileceği akılda tutulmalıdır. Bir yönetim biriminin tüm alt birimlerinde örnek seçmek yerine alt birimler arasında çeşitli değişkenler yolu ile birbirine benzer alt birimleri gruplandırılmalı ve daha az sayıda alt birim seçmelidir. Bu şekilde hem istatistik biliminden faydalanmak hem de zaman ve maliyetten tasarruf etme yolu seçmiş olacaktır.

SONUÇ VE ÖNERİLER

Birincil veri üretmeye dayalı yöntemlerin ele alındığı bu bildiriye söz konusu araştırmacıların alan araştırmasında kullandıkları yöntemler konusunda çerçevesi ve yeterli bilgiye sahip olmadıkları söylenebilir. Sadioğlu ve Yıldız 2007 yılında yayınladıkları 80 yayın üzerinden yaptıkları bibliyografi sonucuna göre makalelerin %21,3'ünün ilk elden toplanan (anket, mülakat, gözlem vd.) verilere dayandığı sonucunu elde etmişlerdir (Sadioğlu ve Yıldız, 2008:336). Bu oranın günümüzde kısmen arttığı tahmin edilse de kuramsal eserlere göre oldukça düşük oranda olduğu göz önüne alındığında alan araştırmasına dayalı yapılan her bir çalışmanın ne kadar büyük önemde olduğu söylenebilir. Ancak alan araştırması sürecindeki her aşama (Doğru araştırma sorusunun belirlenmesi, araştırma sorusuna yönelik uygun araştırma yönteminin belirlenmesi, doğru örnekleme büyüklüğünün tespiti ve örneklem dağılımının yapılması, araştırma sürecinin yönetimi, araştırma süresi gibi.) doğru sonuca ulaşmada büyük öneme sahiptir. Bu bakış açısıyla kuramsal bilginin yanında araştırmacıların araştırma modelini kurarken bilimsel araştırma yöntemleri hakkında yeterli bilgiye sahip olmaları gerektiği gibi diğer araştırmacıların kurdukları iyi uygulama modellerini incelemeleri de önerilmektedir. Karma modelin sosyal bilimlerde daha zengin ve tamamlayıcı sonuçlar sunması açısından kullanılması önerilmektedir.

REFERANSLAR

- Adıgüzel Ş, 2012. 6360 Sayılı Yasa'nın Türkiye'nin Yerel Yönetim Dizgesi Üzerine Etkileri: Eleştirel Bir Değerlendirme. *Toplum ve Demokrasi Dergisi*(13-14), 153-176.
- Altunışık R., Coşkun R., Bayraktaroğlu S, Yıldırım E, 2010. Sosyal Bilimlerde Araştırma Yöntemleri SPSS Uygulamalı (6. Baskı). Sakarya: Sakarya Yayıncılık, Türkiye.

- Avşar Y, 2020. Türkiye'de 6360 Sayılı Kanun Kapsamında Mahalle Ve Köy Yönetiminde Yaşanan Değişim Ve Dönüşüm. *Doktora Tezi*. Kahramanmaraş Sütçü İmam Üniversitesi, Kahramanmaraş. Türkiye.
- Aydın F, 2024. Yeni Büyükşehir Yönetim Modelinde İlçe Belediyelerinin Hizmet Sunumlarına İlişkin Vatandaş Memnuniyet Düzeyini Ölçmeye Yönelik Bir Araştırma: Ankara Örneği. *Doktora Tezi*. Hacettepe Üniversitesi, Sosyal Bilimler Enstitüsü, Ankara, Türkiye.
- Bölük A, Ömürganülşen U, 2020. 6360 Sayılı Kanun ile Büyükşehir Belediyeleri Sınırları İçerisinde Kapatılan ve Mahalle Statüsüne Dönüştürülen Belde Belediyeleri ve Köylerin Durumu: Bodrum Örneği. U. Sadioğlu, M. K. Öktem, & U. Ömürganülşen içinde, *Sanayileşme ve Göç Kıskaçında Kent ve Çevre Yönetimi Üzerine Yazılar II* (s. 262-286). Ankara: Hacettepe Üniversitesi, Türkiye.
- Burns N, Grove S. K, 1993. *The practice of nursing research: Conduct, critique & utilization* (2nd Ed.). Elsevier Science Health Science.
- Çavuşoğlu A L M, 2020. 6360 Sayılı Kanun'un Mahalleye Dönüşen Birimler Üzerine Etkisi: Antalya İli Örneği. *Bartın Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 11(21).
- Çıngı H, 1994. *Örnekleme kuramı* (2. Baskı). Ankara: Hacettepe Üniversitesi Basımevi, Türkiye.
- Fraenkel J R, Wallen N E, 2006. *How to design and evaluate research in education* (6. Ed.). New York: McGraw-Hill International Edition.
- Göküş M A, 2016. 6360 Sayılı Kanun ile Köy Yönetiminden Mahalle Yönetimine Geçiş Sonrası Etkin Hizmet Sunumuna İlişkin Muhtarların Görüşü: Silifke Köyleri Örneği. *Selçuk Üniversitesi Sosyal ve Teknik Araştırmalar Dergisi*(11), 67-86.
- Gültekin R 2021. 6360 Sayılı Kanun İle Mahalleye Dönüştürülen Köylerde Tarımsal Üretim: Kahramanmaraş İli Örneği. *Türk Tarım ve Doğa Bilimleri Dergisi* (8(4)), 974-982.
- Karataş Z, 2015. "Sosyal Bilimlerde Nitel Araştırma Yöntemleri", *Manevi Temelli Sosyal Hizmet Araştırmaları Dergisi* Cilt 1, Sayı 1, Ocak 2015 syf:62-80.
- Karpuz G İ, 2019. 6360 Sayılı Kanun Sonrası Mahalleye Dönüşen Beldeler Açısından Bir Değerlendirme: Akhisar Örneği. *Yüksek Lisans Tezi*. Balıkesir.
- Leech N L, Onwuegbuzie A J, 2009. A typology of mixed methods research designs. *Qual Quant*. 43, 265–275.
- Neuman W L, 2012. *Toplumsal Araştırma Yöntemleri: Nicel ve Nitel Yaklaşımlar I-II*. Cilt (5. Basım). İstanbul: Yayın Odası.
- Ökmen M Ç, 2016. Köylerin Mahalleye Dönüşüm Kararının Etkilerini Belirlemeye Yönelik Bir Araştırma: Manisa Büyükşehir Belediyesi Örneği. *Kamu Yönetiminde Değişimin Yönü Ve Etkileri (KAYFOR 13) Bildiri Kitabı*, (s. 1475-1487).
- Özdemir M, 2010. Nitel Veri Analizi: Sosyal Bilimlerde Yöntembilim Sorunsalı Üzerine Bir Çalışma. *Eskişehir Osmangazi Üniversitesi Sosyal Bilimler Dergisi*, 11(1), 323- 343.
- Sadioğlu U, Yıldız M, 2007. Kamu Yönetimi İle Bilgi Ve İletişim Teknolojileri: Bir Bibliyografik Analiz. *Hacettepe Üniversitesi İktisadi Ve İdari Bilimler Fakültesi Dergisi*, 25(2), 323-365.
- Şahin F, 2019. 6360 Sayılı Kanunla Birlikte Köy Tüzel Kişiliğinin Kaldırılmasıyla Oluşan Mahallelere Verilen Kamu Hizmetlerinin Değerlendirilmesi: Tekirdağ Örneği,. *Yüksek Lisans Tezi*(30), 221-230. Bartın: Bartın Üniversitesi.
- Tunalı B, S, Gözü Ö, Özen G, 2016. Nitel ve Nicel Araştırma Yöntemlerinin Bir Arada Kullanılması "Karma Araştırma Yöntemi". *Kurgu*, 24(2), 106-112.
- Yıldırım A, Şimşek H, 2008. *Sosyal Bilimlerde Nitel Araştırma Yöntemleri* (6. Baskı). Ankara: Seçkin Yayıncılık.

İlaç Sektöründe Gelişmiş Talep Tahmini: Entegre Makine Öğrenmesi ve Derin Öğrenme Yaklaşımları

Recep Yılıkcı^{1*}, Ece Özbilen¹

¹ Eczacıbaşı Bilişim San. ve Tic. A.Ş., Eczacıbaşı Bilişim Ar-Ge Merkezi, İstanbul, Türkiye

Özet

Bu çalışmada, ilaç sektöründe yanlış stok yönetiminin yol açabileceği hayati krizleri ve hammadde israfını önlemek amacıyla, tedarik zincirinin kritik bir bileşeni olan ilaç talep tahminlerinin iyileştirilmesi amaçlanmıştır. Geleneksel zaman serisi yöntemlerinin ötesine geçilerek, modern makine öğrenmesi ve derin öğrenme algoritmaları kullanılmıştır. Bu sayede, yalnızca satış verileri değil, pazar hacmi, döviz kuru, enflasyon ve kampanyalar gibi satışları etkileyebilecek diğer faktörler de dikkate alınmıştır. Analizler sonucunda, pazar verilerinin satışlar üzerinde en yüksek etkiye sahip olduğu, enflasyonun ise daha düşük etkili olduğu tespit edilmiştir. Geliştirilen modeller, 126 farklı ürün üzerinde test edilmiş ve her ürün için ideal model seçimi yapılmıştır. Çalışmada ürün portföyünün yaklaşık %40'ının MAPE değeri %20'nin altında olacak şekilde tahmin doğruluğu sağlanmıştır. Vitamin dışı ürünler için ortalama MAPE %34,77, vitamin ürünleri için ise %55,07 olarak hesaplanmıştır. Geliştirilen modeller, dinamik ve sürekli güncellenen bir sistem olarak canlıya alınmıştır.

Anahtar Kelimeler: Talep Tahmini, Makine Öğrenmesi, Derin Öğrenme, Hiperparametre Optimizasyonu, İlaç Tedarik Zinciri

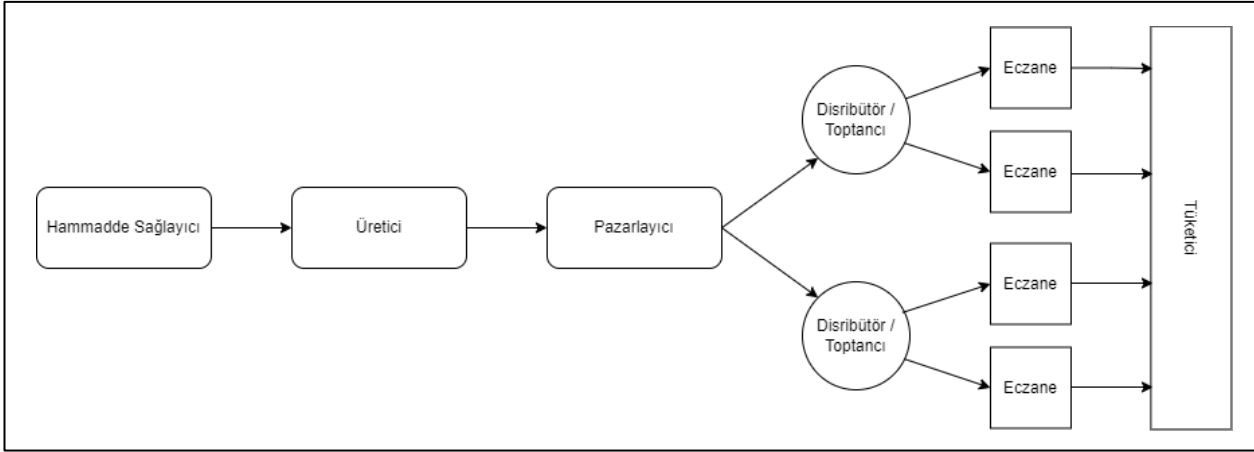
GİRİŞ

Talep tahmini, birçok sektörde tedarik zinciri sürecinin önemli bir bileşenidir. İlaç sektöründe bu süreç kendine özgü ve karmaşık bir yapıda karşımıza çıkar. Ürünlerin hayati öneme sahip olması nedeniyle daha da önemlidir. Epidemiler ve pandemiler gibi beklenmedik sağlık krizleri, bu sektöre özgü düzenleyici kısıtlamalar, ilaçların uzun geliştirme döngüleri ve benzeri belirsizlikler talebin dramatik bir şekilde değişmesine ve stok problemine neden olur (Atasever, 2015). Bu tarz durumları tahmin etmek oldukça güçtür (Merkuryeva vd., 2019).

Doktorların reçetedeği eğilimleri, fiyat, üretim maliyeti, tedarik kıtlıkları, rekabet ortamı ve daha fazlası gibi popülasyonda ilaç kullanımını etkileyen çok sayıda faktör bulunmaktadır. Başarılı bir tahmin modeli tüm bu faktörleri aynı anda hesaba katmalıdır. Ancak mevcut modellerin çoğu yalnızca satış verilerine dayanmaktadır (Pall vd., 2023). Bunun en büyük sebebi tüm bu verileri toplamak oldukça zor ve maliyetlidir. Bu yüzden de tahmin başarısını artırmadaki en büyük rol bahsi geçen faktörlerden sonra algoritmalara kalmaktadır. Bu konuda yapılan çalışmalar incelendiğinde, geçmişteki satış verilerinden yararlanılarak ARMA modeli ile ilaç satış tahmininde başarı elde edilmiştir (Ahmad vd., 2015). Diğer bir çalışmada derin öğrenme algoritmalarıyla ve harici veri kullanmadan satış tahmini iyileştirilmiştir (Chang vd., 2017).

Son yıllarda veriye ve güncel algoritmalara erişim kolaylaştıkça diğer faktörlerin de dikkate alındığı çalışmalar yapılmıştır. Örnek bir çalışmada geçmiş satış verisinin dışında promosyon ve ürün fiyatı gibi harici veriler de kullanılarak 4 farklı (Holt Winters, Random Forest, Ridge Regression ve XGBoost) makine öğrenmesi yöntemi ile başarılı sonuçlar elde edilmiştir (İmece ve Beyca, 2022). ARIMA ve derin sinir ağları yöntemleriyle tahmin modelleri oluşturularak kıyaslama yapılan bir çalışmada derin sinir ağları modellerinin ARIMA modeline göre daha düşük sapmayla talebi tahmin ettiği görülmüştür (Rathipriya vd., 2023). Son olarak, bu alandaki tahmin zorluklarını ele alarak, tahmin doğruluğunu artırmak için kullanılabilir veri ve yöntemlerden bahsedilmiştir (Tolosşi-Halacheva vd., 2024).

Bu çalışmada amacımız çok katmanlı bir tedarik zinciri sürecine sahip olan (bkz. Şekil 1) ilaç sektöründe, tek değişken üzerinden hareket eden geleneksel zaman serileri yöntemleri kullanmak yerine, birçok farklı bağımsız değişkeni (döviz kurları, enflasyon, pazar hacmi vb.) probleme dahil etmek ve Prophet, LSTM (Long-Short Term Memory), XGBoost, LightGBM ve CatBoost olmak üzere 5 modern ve güçlü algoritma kullanarak probleme yenilikçi bir çözüm sunmak ve tahmin doğruluğunu arttırmaktır.



Şekil 1. İlaç Tedarik Süreci

ARAÇ VE YÖNTEMLER

Metodlar

Verilerin Toplanması

Çalışma kapsamında toplam 126 farklı ürün ele alınmıştır. Bu ürünler kendi içinde onkolojik, vitamin, temel tedavi ürünleri gibi farklı gruplara ayrılmaktadır. Her bir ürünün geçmiş satış verisi, satış fiyatları ve bulunduğu pazarın hacim büyüklüğü verileri aylık bazda bulunmaktadır.

18 ürün vitamin grubunda yer almaktadır. Vitamin grubunun tahmin doğruluğunu arttırmak için harici veriler çalışmada toplanmıştır ve 6 değişken dahil edilmiştir. Yapılan ve planlanan kampanyaların maliyeti, resmi tatiller, ÖSYM ve MEB sınav tarihleri, okul dönemleri, enflasyon ve döviz kuru verileri internet üzerinden çekilerek Boosting algoritmaları ile geliştirilen modellere dahil edilmiştir.

Tablo 1. Örnek Veri

İlaç ID	524215363
İlaç	Vitamin X
Dönem	2023-01
Adet	742
Fiyat	125,2
Pazar (Adet)	83.436

Tablo 1. Devamı

Pazar (TL)	10.241.531
USD	18,8080
Aylık TÜFE	6,65
Özel Günler	1
Okul Dönemi	1
Sınav (ÖSYM, MEB)	0
Kampanya	35236

Veri Hazırlığı ve Değişken Türetilmesi

Veride var olan değişkenlerin yanında model açıklayıcılığına katkıda bulunabilecek değişkenler de çalışmada türetilmiştir. Yapılan satışların 4 aylık gecikmeli değişkenleri, hareketli ortalamaları, bir önceki döneme göre yüzdesel değişimi, pazar adet ve TL hacminin bir önceki döneme göre yüzdesel değişimi, ürün fiyatının bir önceki döneme göre yüzdesel değişimi çıkartılarak bu değişkenlerdeki dönemlik değişimin modellere pozitif bir etkisinin olabileceği düşünülmüştür.

Modelleme ve Parametre Optimizasyonu

Model geliştirme aşamasında Prophet, LSTM, XGBoost, LightGBM ve CatBoost olmak üzere 5 farklı algoritma kullanılmıştır. Prophet ve LSTM algoritmaları bütün ürünler için kullanılırken Boosting algoritmaları, yalnızca tahminlemesi diğer ürünlere göre daha zor olan vitamin ürünleri için kullanılmıştır.

Eğitim verisinin büyüklüğüne 12, 24 ve 36 ay olmak üzere kontrollü şekilde karar verilmiştir. Bu sayede pandemi gibi satış verilerini saptıran dönemler model performansını olumsuz etkileyebileceği için dışarda tutulmuştur.

Modelleri geliştirirken, algoritmaların parametrelerini optimize etmek genellikle modelin başarısını artırır. Bu nedenle, XGBoost, LightGBM ve CatBoost algoritmalarıyla oluşturulan modellerde hiperparametre optimizasyonu için Optuna paketi kullanılmıştır.

Model Performansının Ölçülmesi ve İdeal Modele Karar Verilmesi

Modellerin performansları Tablo 2’de paylaşılan Ortalama Mutlak Hata Yüzdesi (MAPE), Ortalama Mutlak Hata (MAE) ve Kök Ortalama Kare Hatası (RMSE) hata metrikleri üzerinden incelenmiştir.

Tablo 2. Hata Metrikleri

Ortalama Mutlak Hata Yüzdesi MAPE (Mean Absolute Percentage Error)	$MAPE = \frac{1}{n} \sum_{i=1}^n \left \frac{\hat{y}_i - y_i}{y_i} \right \times 100$
Ortalama Mutlak Hata MAE (Mean Absolute Error)	$MAE = \frac{1}{n} \sum_{i=1}^n \hat{y}_i - y_i $
Kök Ortalama Kare Hatası RMSE (Root Mean Squared Error)	$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - y_i)^2}{n}}$

İdeal model (Champion Model) seçilirken her ay yeni verilerle modeller güncellenir ve her ürün için en başarılı model otomatik olarak seçilir. Bu süreçte, her ay modeller güncellendikten sonra, her bir ürün için son 3 aylık MAPE değerleri karşılaştırılır. Ürün bazında en düşük MAPE değerine sahip modelin güncel dönemdeki tahminleri nihai tahmin olarak belirlenir. Bu makalede ideal model sonuçları 2023 Kasım, 2023 Aralık ve 2024 Ocak olmak üzere 3 farklı dönemdeki performanslara göre raporlanmıştır.

Sistem Geliştirmesi

Bu çalışmada, verilerin toplanmasından model tahminlerinin arayüze aktarılmasına kadar tüm sürecin otomatik işlediği bir akış tasarlanmıştır. Sistem, her ay düzenli çalışabilmesi için bir sunucu üzerinde canlıya alınmıştır. Yeni ürün veya modellerin kolayca eklenip çıkarılabilmesi için jenerik bir yapıda geliştirilmiştir.

Tüm kodlar Python ile yazılmış, veriler Microsoft SQL Server’da depolanmış ve tahmin sonuçları İş Zekası çözümü MicroStrategy üzerinden sunulmuştur. Model tahminlerinin iş süreçlerinde Tedarik Zinciri Bölümü tarafından doğrudan kullanılabilmesi ve Ürün Müdürleri tarafından kolaylıkla takip edilebilmesi için her ay güncel tahminleri SAP sistemine ileten bir yapı oluşturulmuştur.

Çalışma sonunda ortaya konulan sistem, her ay yeni gelen verilerle modelleri güncelleyip tahminleri revize ederken, ürün bazında en ideal modeli seçerek en yüksek başarı oranını sağlamayı hedeflemektedir. Sonuç olarak, yaşayan ve kendini sürekli güncelleyen bir sistem tasarlanmıştır.

SONUÇLAR

Çalışmada ilk olarak tüm ürünler için Prophet ve LSTM algoritmaları ile modeller geliştirilmiştir. Tablo 3’te model sonlarındaki eklerin açıklamaları bulunmaktadır.

Tablo 3. Türetilmiş Model Yapıları

-LOG	Satış değişkeni üzerinde logaritmik dönüşüm uygulanarak türetilen değişkenin kullanıldığı model.
-LAG	Yalnızca gecikme yöntemiyle türetilen değişkenlerin kullanıldığı model.
-LAGROLL	Gecikmeli değişkenlerle beraber hareketli ortalama yöntemiyle türetilen değişkenlerin de kullanıldığı model.
-EXT	Tüm harici ve türetilen değişkenlerin kullanıldığı model.

Yapılan testler sonucunda ideal eğitim setinin büyüklüğünün modeller için farklılaştığı tespit edilmiştir. Test sonuçlarına göre çalışmanın devamında Prophet modeli için 24 aylık ve diğer 4 model için 36 aylık eğitim seti kullanılmıştır.

Vitamin ürünü olmayan 108 ürün için PROPHET, PROPHET-LOG ve LSTM olmak üzere 3 farklı model geliştirilmiştir. 2024 yılının Ocak ayına kadar olan verilerle modeller eğitilmiştir ve Şubat, Mart, Nisan ve Mayıs ayları için tahminler üretilmiştir.

Tablo 4’te her bir modelin Mart ayı için kaç tane ürünü hangi MAPE aralığında tahmin ettiği bilgisi görülmektedir. Örneğin; PROPHET modeli 25 ürünü en fazla %10 sapmayla tahmin ederken 19 ürünü %100 hatalıya yakın şekilde tahmin etmiştir.

Tablo 4. 2024 Mart MAPE Aralıklarına Göre Ürün Frekansları

Model	MAPE (%)									
	0-10	11-20	21-30	31-40	41-50	51-60	61-70	71-80	81-90	91-100
PROPHET	25	23	11	11	6	4	5	1	3	19
PROPHET-LOG	24	16	15	12	5	8	2	3	2	21
LSTM	26	17	13	11	14	5	3	0	2	17

Tablo 5’te de her bir modelin aylık bazda ortalama MAPE ve MAE değerleri yer almaktadır. Genel değerlere bakıldığında logaritmik dönüşümlere uğramış değişkenlerle eğitilen PROPHET-LOG modelinin hem MAPE hem de MAE değerleri açısından diğer iki modelin arkasında kaldığı görülmektedir. Bu tabloya göre tahmin edilecek dönem uzaklaştıkça modellerin başarısının da düştüğü görülmektedir.

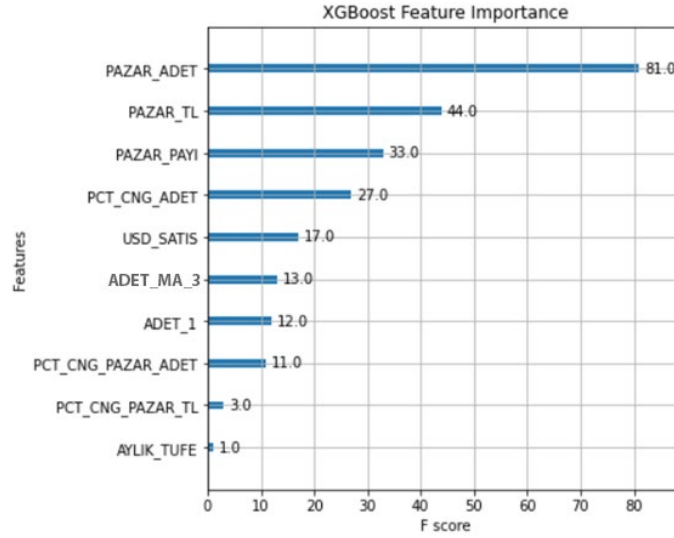
Tablo 5. Vitamin Grubu Hariç Model Performansları

Model	Şubat		Mart		Nisan		Mayıs	
	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE
PROPHET	0.3539	8823.5556	0.3862	9658.3981	0.433	7573.5556	0.4051	8103.1019
PROPHET-LOG	0.3747	10290.0741	0.41	12206.8981	0.4274	8494.2778	0.4416	12271.9722
LSTM	0.333	9385.9074	0.372	8717.2963	0.4322	7352	0.3908	8893.8241
IDEAL	0.3477	9740.8426	0.3917	10174.3426	0.4182	6669.5278	0.4012	9114.1204

Tablo 5’teki sonuçlara göre, ideal modelin seçilmesinin sistemin genel başarısını az da olsa olumlu yönde etkilediği görülmektedir. İdeal model seçiminin temel amacı, dönemsel olarak başarısı düşen modeller yerine daha iyi uyum sağlayan modelleri önermektir.

Çalışmanın devamında, tahmin başarısı düşük vitamin ürünleri için harici veriler kullanılarak farklı değişken kombinasyonlarıyla modeller geliştirilmiştir. Boosting algoritmalarıyla oluşturulan modellerin Değişken Önem

(Feature Importance) değerleri incelenmiştir. Şekil 2’de XGBoost ile oluşturulan bir modelin değişken önem grafiği sunulmuştur.



Şekil 2. Vitamin Grubu Değişken Önemi (Feature Importance)

Bu grafikte, pazar verilerinin ürün satışları üzerinde diğer değişkenlere göre daha etkili olduğu görülmektedir. Diğer ürünler için de benzer grafikler incelenerek hangi harici değişkenlerin gerekli olduğu değerlendirilmiştir. Sonuç olarak, bazı ürünler için harici verilerin etkisiz olduğu, gecikme veya hareketli ortalama yöntemleriyle türetilen değişkenlerin daha etkili olduğu tespit edilmiştir.

Tablo 6’da 18 farklı vitamin ürününe ait ortalama hata değerleri yer almaktadır. MAPE değerleri incelendiğinde XGBOOST-EXT modeli diğer Boosting modellerinden daha iyi performans sergilediği görülmektedir. Bununla birlikte MAE değerine göre sonuçları karşılaştırdığımızda bu model diğerlerinden daha düşük performans sergilemektedir. Bu durum XGBOOST-EXT modelinin ani iniş ve çıkışlarda daha iyi uyum sağladığını fakat sayısal olarak daha yüksek hatalar yaptığını göstermektedir. Ayrıca bu tabloya göre harici verilerle eğitilen modellerin, vitamin ürünlerinin genel başarısında çok etkili olmadığı, hatta bazı modellerin herhangi bir harici değişken kullanmayan Prophet modellerinden daha kötü performansa sahip olduğu da görülmektedir.

Testlerin sonunda Tablo 5’te vitamin dışı ürünlerinin IDEAL model performansının Tablo 6’da paylaşılan vitamin ürünlerinin IDEAL model performansından daha başarılı sonuç verdiği gözlemlenmiştir.

Tablo 6. Vitamin Grubu Model Performansları

Model	Şubat		Mart		Nisan		Mayıs	
	MAPE	MAE	MAPE	MAE	MAPE	MAE	MAPE	MAE
PROPHET	0.4927	559.2778	0.5167	1184.3333	0.5601	929.7778	0.8461	1273.1667
PROPHET-LOG	0.4352	608.1667	0.4715	1192.1111	0.6013	794.8333	0.7963	1015.1111
LSTM	0.4562	523.1667	0.5343	1176.8889	0.6208	787.5556	0.823	928.0556
XGBOOST-LAG	0.5694	795.6111	0.5864	1338	0.6626	937.5	0.5302	663.1111
XGBOOST-LAGROLL	0.5833	825.0556	0.6139	1439.5556	0.6102	1075.5	0.7269	1047.8889
XGBOOST-EXT	0.4672	1519.2222	0.5152	2145	0.4782	1533.3333	0.7507	2063.9444
LIGHTGBM-LAG	0.6377	1020.9444	0.5293	1401.7778	0.5301	1298.8333	0.7522	1163.7222
LIGHTGBM-LAGROLL	0.6056	1045.0556	0.5143	1356.6111	0.5351	1170.1667	0.818	1225.1111
LIGHTGBM-EXT	0.5177	1421.3333	0.4836	2068.0556	0.5491	1607.1667	0.8705	2171.2222
CATBOOST-LAG	0.5711	995.7778	0.5601	1369.6111	0.5437	1255.8889	0.7111	1154.6111
CATBOOST-LAGROLL	0.5607	883.8333	0.543	1189.6111	0.5367	1169.5	0.8507	1349.8889
CATBOOST-EXT	0.5355	1087	0.5531	1730.6667	0.5722	1300.1667	0.8589	1916.5556
IDEAL	0.5506	744.5	0.5056	1195	0.5642	864.9444	0.807	1108.8333

TARTIŞMA VE SONUÇ

Yapılan çalışmalar sonucunda eğitim seti büyüklüğünü doğru seçmenin model performansına etki ettiği görülmüştür. İncelenen benzer çalışmalarda kampanya gibi harici verilerin talep tahmininde doğruluk oranını artırabileceği belirtilmekteydi (Tolosşi-Halacheva vd., 2024). Bu bilgiye istinaden incelenen bir diğer çalışmada da kampanya gibi harici veriler ile XGBoost modeli oluşturularak iyi sonuçlar elde edildiği görülmüştür (İmece ve Beyca, 2022). Fakat yapılan bu çalışmada toplanan ve türetilen harici değişkenlerin, tahminlemesi zor olan vitamin ürünlerini tahmin etmede yeterli seviyede etkili olmadığı görülmüştür. Bu ürünlerdeki ani iniş ve çıkışları daha iyi açıklayabilecek değişkenlerin araştırılması ve denenmesiyle model başarıları daha iyi seviyeye getirilebilir.

Çalışma kapsamında bir RNN (Recurrent Neural Networks) yöntemi olan LSTM ile bazı ürünler özelinde daha iyi çıktılar üretildiği gözlenmiştir. Literatür taramasında benzer bir çalışmanın DNN (Deep Neural Networks) ile yapıldığı ve başarılı sonuçlar elde edildiği görülmüştür (Chang vd., 2017). Bu çalışmada geliştirilen LSTM modelleri üzerinde parametre optimizasyonu yapılarak daha iyi sonuçlar elde edilebilir.

İdeal modele karar veren yapının, sistemin başarısının stabil kalmasına büyük katkı sağladığı görülmektedir. Fakat bu yapı ürün bazında daima en iyi modeli seçememektedir. İdeal modeli seçen yapı daha da geliştirilerek sistemin genel başarısını artırmak mümkün olabilir.

Canlıya alınan bu sistem ile her dönem her ürün için seçilen ideal model ile tüm modellerin hata metrikleri veri tabanına kaydedilmektedir. Toplanan bu veriler ile her dönem hangi modellerin daha iyi veya daha kötü olduğu analiz edilerek yeni içgörüler elde edilebilir ve bu içgörüler ile sistem başarısını artıracak aksiyon planları oluşturulabilir.

Referanslar

- Atasever M, Bağcı H, 2021. Türkiye İlaç Sektörü Analizi, 2. Bölüm: İlaç Sektörünün Değerlendirilmesi, Nobel Yayın, pp. 53-81
- Chang O, Naranjo I, Guerron C, Criollo D, 2017. A Deep Learning Algorithm to Forecast Sales of Pharmaceutical Products, Future Technologies Conference (FTC)
- İmece S, Beyca Ö, 2022. Demand Forecasting with Integration of Time Series and Regression Models in Pharmaceutical Industry, International Journal of Advances in Engineering and Pure Sciences. 34(3): 415-425
- Merkuryeva G, Valberga A, 2019. Demand forecasting in pharmaceutical supply chains: A case study, Procedia Computer Science, 149: 3-10
- Pall R, Gauthier Y, Auer S, Mowaswes W, 2023. Predicting drug shortages using pharmacy data and machine learning, Health Care Management Science, 26: 395-411
- Rathipriya R, Abdul A, Dhamodharavadhani S, Meero A, Yoganandan G, 2023. Demand forecasting model for time-series pharmaceutical data using shallow and deep neural network model, Neural Computing and Applications, 35(2): 1945-1957
- Tolosşi-Halacheva L. -M, Andreev R, Shcherbakov O, Nevo E, 2024. Challenges in Modeling Drug Shortage Events in the Pharmaceutical Domain, 2024 IEEE 40th International Conference on Data Engineering Workshops (ICDEW), pp. 140-143

Comparative Analysis of LLM-Based and Conventional NLP-Based Methods for MultiClass Text Classification in Turkish

Furkan Ayık^{1*}, Mehmet Yalçın², Celal Akçelik³ ve Serap Mungan Tanhan⁴

^{1,4}Turkish Techology, Istanbul, Turkiye

²Statistics, Yıldız Technical University, Istanbul, Turkiye

³RDC Partner, Istanbul, Turkiye

Abstract

The increasing availability of data sources has made text-based information extraction essential. Text classification, a key task in this field, enables high-accuracy categorization through NLP models, reducing time and costs for faster responses. While text classification is gaining importance in Turkish, its morphological complexity has limited research compared to English. This study developed multi-class text classification models using Turkish cabin report data from an airline, comparing machine learning-based boosting models and large language models (LLMs) (e.g., LLaMA, BERT Cased, BERT Uncased, ConvBERT Cased, ConvBERT Uncased, XLM-RoBERTa). Traditional text representations (FastText, Word2Vec, TF-IDF) were used in ML models, while LLMs were fine-tuned.

The results showed that the Bert Base Uncased - Turkish model achieved the highest accuracy at 93%, demonstrating LLMs' effectiveness in capturing Turkish linguistic nuances. Among boosting models, CatBoost with N-gram TF/IDF reached 86% accuracy. This study highlights LLMs' potential in Turkish text classification, confirming the validity of traditional NLP methods, especially in resource-limited environments.

Key words: NLP, Large Language Model, Text Classification, Turkish Language, BERT

INTRODUCTION

Extracting information from text data has gained increasing importance with the growth in data sources. Text classification is a foundational task in this area, enabling efficient and accurate categorization through natural language processing (NLP) models, saving both time and cost. Although methods for text classification have advanced, their applications in the Turkish language remain limited due to its complex morphological structure, which presents unique challenges compared to languages like English.

This study aims to address these challenges by developing multi-class text classification models using a four-class dataset based on Turkish cabin reports from an airline. Both traditional machine learning (ML) methods and large language models (LLMs), particularly the Attention-based Large Language Model (LLaMA) and Bidirectional Encoder Representations from Transformers (BERT), were evaluated for performance. For ML models, text representation techniques such as TF-IDF, FastText, and Word2Vec were employed, while BERT and its variants were fine-tuned, and LLaMA 3.1 was applied using zero-shot learning through prompt engineering.

Our analysis provides a comprehensive examination of conventional ML and advanced transfer learning models within NLP, focusing on preprocessing and word embedding techniques. By leveraging fine-tuning and zero-shot methods, we assess the effectiveness of these models in addressing the Turkish text classification problem, aiming for substantial efficiency improvements in terms of labor and time.

MATERIAL AND METHODS

This study utilized both machine learning (ML)-based boosting models and large language models (LLMs) for the text classification task.

For traditional NLP models based on machine learning, key text representation techniques such as FastText, Word2Vec, and TF-IDF were applied. Among the boosting models, CatBoost, LightGBM, and XGBoost were chosen. The modeling and data preprocessing steps were executed with the help of Python's Sklearn library and related packages.

In the LLM-based models, fine-tuning was applied to achieve optimal performance. Models used included LLaMA, BERT Cased, BERT Uncased, ConvBERT Cased, ConvBERT Uncased, and XLMRoBERTa. These models were downloaded from Hugging Face and run in GPU-supported PyTorch environments.

Material

Natural language processing (NLP) involves processing and interpreting spoken language as used in human communication, translating it into a form that computers can process and analyze using mathematical formulas (Adalı, 2012). NLP methods are essential for a range of applications, such as virtual assistants, text summarization, text classification, speech-to-text conversion, and language learning (Kurtuluş, 2023).

Within the scope of this study, NLP methods were applied for text classification, which aims to predict categories corresponding to predefined classes or labels in a classification system. Various text classification methods have been developed in the literature, mostly based on statistical, machine learning, and deep learning approaches. In this study, machine learning was applied as a conventional NLP method alongside transfer learning models within the field of deep learning, where a substantial amount of recent research has been conducted.

MACHINE LEARNING

Machine learning is a scientific field focused on developing algorithms, models, and techniques that enable computer systems to learn in a way similar to humans (Alpaydm, 2010). In this study, we used supervised learning methods within machine learning to establish prediction models for multi-class classification. Boosting models, known for delivering high performance in recent years, were preferred for the classification task. Boosting is an ensemble learning technique that improves predictions generated by a decision tree (Coşkun & Turanlı, 2023).

In boosting decision tree methods, each tree is sequentially constructed using information from previous trees to reduce errors in the predictions of the current tree (Giussani, 2019). The specific boosting methods used in this study include XgBoost, CatBoost, and LightGBM.

LARGE LANGUAGE MODELS BERT (BIDIRECTIONAL ENCODER REPRESENTATIONS FROM TRANSFORMERS)

BERT (Bidirectional Encoder Representations from Transformers) is a transformer model designed to capture contextual relationships in text by analyzing the words bidirectionally—from both left-to-right and right-to-left (Devlin et al., 2018). After pre-training on a large dataset, BERT can be fine-tuned for specific tasks such as multi-class classification. BERT converts words into embedding vectors, adds positional information, and passes the output for the [CLS] token through a fully connected layer to predict class labels. For languages like Turkish, with rich morphology, models fine-tuned for Turkish, such as BERTBase-Cased-Turkish, offer improved performance by capturing complex linguistic features.

ConvBERT enhances BERT's capabilities by incorporating convolutional layers, which capture local dependencies more effectively (Jiang et al., 2020), while XLM-RoBERTa leverages cross-lingual learning, showing high performance in various languages, including Turkish (Conneau et al., 2019).

Overall procedures for BERT's pre-training and fine-tuning in Figure 1. The same architectures are utilized in both phases, with differences only in the output layers.

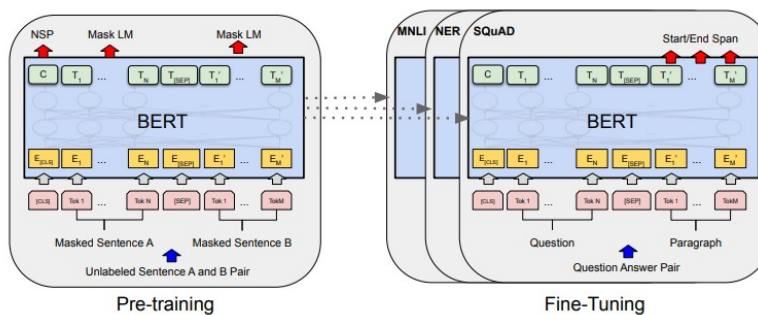


Figure 1 Bert's Pre-training and Fine-tuning

LAMAA 3.1

LLAMA 3.1 is an attention-based large language model recognized for its high performance in natural language processing (NLP) tasks. With extensive training on large datasets, LLAMA 3.1 has demonstrated strong performance in tasks such as text understanding, summarization, and translation (Dubey et al., 2024). In this study, LLAMA 3.1

was used in zero-shot mode for text classification, where specific prompts were designed to direct the model. For example, instructions like “Classify this text according to the given categories” help the model interpret the classification task. Model parameters such as `max_tokens: 512`, `top_p: 0.95`, `temperature: 0.01`, `presence_penalty: 1.03`, and `streaming: False` were optimized for enhanced performance (LLamaaimodel, 2024; Github, 2024).

EVALUATION METRICS

In evaluating classification models, Confusion Matrix is commonly used (Han et al., 2012). However, since this study addresses a multi-class classification problem with an imbalanced target variable, macro and weighted metrics were also considered. F1-Macro calculates the average of F1 scores across all classes, providing an overall performance measure that accounts for class imbalance, while F1-Weighted also measures F1 for each class but weights it by the number of instances in each class to reflect general performance. Precision-Macro represents the average precision, calculated as the ratio of true positive classifications to all positive predictions across classes, whereas Precision-Weighted adjusts these values based on class instance count. Similarly, Recall-Macro measures the average true positive rate across classes, and Recall-Weighted weights each recall score by instance count to represent overall performance (Kenneth Leung, 2024; Scikit-learn, 2024).

Methods

The Collection of the Data

In this study, a dataset of Turkish cabin reports containing free-text fields was classified using machine learning and large language models, allowing for performance comparison. The dataset consists of 9,463 entries recorded by cabin managers and contains three columns: Date, text, and target variable. The target variable has four classes: A, B, C, and D. The dataset covers the period from July 2023 to July 2024, with a time-based split where July 2023 to February 2024 serves as the training set (7,413 entries), and March to July 2024 serves as the test set (2,050 entries), representing 25% of the data. The distribution of the target variable is imbalanced, with class frequencies of A: 2717, B: 4772, C: 397, and D: 1577.

Before classification, preprocessing steps were applied, including converting text to lowercase, removing special characters, punctuation, numbers, and URLs. For machine learning models, additional steps were taken, such as converting Turkish characters to their English equivalents and removing stop words.

Statistical Analysis

MACHINE LEARNING-BASED NLP METHODS

In this study, machine learning-based natural language processing (NLP) methods were applied for multiclass text classification. After cleaning characters and numbers in the text, stemming and lemmatization were used to reduce words to their root forms, creating unique units (Kurtuluş, 2023). Turkish Stemmer library was used for stemming, while Spacy library was used for lemmatization.

Following preprocessing, words were represented numerically through word embedding or feature extraction, with methods falling into two categories: frequency-based and prediction-based (Polat & Körpe, 2018). Frequency-based representation methods like TF-IDF (at word and n-gram levels) and predictionbased methods like Word2Vec (Skip-gram and CBOW) and FastText were used, with parameters set to 400 for vector size and 10 for window size. For a comprehensive evaluation, the ML approach included two word normalization methods (stemming and lemmatization), five vectorization techniques (TF-IDF, Ngram TF-IDF, Word2Vec, FastText), and three decision tree boosting models (CatBoost, XGBoost, and LightGBM), yielding 30 different results to determine the most effective approach.

BERT

In the application and experiment phase of our study, several BERT-based models were utilized for Turkish multi-class text classification. Developed by the MDZ Digital Library team (dbmdz) at the Bavarian State Library (MDZ Digital Library, 2024), which open sources models specifically designed for various languages, the bert-base-turkish-128k-cased and uncased models are optimized for Turkish syntax and morphology, making them well-suited to capture the language’s linguistic nuances. Additionally, XLMRoBERTa, a multilingual model, leverages cross-lingual pretraining to provide robust generalization for Turkish. We also used ConvBERT-Turkish models (cased

and uncased), which incorporate convolutional layers to better capture local dependencies in the text. To assess the impact of case sensitivity, we compared the performance of cased and uncased versions of both BERT and ConvBERT models. For fair comparison, all models were trained with the same parameters using the Hugging Face Trainer API's

“TrainerArguments,” with learning_rate: 0.005, weight_decay: 0.01, num_train_epochs: 10, and eval_strategy: “epoch” (Hugging Face, 2024).

The results indicate that the bert-base-turkish-128k-cased model achieved the highest performance according to the F1-Macro metric, making it the best model for our classification task. The bert-baseturkish-128k-uncased version followed closely, ranking as the second-best model based on F1-Macro.

LLAMA ZERO-SHOT

In this study, we evaluated the zero-shot multi-class text classification performance of LLaMA 3.1, an attention-based large language model known for its high accuracy in natural language processing (NLP) tasks. Due to extensive pre-training on large datasets, LLaMA 3.1 demonstrates strong capabilities in tasks such as text understanding, summarization, and translation (Dubey et al., 2024). For zero-shot classification, prompts were carefully crafted to specify the task, ensuring the model’s responses were limited to class labels only. For instance, instructions such as "Classify this text according to the given categories" helped guide the model effectively in interpreting the classification task. Model parameters were optimized to improve performance, with settings as follows: max_tokens: 512, top_p: 0.95, temperature: 0.01, presence_penalty: 1.03, and streaming: False (LLamaaimodel, 2024; Github, 2024).

RESULTS

The classification models were ranked based on F1 Macro scores, as shown in Table 1. The fine-tuned **bertbase-turkish-128k-cased** model emerged as the top performer with an accuracy of 93% and an F1 Macro score of 89%. This model was followed by bert-base-turkish-128k-uncased and DistilBERT – Turkish Cased models, demonstrating the effectiveness of LLMs in capturing Turkish language nuances. The zeroshot classification performance of the LLaMA 3.1 model reached an accuracy of 81%, showing promising results for zero-shot applications. Among the boosting models, CatBoost achieved an accuracy of 86% when paired with N-gram TF/IDF vectorization. These findings validate the efficiency of traditional NLP methods, especially in resource-constrained environments, while highlighting the superior performance of LLMs in Turkish text classification.

Table 1 Model Results

Methods	Model and vectorizer Methods	Accuracy	F1 Weighted	F1 Macro	Precision Weighted	Precision Macro	Recall Weighted	Recall Macro
LLM	dbmdz/bert-base-turkish-128k-cased_multiclass	0,92	0,92	0,89	0,92	0,91	0,92	0,88
LLM	dbmdz/bert-base-turkish-128k-uncased_multiclass	0,91	0,91	0,87	0,92	0,91	0,91	0,87
LLM	dbmdz/distilbert-base-turkish-cased_multiclass	0,87	0,82	0,83	0,85	0,85	0,87	0,84
LLM	xlm-roberta-base_multiclass	0,87	0,87	0,82	0,87	0,83	0,87	0,81
LLM	dbmdz/convert-base-turkish-mc4-uncased_multiclass	0,90	0,90	0,87	0,90	0,84	0,90	0,84
LLM	dbmdz/convert-base-turkish-mc4-cased_multiclass	0,91	0,91	0,87	0,87	0,87	0,91	0,87
ML - Boosting	CatBoost - ngram TF/IDF	0,85	0,85	0,80	0,87	0,82	0,85	0,79
ML - Boosting	LightGBM - ngram TF/IDF	0,84	0,84	0,78	0,85	0,82	0,84	0,76
ML - Boosting	CatBoost - word TF/IDF	0,84	0,84	0,78	0,86	0,81	0,84	0,77
ML - Boosting	LightGBM - word TF/IDF	0,83	0,83	0,77	0,84	0,80	0,83	0,75
ML - Boosting	CatBoost - Cbow	0,79	0,78	0,71	0,79	0,77	0,79	0,68
Zero Shot	Llama 3.1 (8b)	0,82	0,82	0,76	0,82	0,73	0,82	0,81

DISCUSSION AND CONCLUSION

With the increase in data sources, the significance of extracting information from text data has grown, particularly in Turkish text classification. However, due to the complex morphological structure of Turkish, research in this area has been more limited compared to English.

In our study, we developed multi-class text classification models using Turkish cabin report data from an airline, covering the period from July 2023 to July 2024. Both machine learning models and large language models (LLMs) were evaluated using fine-tuning and zero-shot learning methods. Among the models, bert-base-turkish-128k-uncased achieved the highest accuracy at 93%, showcasing the superior performance of LLM-based approaches in capturing Turkish linguistic nuances. The zero-shot classification accuracy of LLaMA 3.1 reached 81%, while the CatBoost model with FastText vectorization achieved 86% accuracy.

This analysis demonstrates that automating text classification tasks can lead to significant savings in time and labor. It also highlights the effectiveness of both machine learning and large language models in Turkish text classification, confirming the potential of traditional NLP methods, particularly in resourcelimited environments.

References

- Alpaydın, E., 2010. Introduction to Machine Learning. USA: Massachusetts Institute of Technology Press
- Coşkun, S. B., & Turanlı, M. (2023). Credit risk analysis using boosting methods. *J. Appl. Math. Stat. Inform.*, 19(1), 5–18. <https://doi.org/10.2478/jamsi-2023-0001>.
- Dubey A., Jauhri A., et al.,2024, The Llama 3 Herd of Models, <https://arxiv.org/abs/2407.21783>, <https://doi.org/10.48550/arXiv.2407.21783>
- GitHub. (2024). *LLaMA 3.1 Model Card*. Retrieved from https://github.com/meta-llama/llama-models/blob/main/models/llama3_1/MODEL_CARD.md (Accessed: 30.09.2024).
- Kenneth Leung,2024, Micro, Macro & Weighted Averages of F1 Score, Clearly Explained (readmedium.com),<https://readmedium.com/micro-macro-weighted-averages-of-f1-score-clearlyexplained-b603420b292f>
- Kurtuluş T., 2023. Analysis Of Customer Feedback Replied To An Airlines Support Account Via Twitter With Natural Language Processing Methods, Master Thesis. Yıldız Technical University, Graduate School Of Science And Engineering, Istanbul, Turkiye.
- LLamaaimodel,2024, Llama 3.1 Requirements(What you Need to Use It), <https://llamaaimodel.com/requirements/>, Date of Access:25.09.2024
- Polat, H., Körpe, M. 2018. TBMM Genel Kurul Tutanaklarından Yakın Anlamlı Kavramların Çıkarılması. *International Journal of Informatics Technologies*, 11, 3.
- Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. arXiv preprint arXiv:1810.04805.
- Conneau, A., Khandelwal, K., Goyal, N., Chaudhary, V., Wenzek, G., Guzmán, F., ... & Stoyanov, V. (2020). Unsupervised Cross-lingual Representation Learning at Scale. arXiv preprint arXiv:1911.02116 (for XLM-RoBERTa).
- Jiang, Z., Xu, F. F., Araki, J., & Neubig, G. (2020). ConvBERT: Improving BERT with Span-based Dynamic Convolution. arXiv preprint arXiv:2008.02496.
- Hugging Face. (2024). Text classification with Hugging Face Transformers. Retrieved from https://huggingface.co/docs/transformers/en/tasks/sequence_classification
- MDZ Digital Library. (2024). dbmdz/bert-base-turkish-128k-cased. Hugging Face. Retrieved from <https://huggingface.co/dbmdz/bert-base-turkish-128k-cased>

An Overall Equipment Efficiency Predictive Analysis of a Hydraulic Press System by Time Series Forecasting with Topological Features

Korkut Anapa^{1*}, İsmail Güzel², and Ceylan Talu Yozgatgil³

^{1,2}Institute of Applied Mathematics, Middle East Technical University, Ankara, Turkey

²Network Technologies Department, TÜBİTAK ULAKBİM, Ankara, Turkey

³Department of Statistics, Middle East Technical University, Ankara, Turkey

Abstract

This research is about forecasting Overall Equipment Efficiency (OEE) in hydraulic press systems critical to production plants. These systems represent significant capital and operational expenditures, making high OEE essential for economic sustainability in manufacturing. The study focuses on short-term, highly volatile OEE data, reflecting frequent process changes within production plants that directly impact OEE. Traditional time series forecasting methods are enhanced by incorporating topological features to address the challenges posed by the short-term nature of the data, where conventional machine learning and artificial intelligence approaches fall short. This study proposes a novel methodological enhancement for more accurate OEE forecasts by integrating topological data analysis. It demonstrates how these advanced analytical techniques can optimize predictive maintenance strategies, thereby contributing to manufacturing operations' overall efficiency and stability.

Keywords: Overall Equipment Efficiency, Forecasting, Time Series, Topological Data Analysis

INTRODUCTION

Forecasting is essential across industries, providing insights that drive decision-making and optimize operations. In manufacturing, accurate predictions help plan production schedules, manage inventory, and utilize machinery efficiently. Traditional models like Autoregressive Integrated Moving Average (ARIMA) have been foundational, using historical data to identify patterns and project future values. Additionally, the rise of machine learning (ML) and artificial intelligence (AI) has significantly improved forecasting accuracy. These advanced techniques, such as neural networks and ensemble methods, capture complex, non-linear relationships in data, making them ideal for real-time applications.

The persistent homology technique, which is a tool from Topological Data Analysis (TDA), has provided novel methods and insights for investigating the shape and topological properties of complex data sets (Munch, 2017). TDA focuses on the shape and structure of data, helping identify patterns and anomalies in highly variable datasets. This method provides deeper insights into data dynamics, improving forecast reliability in complex environments.

In manufacturing, especially with hydraulic press systems, forecasting Overall Equipment Efficiency (OEE) is crucial. OEE measures the efficiency and productivity of equipment, and accurate forecasting helps maximize utilization, reduce downtime, and optimize maintenance. With the Fourth Industrial Revolution (4IR) emphasizing smart manufacturing and data-driven decision-making, advanced forecasting techniques are vital. They help manufacturers achieve higher OEE levels, ensuring cost-effectiveness and competitiveness in the industrial landscape.

METHODOLOGY

Literature Survey

Forecasting OEE in manufacturing industries has earned significant attention due to its implications for efficiency and productivity. Various studies have explored methods and approaches for predicting OEE values, leveraging machine learning techniques and historical data analysis.

The article by Fatima et al. (2024) reviews time-series forecasting models used in industrial settings, highlighting the strengths and limitations of traditional statistical methods like ARIMA and Exponential Smoothing, as well as the flexibility offered by machine learning techniques such as Support Vector Machines and Artificial Neural Networks. The paper emphasizes the effectiveness of hybrid models and suggests selecting models based on specific industry needs, data characteristics, and forecasting objectives while noting the increasing importance of AI and ML in handling complex data. Saylam and Atlı's study (2023), focuses on the use of advanced models like XGBoost, Prophet, MLP, and LSTM to predict production line downtime, impacting OEE. The research highlights feature selection, model robustness, and the use of cross-validation and early-stopping mechanisms to enhance predictive

accuracy. Dobra and Jósvali (2022) utilize decision trees to forecast the effects of production type changes on OEE, stressing the importance of understanding operational adjustments for better resource planning and efficiency. Finally, the article by Anusha and Umasankar (2020) explores the use of statistical methods like simple moving averages and Holt's double exponential smoothing to forecast OEE, demonstrating their applicability for real-time performance monitoring.

The studies reviewed highlight a promising trend toward combining advanced analytical techniques with traditional methods for forecasting OEE. This integration can significantly improve predictive accuracy and operational efficiency in manufacturing. Machine learning is not just a tool for better predictions; it also offers manufacturers a proactive approach to optimize their operations. In forecasting concepts, feature engineering is also an important step to get more improved models as well as choosing ML models. A novel method can be using TDA due to its ability to capture complex relationships and patterns within the raw data that may not be easily discernible through traditional methods. The combination of TDA and ML models in the forecasting contexts has improved model performance and better decision-making for various industries (Senekane et al., 2021). Claudia Caudai and colleagues (2024) present a Topological Machine Learning (TML) pipeline that integrates persistent homology with machine learning, offering a robust framework for capturing topological invariants and differentiating spaces in various data types. This pipeline is particularly successful in medical applications like cancer grading and Alzheimer's disease classification. In Zheng et al., 2024, they propose a TDA-based method for analyzing multivariate time series, specifically EEG signals, using the Hilbert–Huang Transform to capture dynamic frequency and amplitude information. This method excels in capturing temporal and spatial information, outperforming other TDA approaches in classifying EEG signals (Zheng et al., 2024). Additionally, Oseko et al. (2024) explore TDA's application in financial forecasting, utilizing persistence homology to identify critical topological features as early warning signals for financial crises. The study highlights TDA's potential to complement existing risk assessment frameworks.

While TDA has been successfully applied in many fields (Chazal et al., 2021) like image analysis, material science, and biology, its application in forecasting OEE is still in the early stages. In this study, we examine and compare the combination of traditional methods like SARIMA, TDA, and ML methods.

Data

The dataset that we examine contains 40,320 entries with operational data logged every minute, indicating whether hydraulic presses were active ('1') or inactive ('0'). For analysis, this data is aggregated into 60-minute intervals, yielding 672 observations that reflect the number of active minutes per hour, with a range of 1 to 60. The mean operational value is 26.25, with a high standard deviation of 23.29, indicating significant variation. The skewness is -0.0046, suggesting a symmetric distribution, while the kurtosis of -1.777 indicates a flatter distribution compared to normal. The Shapiro-Wilk test confirms the data does not follow a normal distribution.

In Figure 1, “original data” graph displays the hourly Operational Equipment Efficiency (OEE) values from March 9, 2024, to April 5, 2024. The OEE values fluctuate significantly throughout the period, indicating variations in operational performance. Figure 1. also presents a time series decomposition of operational data in the same period by trend, seasonality, and residuals. The decomposition plot reveals that the time series data exhibits both daily and weekly seasonality, as evidenced by the distinct repeating patterns in the corresponding components. The trend component shows a volatile pattern with an initial decline followed by a slight increase over time.

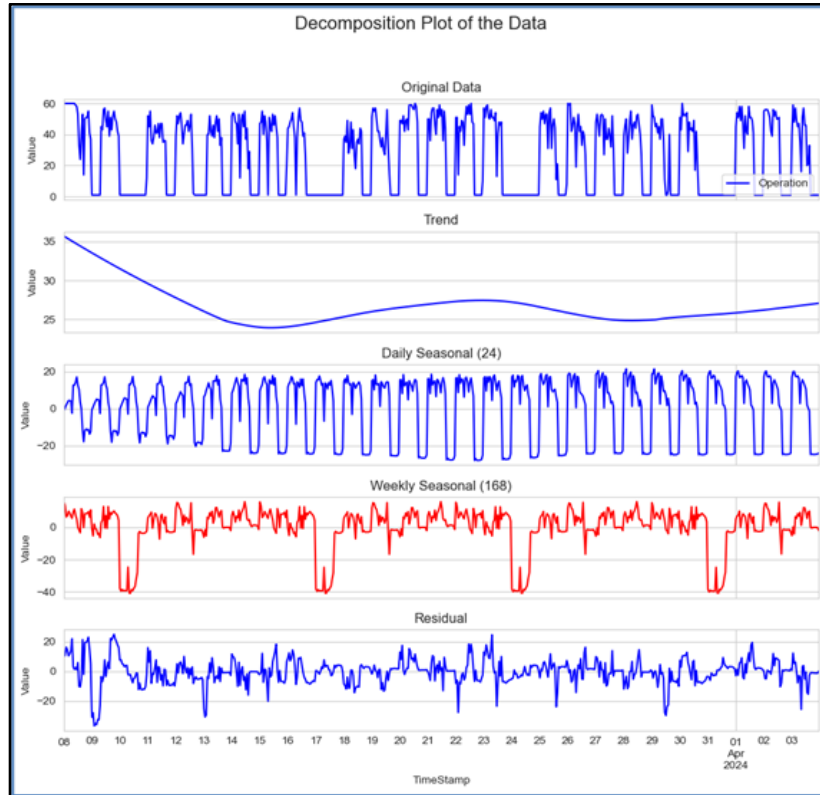


Figure 1. Time series decomposition plot of the OEE values

Figure 2 displays the ACF and PACF plots of the series. Significant peaks at lags 24, 48, 72, and 96 in the ACF suggest strong daily seasonality. Significant spikes in the ACF at lag 1 and decaying pattern suggest MA (1) or higher order. Significant spikes in the PACF at lag 1, decaying quickly, suggest AR (1) or higher order.

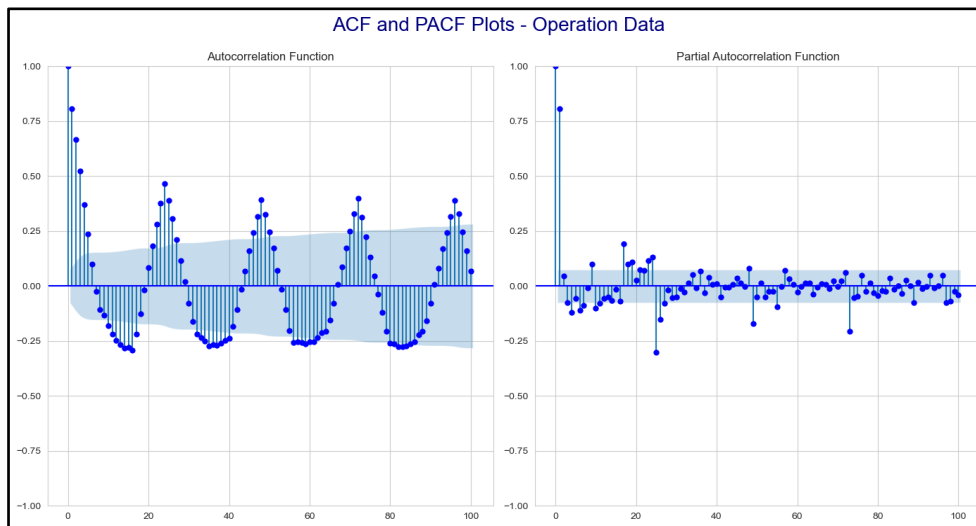


Figure 2. ACF and PACF plots of the hourly OEE series.

Methods

The aim is to forecast the OEE of a hydraulic press system, employing a range of forecasting models. These models are categorized into three distinct groups, each utilizing different methodologies and techniques. The effectiveness of these categories is then evaluated to identify the most proficient forecasting model.

Modeling with Traditional Time Series Methods (Model 1)

This category primarily focuses on traditional time series forecasting methods, with an emphasis on the Seasonal ARIMA (SARIMA) model. We explore several variations of the SARIMA model. The models are Standard SARIMA, SARIMA with Fourier Transformations, and SARIMA with Dual Seasonality. In addition to these models,

this category includes feature extraction techniques applied to the time series data. ML algorithms are then employed in regression models to utilize these features effectively, with feature selection algorithms playing a crucial role in optimizing the input features to enhance model performance. 16 features are extracted, which are: Absolute Energy, Mean Absolute Change, Mean, Standard Deviation, Variation Coefficient, Skewness, Kurtosis, Root Mean Square, Absolute Sum of Changes, C3 lag 1, C3 lag 2, Quantile 0.1, Quantile 0.9, Autocorrelation lag 1, Autocorrelation lag 2, and Autocorrelation lag 3.

Modeling with Topological Data Analysis (Model 2)

This category introduces TDA features to the forecasting framework. The process begins with the application of sliding window techniques to the time series data to generate samples. These samples are then used to extract topological features via a persistent diagram that summarizes persistent homology for the homology dimension 0 and 1 encodes connected components and holes simultaneously, focusing on capturing the underlying data structure that traditional methods might overlook. More details about persistent homology and time series tasks can be found in the related studies (Wu and Hargreaves, 2021; Ravishanker and Renjie, 2021).

Topological Feature Extraction: The method involves identifying shapes and connectivity patterns within the data that are invariant under certain transformations, providing a robust basis for the subsequent analysis. The persistent diagrams spaces, the output of persistent homology, are not suitable for use in ML algorithms because of the lack of inner product in those spaces. To deal with this issue, there are many methods to vectorize the persistent diagram (Chung, and Austin, 2022). Here, the extracted topological feature include Persistence Entropy, measuring the complexity and diversity of topological features; Bottleneck and Wasserstein Amplitudes, which quantify dissimilarities between persistence diagrams; Betti Curves, which track the number of topological features across filtration values; Persistence Landscapes, which transform diagrams into continuous, functional landscapes; Silhouette, a simplified summary of the topological structure; and Heat Kernel, a method that emphasizes long-lasting features by generating dense summaries of persistence diagrams. These transformations enable the effective use of topological features in various models.

Feature Selection and ML/AI Regression: Like the first category, this phase involves selecting the most informative topological features and using them in ML and AI algorithms for effective regression modeling.

Modeling with Hybrid SARIMA-Topological Approach (Model 3)

The third category explores a hybrid approach that combines the strengths of SARIMA and TDA. Here, the topological features extracted from the TDA methods are used as exogenous inputs in the SARIMA model. The integration aims to enhance the forecasting capabilities of the SARIMA model by incorporating complex patterns and structures detected by TDA, thus potentially leading to significant improvements in forecast accuracy.

Each category is carefully analyzed to determine its effectiveness in forecasting OEE values. The results are compared to establish which method offers the most significant improvement over traditional forecasting techniques, thereby guiding the selection of the optimal approach for real-world applications in hydraulic press systems.

Evaluation Methods

This study uses a dataset of 28 days of OEE values from a hydraulic press system to compare different forecasting methods. The first 27 days are used for training, while the final day, characterized by high volatility, serves as the test set. The effectiveness of the models is measured using Mean Absolute Percentage Error (MAPE). While overall performance across the test day is important, special attention is given to the first five values of the forecast, as these initial predictions often set the tone for the model's accuracy and responsiveness. Achieving an MAE of 5 and a MAPE of less than 10% in these early forecasts is considered a good balance between accuracy and responsiveness. This focus ensures that the chosen forecasting method can effectively handle the volatility in OEE values, which is critical for reliable real-time monitoring and decision-making in manufacturing environments.

RESULTS

In this study, the first approach to forecasting the OEE of a hydraulic press system involves traditional time series analysis methods. The results are tabulated in Table 1.

Among the traditional models, the SARIMAX (0, 1, 1) (2, 1, 0, 24) model, identified via the Auto-ARIMA process, performed best with an MAE of 6.35 and a MAPE of 36. The model accurately captured the dataset's dynamics and responded well to variations in OEE values. The accompanying graph, Figure 3. shows the actual versus forecasted OEE values, illustrating the model's predictive accuracy. Focusing on the **first five samples** of each model's output,

the SARIMAX model achieved an MAE of 3.6 and a MAPE of 8%, which is within the acceptable range for immediate operational adjustments.

Table 1. The results of traditional models (24 test sizes).

Model	MAE	MAPE
SARIMAX (0, 1, 1) (2, 1, 0, 24)	6.35	36
SARIMAX (1, 0, 1) (1, 0, 1, 24)	9.95	40
Holt-Winters (24)	10.65	65
SARIMAX w Fourier (0, 1, 1) (2, 1, 0, 24)	6.42	37
SARIMAX (1, 0, 1) (1, 0, 1,24,168)	8.15	57

Defined features of Model 1 are extracted, and before using them in regression models, feature selection is conducted to enhance the model's robustness and accuracy, minimizing redundancy and optimizing feature utility, thus preventing overfitting and ensuring good generalization to new data. After feature selection, various ML and deep learning (DL) algorithms are applied to perform regressions on the dataset, including RFR, SVR, GBR, XGBoost, LightGBM, LSTM, GRU, and TCN. Models are tested with both the full feature set and a selected subset, generally showing better performance with selected features due to reduced noise and avoidance of overfitting by hyperparameter tuning. However, despite optimization efforts, the models struggled to accurately forecast OEE values, indicating potential issues with feature engineering or model capabilities. XGBoost achieved the best results with an MAE of 13.11 and a MAPE of 280.

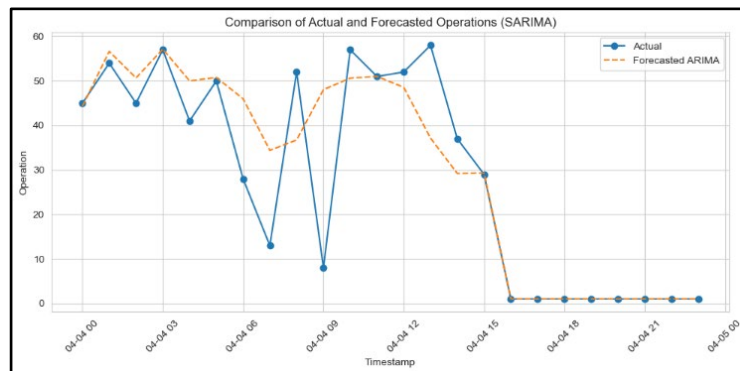


Figure 3. Forecast – actual OEE values plot for SARIMAX (0, 1, 1) (2, 1, 0, 24) model.

In the second part of the study, topological features derived from time series data are explored, initially extracting 3,179 features, later reduced to 24 through rigorous selection to address the High Dimensional Low Sample Size (HDLSS) problem. The XGBoost Regressor, after hyperparameter tuning, using these features, reports an MAE of 21.71 and a MAPE of 694, indicating predictive challenges.

In the final phase, topological features are integrated into a SARIMAX model as exogenous variables. This approach, using features like Amplitude (Landscape), Betti Curve, and others, improved the SARIMAX model's performance, reducing the **MAE from 6.35 to 6.20** and the **MAPE from 36 to 33** as shown in Figure 4. The accuracy of the **initial five forecasts** also improved, with the **MAPE decreasing from 8 to 6**, demonstrating the effectiveness of incorporating TDA features into forecasting models.

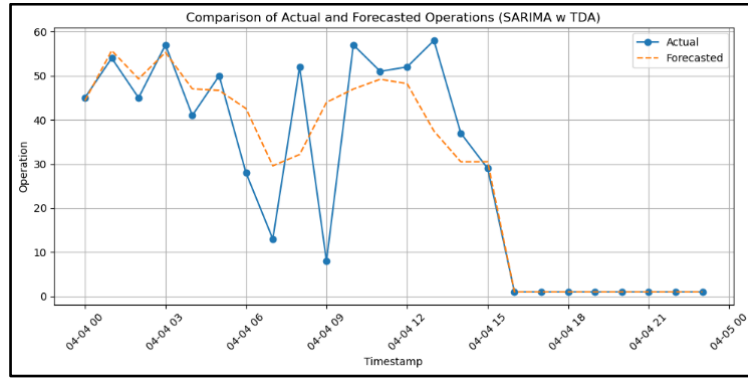


Figure 4. The plot of the best-achieved forecast for 24 samples with SARIMAX with topological features.

This research phase demonstrates that integrating topological features into traditional time series forecasting models can significantly enhance their predictive capabilities. By capturing the underlying data structure that traditional models might overlook, TDA provides a robust framework for improving model accuracy and reliability, confirming its utility in complex and highly volatile forecasting scenarios such as predicting OEE values.

DISCUSSION AND CONCLUSION

In conclusion, integrating topological features with traditional forecasting methods has shown significant potential in improving model accuracy for volatile scenarios like OEE prediction. This study demonstrates that combining traditional time series analysis with TDA creates a hybrid model capable of handling complex, short-term data sequences. Traditional methods alone often struggle with rapid data fluctuations, but TDA captures underlying structures, enhancing forecasting in environments with limited and volatile data. The hybrid approach also addresses the challenges of machine learning algorithms requiring large datasets, providing a more accurate and responsive solution for short-term predictions. This methodology's success in OEE forecasting suggests its broader applicability in various fields where traditional methods and standalone algorithms may fall short, highlighting the value of interdisciplinary approaches in predictive analytics.

References

- Munch, Elizabeth. (2017). A user's guide to topological data analysis. *Journal of Learning Analytics* 4.2 (2017): 47-61.
- Fatima, Syeda & Rahimi, Afshin. (2024). A Review of Time-Series Forecasting Algorithms for Industrial Manufacturing Systems. *Machines*. 12. 380. 10.3390/machines12060380.
- Saylam, A., & Atlı, H. (2023). Predictive Analytics for Production Line Downtime: A Comprehensive Study Using Advanced Machine Learning Models. *The European Journal of Research and Development*, 3(4), 88-94.
- Dobra, P., & Jósvali, J. (2022). Assembly Line Overall Equipment Effectiveness (OEE) Prediction from Human Estimation to Supervised Machine Learning. *J. Manuf. Mater. Process.*, 6(59).
- Anusha, Chintada, & Umasankar, Venkata. (2020). Performance Prediction through OEE-Model. *International Journal of Industrial Engineering and Management*, 11, 93-104.
- M. Senekane, N. J. Matjelo and B. M. Taele. 2021. Improving Short-term Output Power Forecasting Using Topological Data Analysis and Machine Learning, 2021 International Conference on Electrical, Computer and Energy Technologies (ICECET), Cape Town, South Africa, 2021, pp. 1-6.
- Chazal, Frédéric, and Bertrand Michel. 2021. An introduction to topological data analysis: fundamental and practical aspects for data scientists. *Frontiers in artificial intelligence* 4 (2021): 667963.
- Caudai, Claudia, Colantonio, Sara, Conti, Francesco, D'acunto, Mario, Moroni, Davide, & Pascali, Maria. (2024). Exploring Data through the Topological Lens: A Topological Machine Learning Pipeline for Data Analysis.
- Zheng, Jingyi, Feng, Ziqin, & Ekstrom, Arne .D. (2024). Towards Analysis of Multivariate Time Series Using Topological Data Analysis. *Mathematics*, 12(1727).

Oseko, Naomi, Omondi, Achuo, Onyango, Hassan, Olwa, Desma, Maina, Gabriel, Morara, Moses, & Thiong'o, Kelvin. (2024). Forecasting Financial Crisis using Topological Data Analysis Approach. *African Scientific Annual Review*, 1, 1-17.

Wu, Chengyuan, and Carol Anne Hargreaves. 2021. Topological Machine Learning for Multivariate Time Series. *Journal of Experimental & Theoretical Artificial Intelligence* 34 (2): 311–26.

Ravishanker, Nalini, and Renjie Chen. 2021. An introduction to persistent homology for time series. *Wiley Interdisciplinary Reviews: Computational Statistics* 13.3 (2021): e1548.

Chung, Yu-Min, and Austin Lawson. 2022. Persistence curves: A canonical framework for summarizing persistence diagrams. *Advances in Computational Mathematics* 48.1 (2022): 6.

Sigortacılıkta Araç Sigortası Talep Sınıflandırmasında Gelişmiş Makine ve Derin Öğrenme Yöntemleri: Yenilikçi Yaklaşımlar ve Performans Değerlendirmeleri

Volkan OBAN¹, Bulut Karadağ¹, Gökhan Görmüş¹, Aslıhan Çandır¹

¹ Neova Sigorta A.Ş., İstanbul, Türkiye

Özet

Sigortacılık sektöründe, müşteri taleplerinin ve hasar iddialarının doğru bir şekilde sınıflandırılması, maliyet optimizasyonu, müşteri memnuniyetinin artırılması ve etkin risk yönetiminin sağlanması açısından kritik bir öneme sahiptir. Sigorta şirketleri, bu amaçlara ulaşmak için gelişmiş makine öğrenmesi ve derin öğrenme yöntemlerine giderek daha fazla ihtiyaç duymaktadır. Bu çalışmada, araç sigortası taleplerine ilişkin veriler kullanılarak, çeşitli makine öğrenmesi ve derin öğrenme modellerinin sınıflandırma performansları kapsamlı bir şekilde karşılaştırılmıştır.

Araştırmamızda, TabNet, Graph Neural Network (GNN), Graph Attention Network (GAT), GraphSAGE, LSTM, BiLSTM, Random Forest, XGBoost, LightGBM, CatBoost ve TensorFlow Decision Forests gibi modern ve çeşitli modellerin performans değerlendirilmesi yapılmıştır. Bu modeller, farklı algoritmik yaklaşımlar ve veri işleme stratejileri kullanarak sigorta taleplerinin yüksek doğruluk oranları ile sınıflandırılmasını sağlamaktadır. Çalışma, bu modellerin doğruluk oranlarını ve genel performanslarını değerlendirerek, sigortacılık sektöründeki sınıflandırma problemlerine yönelik yenilikçi yaklaşımların etkinliğini ve potansiyelini ortaya koymaktadır.

Elde edilen bulgular, farklı makine öğrenmesi ve derin öğrenme modellerinin sigorta taleplerinin sınıflandırılması açısından anlamlı ölçüde farklılık gösterdiğini ve bu modellerin doğru seçiminin sigorta şirketlerinin operasyonel verimliliğini ve müşteri memnuniyetini artırmada hayati bir rol oynayabileceğini vurgulamaktadır. Sonuç olarak, bu çalışma, sigorta sektöründe sınıflandırma performansını optimize etmek için hangi modellerin daha uygun olduğuna dair önemli içgörüler sunmaktadır.

Anahtar Kelimeler: Araç sigortası, Makine öğrenmesi, Derin öğrenme, Sınıflandırma, Sigorta talebi, Performans değerlendirme, TabNet, Graph Neural Network (GNN), XGBoost, LightGBM, CatBoost, LSTM, Random Forest, BiLSTM, TensorFlow Decision Forests, Veri analizi.

I. Giriş

Sigorta şirketleri, müşteri taleplerini ve hasar iddialarını değerlendirirken yüksek doğrulukta sınıflandırma yapma zorunluluğundadır. Sigorta taleplerinin doğru bir şekilde sınıflandırılması, yalnızca finansal kayıpların en aza indirilmesiyle sınırlı kalmayıp, aynı zamanda müşteri memnuniyeti ve güvenini artırma açısından da büyük bir öneme sahiptir. Doğru ve zamanında yapılan sınıflandırma, sigorta şirketlerinin etkin risk yönetimi gerçekleştirmesine ve böylece operasyonel süreçlerin daha verimli bir şekilde yürütülmesine olanak tanır. Günümüzde, bu ihtiyacın karşılanması amacıyla gelişmiş makine öğrenmesi ve derin öğrenme yöntemlerine olan talep artmaktadır. Bu bağlamda, çalışmada araç sigortası talep verileri kullanılarak çeşitli makine öğrenmesi ve derin öğrenme modellerinin performansı incelenmiştir. Araştırma, sektördeki sınıflandırma problemlerine yönelik hem geleneksel hem de ileri düzey yaklaşımların etkinliğini değerlendirmeyi amaçlamaktadır.

Literatürde, sigorta taleplerinin sınıflandırılması üzerine yapılan araştırmalar, bu süreçlerin daha verimli bir şekilde yürütülmesi için çeşitli makine öğrenmesi ve derin öğrenme tekniklerinin kullanılabilirliğini ortaya koymaktadır. Örneğin, XGBoost, Random Forest ve LightGBM gibi gradyan artırma modelleri, hızlı ve yüksek performansları ile dikkat çekmekte ve sigorta taleplerinin doğrulukla sınıflandırılmasında önemli bir rol oynamaktadır. Bunun yanı sıra, LSTM ve GRU gibi zaman serisi analizine odaklanan derin öğrenme modelleri, özellikle geçmiş verilerin önemli olduğu senaryolarda yüksek performans göstermektedir. Yapılan istatistiksel analizler, doğru model seçiminin sigorta taleplerinin sınıflandırılmasında %15'e varan performans artışları sağlayabileceğini göstermektedir. Bu nedenle, sigorta şirketleri için model çeşitliliği ve uygun model seçimi, operasyonel verimlilik ve müşteri memnuniyeti açısından kritik bir faktör olarak öne çıkmaktadır.

Önümüzdeki yıllarda, özellikle otonom ve yarı otonom araçların sigorta piyasasına giriş yapmasıyla birlikte, sigorta taleplerinin yapısı ve karmaşıklığı da önemli ölçüde değişecektir. Otonom araçların karıştığı kazaların artması, sigorta şirketlerinin daha karmaşık veri setleriyle ve yeni tip taleplerle başa çıkmasını gerektirecektir. Bu gelişmeler, sigorta taleplerinin sınıflandırılmasında ileri düzey makine öğrenmesi ve derin öğrenme modellerinin kullanımını daha da kritik hale getirmektedir. Örneğin, Graph Neural Networks (GNN) gibi grafik tabanlı modeller, otonom

araçların yolculuk verilerini ve araçlar arasındaki etkileşimleri modelleyerek sigorta taleplerinin değerlendirilmesinde önemli bir rol oynayabilir. Bu durum, sigorta sektöründe kullanılan sınıflandırma algoritmalarının ve veri işleme tekniklerinin sürekli olarak geliştirilmesini ve güncellenmesini zorunlu kılmaktadır.

Sonuç olarak, bu çalışmada çeşitli makine öğrenmesi ve derin öğrenme modellerinin performansı değerlendirilmiş ve sigorta taleplerinin sınıflandırılmasında en uygun modellerin belirlenmesine yönelik öneriler sunulmuştur. Gelecek çalışmalarda, otonom ve yarı otonom araçlardan elde edilecek yeni veri tiplerinin kullanılması ve bu verilere dayalı modellerin geliştirilmesi planlanmaktadır. Bu kapsamda, sigorta şirketlerinin risk yönetim stratejilerini optimize etmelerine yardımcı olacak daha yenilikçi ve etkili sınıflandırma yöntemlerinin geliştirilmesi hedeflenmektedir.

II. Veri Seti

Veri Seti Tanıtımı: Çalışmamızda kullanılan veri seti, araç sigortası taleplerine ilişkin çeşitli demografik ve araç bilgilerini içermektedir. Veri seti, [Kaggle](#) platformundan alınmıştır ve araç sigortası talep verilerini içermektedir, zenginleştirilmiştir. Veri seti, 10,000 satır ve çeşitli özellikleri temsil eden 23 sütundan oluşmaktadır.

Veri setindeki ana sütunlar şunlardır:

ID: Benzersiz kimlik numarası

AGE: Sürücünün yaş grubu

GENDER: Sürücünün cinsiyeti

RACE: Sürücünün ırkı

DRIVING_EXPERIENCE: Sürüş deneyimi (yıl)

EDUCATION: Eğitim seviyesi

INCOME: Gelir seviyesi

CREDIT_SCORE: Kredi skoru

VEHICLE_OWNERSHIP: Araç sahipliği (0 = Sahip değil, 1 = Sahip)

VEHICLE_YEAR: Araç yılı

ANNUAL_MILEAGE: Yıllık kilometre

VEHICLE_TYPE: Araç tipi (sedan, suv, vb.)

SPEEDING_VIOLATIONS: Hız ihlali sayısı

DUIS: Alkol ile araç kullanma durumu

PAST_ACCIDENTS: Geçmiş kaza sayısı

OUTCOME: Sigorta talebinin sonucu (0 = Onaylanmadı, 1 = Onaylandı)

Özellik Mühendisliği ve Veri Hazırlama: Özellik mühendisliği sürecinde, veri setine yeni değişkenler eklenmiştir:

VEHICLE_AGE: Aracın yaşı (yıl cinsinden).

VEHICLE_MILEAGE: Aracın toplam kilometresi.

PREVIOUS_DAMAGE: Aracın önceki hasar durumu (0 = Hasar yok, 1 = Hasar var).

DRIVER_PENALTY_POINTS: Sürücünün ceza puanları.

Özellik Mühendisliği Detayları:

Araç Yaşı (VEHICLE_AGE): Araçların yaşları, araçların sigorta talebi ile olan ilişkisini belirlemek için önemli bir faktördür. Bu özellik, aracın üretim yılı ve veri setinin toplandığı yıl arasındaki fark olarak hesaplanmıştır.

Araç Kilometresi (VEHICLE_MILEAGE): Araçların ne kadar kullanıldığını gösteren bu özellik, sigorta taleplerinin değerlendirilmesinde kritik bir role sahiptir. Daha yüksek kilometre, aracın daha fazla risk altında olabileceğini gösterebilir.

Önceki Hasar Durumu (PREVIOUS_DAMAGE): Aracın daha önce hasar görüp görmediği bilgisi, gelecekteki sigorta taleplerinin öngörülmesinde önemlidir. Bu özellik, önceki hasar durumunu 0 veya 1 olarak kodlar.

Sürücü Ceza Puanları (DRIVER_PENALTY_POINTS): Sürücünün geçmişte aldığı ceza puanları, sürücünün risk profili hakkında bilgi verir. Daha yüksek ceza puanları, daha yüksek risk anlamına gelebilir.

Veri Hazırlama:

Eksik Değerlerin Doldurulması: Veri setindeki eksik değerler, ilgili *sütunların ortalamaları* ile doldurulmuştur.

Kategorik Değişkenlerin Etiketlenmesi: Kategorik değişkenler, sayısal değerlere dönüştürülmüştür. Örneğin, cinsiyet (GENDER) sütunu kadın (0) ve erkek (1) olarak kodlanmıştır.

Sayısal Özelliklerin Standardize Edilmesi: Sayısal değişkenler, model performansını artırmak için standardize edilmiştir. Bu işlem, her bir değişkenin ortalamasının 0 ve standart sapmasının 1 olacak şekilde dönüştürülmesini içerir.

Veri Analizi: Veri analizi aşamasında, veri setindeki demografik ve araç bilgileri incelenmiş, bu özelliklerin sigorta talebi sonucu üzerindeki etkisi değerlendirilmiştir. İstatistiksel analizler ve görselleştirmeler kullanılarak, her bir özelliğin dağılımı ve sınıflandırma probleminin genel yapısı ortaya konmuştur.

Yaş Gruplarına Göre Sigorta Talebi Sonucu

Aşağıda, farklı yaş gruplarına göre sigorta talebi sonucu (OUTCOME) 0 ve 1 olanların sayısı verilmiştir:

YAŞ ARALIĞI	SONUÇ: 0	SONUÇ:1
16-25	568	1448
26-39	2031	1032
40-64	2474	457
65+	1794	196

TABLO 1: Yaş Gruplarına Göre Sigorta Talebi Sonucu

Ceza Puanı	SONUÇ: 0	SONUÇ: 1
0	583	236
1	533	261
2	619	241
3	574	286
4	615	289
5	552	253
6	558	275
7	572	253
8	582	244
9	533	256
10	566	241
11	580	298

TABLO 2: Ceza Puanına Göre Talep Sonuçları

Gelir Seviyesine Göre Sigorta Talebi Sonucu

Aşağıda, farklı gelir seviyelerine göre sigorta talebi sonucu (OUTCOME) 0 ve 1 olanların sayısı verilmiştir:

Gelir Seviyesi	SONUÇ: 0	SONUÇ: 1
Orta Sınıf (Middle Class)	1546	592
Yoksulluk (Poverty)	628	1186
Üst Sınıf (Upper Class)	3757	579
Çalışan Sınıf (Working Class)	936	776

TABLO 3: Gelir Seviyesine Göre Talep Sonuçları



Şekil 1 - Evli Olma Durumuna Göre Sigorta Talebi Sonucu

Evli Değil:

Sigorta talebi onaylanmayan (OUTCOME = 0) kişi sayısı 2838.

Sigorta talebi onaylanan (OUTCOME = 1) kişi sayısı 2180.

Evli:

Sigorta talebi onaylanmayan (OUTCOME = 0) kişi sayısı 4029.

Sigorta talebi onaylanan (OUTCOME = 1) kişi sayısı 953.

Çocuk Varlığına Göre Sigorta Talebi Sonucu

Olmayanlar:

Sigorta talebi onaylanmayan (OUTCOME = 0) kişi sayısı 1637.

Sigorta talebi onaylanan (OUTCOME = 1) kişi sayısı 1475.

Olanlar:

Sigorta talebi onaylanmayan (OUTCOME = 0) kişi sayısı 5230.

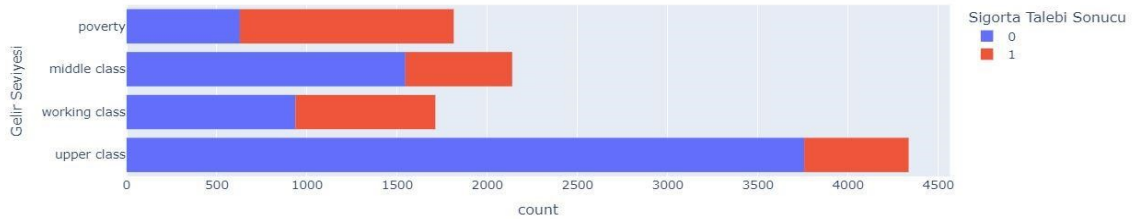
Sigorta talebi onaylanan (OUTCOME = 1) kişi sayısı 1658.

Geçmiş Kazalar ve Sigorta Talebi Sonucu

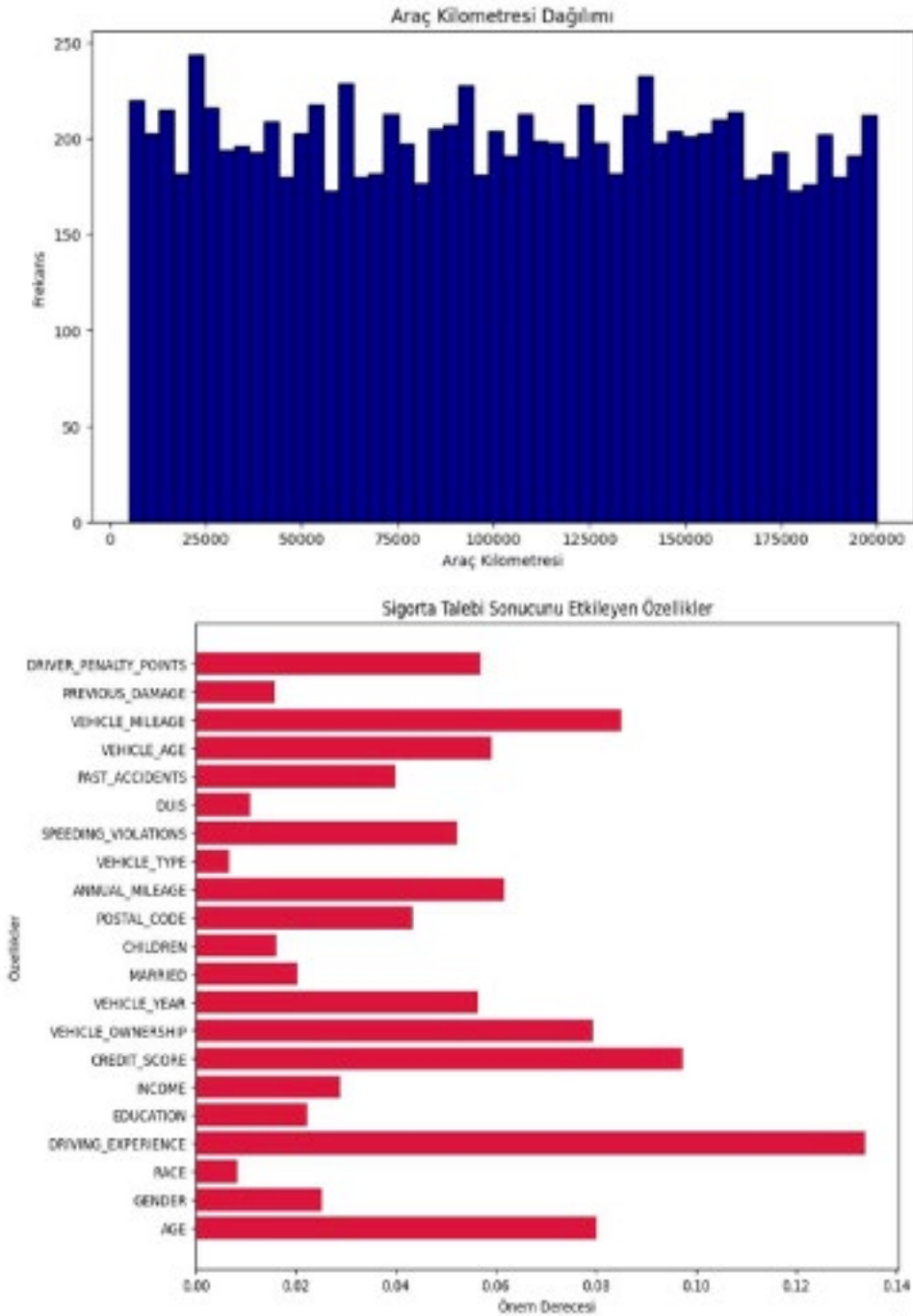


Şekil 2 – Geçmiş Kazalar ve Sigorta Talebi Sonucu

Gelir Seviyesine Göre Sigorta Talebi Sonucu Dağılımı



Şekil 3 – Gelir Seviyesine Göre Sigorta Talebi Sonucu Dağılımı



Şekil 4 – Araç Kilometresi ve Sigorta Talebi Sonucunu Etkileyen Özellikler

III. Model Çalışmaları:

Model Eğitim ve Değerlendirme: Çalışmamızda kullanılan modellerin açıklamaları ve test doğruluk oranları aşağıda sunulmuştur:

TabNetClassifier: Derin öğrenme tabanlı, tabular veriler için özel olarak tasarlanmış bir model. TabNet, dikkat mekanizmalarını kullanarak, tabular verilerdeki önemli özellikleri otomatik olarak seçer ve işler.

Graph Neural Network (GNN): Grafik verileri işlemek için kullanılan sinir ağı modeli. GNN, düğümler arasındaki ilişkileri ve bu ilişkilerin düğüm özellikleri üzerindeki etkilerini öğrenir.

Graph Attention Network (GAT): Grafik verilerdeki düğümler arasındaki ilişkileri öğrenmek için dikkat mekanizmalarını kullanan model. GAT, düğümler arasındaki önemli bağlantıları vurgulamak için dikkat katsayılarını kullanır.

GraphSAGE: Büyük grafiklerdeki düğüm özelliklerini öğrenmek için kullanılan bir model. GraphSAGE, düğüm komşuluk bilgilerini kullanarak düğüm temsillerini öğrenir.

LSTM (Long Short-Term Memory): Zaman serisi verilerini işlemek için kullanılan derin öğrenme modeli. LSTM, uzun süreli bağımlılıkları öğrenme kapasitesine sahiptir.

Random Forest: Birden fazla karar ağacının çoğunluk oyu ile karar verdiği ansambl öğrenme modeli. Random Forest, overfitting riskini azaltarak genel performansı artırır.

XGBoost: Ekstrem gradyan artırma algoritması, hızlı ve yüksek performanslı bir model. XGBoost, modelin hatalarını iteratif olarak düzelterek yüksek doğruluk sağlar.

LightGBM: Işık gradyan artırma makinesi, büyük veri setleri için optimize edilmiş bir model. LightGBM, hızı ve performansı ile bilinir.

CatBoost: Kategorik özellikleri otomatik olarak işleyebilen gradyan artırma algoritması. CatBoost, kategorik verilerde yüksek performans gösterir.

Decision Forest: TensorFlow Decision Forests (TF-DF), TensorFlow içerisinde karar ormanı modellerini (örneğin, Random Forests, Gradient Boosted Trees) eğitmek, çalıştırmak ve yorumlamak için kullanılan bir kütüphanedir. TF-DF, sınıflandırma, regresyon, sıralama ve uplift modellerini destekler. Bu model, çeşitli karar ağaçlarının birleşiminden oluşur ve her bir ağacın tahminlerini birleştirerek daha doğru ve genelleştirilebilir sonuçlar elde eder.

BiLSTM (Bidirectional LSTM): Çift yönlü LSTM modeli. BiLSTM, verinin hem geçmiş hem de gelecek bağlamını öğrenir.

Feature Bagging + Neural Networks: Özellik torbalama ve sinir ağları kombinasyonu. Bu yöntem, çeşitli özellik torbaları üzerinde sinir ağları eğiterek performansı artırır.

Wide and Deep Learning Model: Geniş ve derin öğrenme modeli. Bu model, hem geniş lineer modellerin hem de derin sinir ağlarının gücünü birleştirir.

DNN (Deep Neural Network): Derin sinir ağı. DNN, çok katmanlı yapısıyla karmaşık verileri öğrenme kapasitesine sahiptir.

VAE (Variational AutoEncoder) with TabNet: Değişken otoenkoder ve TabNet kombinasyonu. VAE, veri dağılımını öğrenirken TabNet, önemli özellikleri seçer.

Self-Attention Mechanisms with Transformer: Dönüştürücü modellerdeki kendine dikkat mekanizmaları. Bu mekanizmalar, verinin önemli kısımlarına odaklanarak öğrenmeyi optimize eder.

Meta-Embedding Neural Networks: Meta-gömme sinir ağları. Bu ağlar, çeşitli gömme tekniklerini birleştirerek daha zengin veri temsilleri öğrenir.

SENet (Squeeze-and-Excitation Networks) with Dense Layers: Sıkıştırma ve uyarma ağları ile yoğun katmanlar. SENet, özellik haritalarının önemini yeniden ölçeklendirmek için sıkıştırma ve uyarma mekanizmalarını kullanarak, verinin önemli kısımlarını vurgular.

CVAE (Conditional Variational Autoencoder) for Classification: Şartlı değişken otoenkoder. CVAE, belirli koşullar altında veriyi yeniden yapılandırarak sınıflandırma problemlerine uygulanabilir.

GANs (Generative Adversarial Networks) for Data Augmentation and Classification: Veri artırma ve sınıflandırma için generatif adversarial ağlar. GANs, sahte veri üretip sınıflandırıcıyı eğiterek performansı artırır.

Hybrid CNN-RNN with GBDT: Hibrit konvolüsyonel sinir ağı ve tekrarlayan sinir ağı ile gradyan artırma karar ağaçları. Bu model, CNN ve RNN'in özellik çıkarma yeteneklerini GBDT'nin güçlü sınıflandırma yetenekleriyle birleştirir.

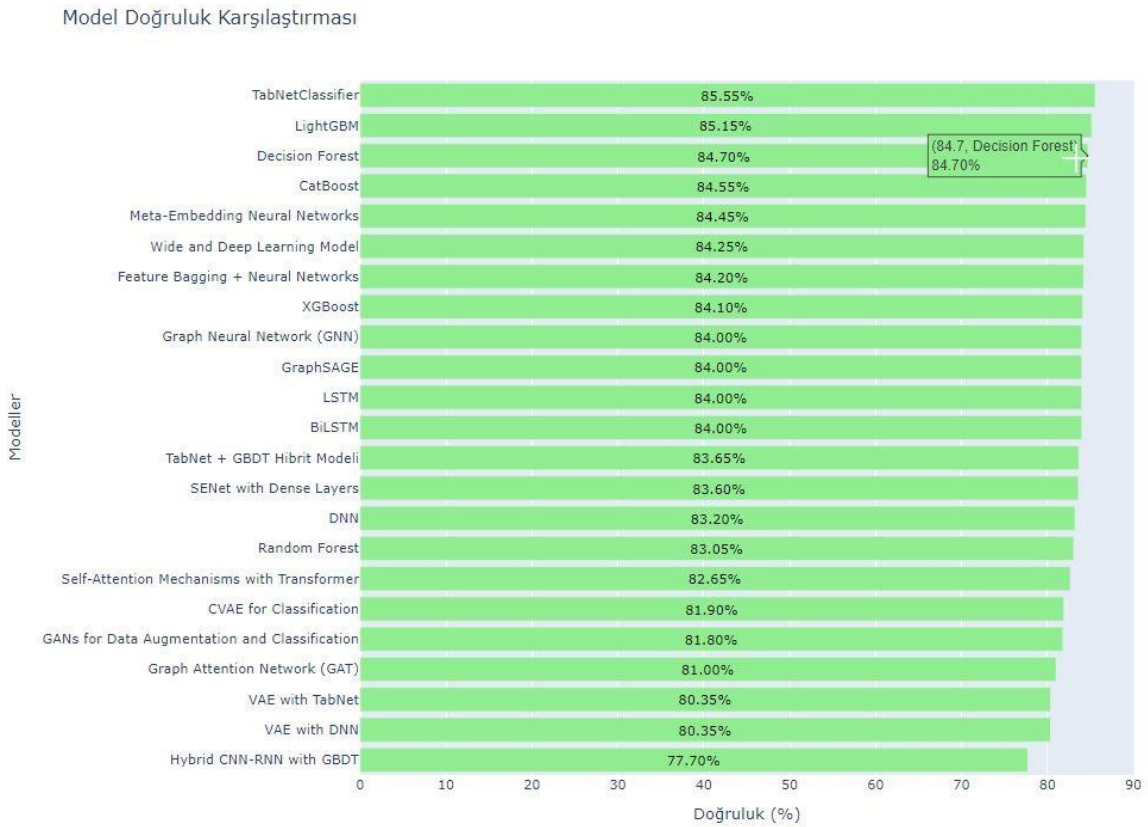
TabNet + GBDT Hibrit Modeli: TabNet ve gradyan artırma karar ağaçlarının hibrit modeli. Bu model, TabNet'in güçlü özellik seçimi ve GBDT'nin güçlü sınıflandırma yeteneklerini birleştirir.

Çalışmamızda kullanılan modellerin test doğruluk oranları aşağıda sunulmuştur:

<i>Model</i>	<i>Accuracy (%)</i>
TabNetClassifier	85.55
Decision Forest	84.70
Graph Neural Network (GNN)	84.00
Graph Attention Network (GAT)	81.00

GraphSAGE	84.00
LSTM (Long Short-Term Memory)	84.00
Random Forest	83.05
XGBoost	84.10
LightGBM	85.15
CatBoost	84.55
BiLSTM (Bidirectional LSTM)	84.00
Feature Bagging + Neural Networks	84.20
Wide and Deep Learning Model	84.25
DNN (Deep Neural Network)	83.20
VAE (Variational AutoEncoder) with TabNet	80.35
Self-Attention Mechanisms with Transformer	82.65
VAE with DNN	80.35
Meta-Embedding Neural Networks	84.45
SENet with Dense Layers	83.60
CVAE for Classification	81.90
GANs for Data Augmentation and Classification	81.80
Hybrid CNN-RNN with GBDT	77.70
TabNet + GBDT Hibrit Modeli	83.65

Tablo 4: Model Karşılaştırmaları(Accuracy Metriğine Göre)



Şekil 5 – Modellerin Doğruluk Metriğine Göre Karşılaştırılmasını Gösteren Grafik

IV. Sonuç

Bu çalışma, sigortacılık sektöründe araç sigortası taleplerinin sınıflandırılmasında kullanılan ileri düzey makine öğrenmesi, derin öğrenme ve hibrit tekniklerin performansını kapsamlı bir şekilde analiz etmektedir. Sigorta taleplerinin doğru ve hızlı bir şekilde sınıflandırılması, operasyonel maliyetlerin azaltılması, müşteri memnuniyetinin artırılması ve etkin risk yönetimi sağlanması açısından kritik bir öneme sahiptir. Çalışmada, farklı model ve

tekniklerin sınıflandırma yetenekleri doğruluk metriği temelinde karşılaştırılmış ve en uygun yaklaşımlar belirlenmiştir. Sigorta şirketlerinin rekabet avantajını sürdürebilmeleri için doğru model seçimi hayati bir gerekliliktir. Bu bağlamda, yapılan analizler, sigorta taleplerinin sınıflandırılmasında hangi tekniklerin daha etkili sonuçlar verdiğini net bir şekilde ortaya koymaktadır.

Model çeşitliliği, sigorta taleplerinin doğru ve hızlı bir şekilde sınıflandırılmasında önemli bir rol oynamaktadır. Örneğin, **TabNetClassifier**, tabular veriler üzerinde yüksek performans gösterirken, dikkat mekanizmaları kullanarak önemli özellikleri otomatik olarak seçme kabiliyeti sunar. **Graph Neural Network (GNN)** ise, grafik tabanlı veri yapılarında müşteriler arasındaki ilişkileri modelleyerek, talepler üzerindeki etkilerini analiz eder. **LSTM (Long Short-Term Memory)** modeli, zaman serisi verilerini işlemek için uygun olup, zaman faktörünün kritik olduğu durumlarda yüksek doğruluk sağlamaktadır. **Random Forest** gibi ansambl öğrenme modelleri, overfitting riskini azaltarak genel performansı artırır. **XGBoost**, **LightGBM** ve **CatBoost** gibi gradyan artırma modelleri ise yüksek hız ve performanslarıyla sigorta taleplerinin sınıflandırılmasında hassas ve doğru sonuçlar vermektedir.

Çalışmanın performans değerlendirmesi, farklı modellerin güçlü ve zayıf yönlerini ortaya koyarak sigorta şirketlerinin en etkili modelleme stratejisini seçmelerine yardımcı olmuştur. Modellerin doğruluk (accuracy) metriği ile değerlendirilmesi, her bir modelin sınıflandırma başarısını objektif bir şekilde ölçmüştür. Bu yaklaşım, karar verme sürecini daha bilinçli ve stratejik hale getirmiştir. Ayrıca, ansambl ve hibrit modellerin kullanılması, model performansını optimize etmiş ve çeşitli hatalara karşı daha dirençli hale gelmelerini sağlamıştır. Bu çeşitlilik, veri setindeki varyasyonlara karşı model dayanıklılığını ve genelleme kapasitesini artırarak, modellerin gerçek dünya verileri üzerinde daha iyi performans göstermesini sağlamıştır.

Sonuç olarak, bu çalışmanın bulguları, sigorta şirketlerine, sigorta taleplerinin sınıflandırılmasında hangi modellerin daha etkili olduğunu göstererek, daha verimli süreçler oluşturmalarına olanak tanımaktadır. Özellikle **TabNetClassifier** ve **Graph Neural Network (GNN)** gibi modellerin belirli durumlar için uygunluğu, sigorta taleplerinin hızlı ve doğru bir şekilde işlenmesini sağlamaktadır. Gelecek çalışmalarda, otonom ve yarı otonom araçlardan elde edilecek yeni veri tiplerinin kullanılması ve bu verilere dayalı modellerin geliştirilmesi hedeflenmektedir. Bu kapsamda, sigorta şirketlerinin risk yönetim stratejilerini optimize etmelerine yardımcı olacak daha yenilikçi ve etkili sınıflandırma yöntemlerinin geliştirilmesi planlanmaktadır. Farklı model ve tekniklerin kullanımı, sigorta sektöründe daha doğru ve güvenilir sonuçlar elde edilmesini sağlayarak, genel risk yönetimini ve karar verme süreçlerini iyileştirecektir.

Bu çalışma, sigorta taleplerinin sınıflandırılmasında kullanılan çeşitli makine öğrenmesi ve derin öğrenme modellerinin performansını kapsamlı bir şekilde değerlendirerek, sigorta sektöründe önemli kazanımlar sağlamaktadır. Doğru sınıflandırma modellerinin kullanılması, sigorta taleplerinin daha hızlı ve doğru bir şekilde işlenmesine olanak tanır, bu da operasyonel maliyetlerin azaltılmasına ve müşteri memnuniyetinin artırılmasına katkı sağlar. Ayrıca, doğru sınıflandırma, risk değerlendirmelerinin daha hassas yapılmasını mümkün kılarak, sigorta şirketlerinin risk yönetimi stratejilerini iyileştirmelerine yardımcı olur. Bu durum, şirketlerin rekabet avantajlarını sürdürebilmeleri için kritik bir gerekliliktir.

Model çeşitliliği ve uygun model seçimi, çalışmanın en önemli bilimsel katkılarından biridir. Farklı model yaklaşımlarının bir arada kullanılması, her bir modelin güçlü ve zayıf yönlerini dengeler ve genel performansı artırır. Örneğin, gradyan artırma modelleri (XGBoost, LightGBM, CatBoost) hızlı ve yüksek performanslarıyla öne çıkarken, derin öğrenme modelleri (LSTM, GNN) daha karmaşık veri yapılarının anlaşılmasında üstünlük sağlar. Bu çeşitlilik, modellerin genelleme kapasitesini artırarak, yeni ve görülmemiş veri üzerinde daha iyi performans göstermelerini sağlar. Ayrıca, farklı modellerin kombinasyonu, dayanıklılığı artırır ve çeşitli hatalara karşı daha dirençli hale getirir. Bu çalışma, sigorta sektöründe daha doğru ve güvenilir sonuçlar elde edilmesine katkıda bulunarak, genel risk yönetimi ve karar verme süreçlerini optimize etmektedir.

Ek:

Kullanılan modeller, sigorta taleplerinin sınıflandırılması gibi görevler için belirli avantajlar ve dezavantajlar sunar. Aşağıda, her bir modelin temel özelliklerini, üstünlüklerini ve dezavantajlarını ayrıntılı olarak ele alınmıştır:

1. TabNet

Avantajlar:

Özellik Seçimi: TabNet, dikkat mekanizmalarını kullanarak tabular verilerde önemli özellikleri otomatik olarak seçer ve işler. Bu özellik, modelin yüksek boyutlu verilerle başa çıkma yeteneğini artırır.

Yorumlanabilirlik: TabNet'in dikkat mekanizmaları, modelin karar sürecini daha anlaşılır hale getirir ve yorumlanabilirlik sağlar.

Performans: Tabular verilerde genellikle yüksek performans sergiler.

Dezavantajlar:

Veri Türü Kısıtlaması: Performansı, çoğunlukla tabular veri ile sınırlıdır; diğer veri türleriyle (örneğin, metin veya görüntü) performansı sınırlı olabilir.

Model Kompleksliği: Dikkat mekanizmaları ve tabular veri işlemleri nedeniyle daha karmaşık bir yapıdadır, bu da eğitim ve çıkarım sürecini yavaşlatabilir.

2. Graph Neural Network (GNN)

Avantajlar:

Grafik Verilerle Çalışma: GNN, grafik verileri (örneğin, sosyal ağlar veya etkileşim grafikleri) modellemede etkilidir. Sigorta taleplerinde, müşteriler arasındaki ilişkileri ve bu ilişkilerin talepler üzerindeki etkilerini modelleyebilir.

Veri Bağlantılarını Öğrenme: GNN, düğümler arasındaki ilişkileri ve bu ilişkilerin özellikler üzerindeki etkilerini öğrenme kapasitesine sahiptir.

Dezavantajlar:

Ölçeklenebilirlik Sorunları: Büyük ölçekli grafik verilerde hesaplama maliyetleri artabilir.

Veri Türü Kısıtlaması: GNN'ler, grafik yapısına sahip verilerde en iyi performansı gösterir; diğer veri türleri için uygun olmayabilir.

3. XGBoost

Avantajlar:

Hız ve Performans: XGBoost, gradyan artırma algoritmasının optimize edilmiş bir versiyonudur ve yüksek hız ve doğruluk ile bilinir.

Genel Performans: Dengeli bir şekilde hem doğruluk hem de hesaplama hızında mükemmel performans sunar.

Özelleştirilebilirlik: Birçok hiperparametre ayarına sahiptir, bu da modeli belirli veri setlerine uyarlamayı kolaylaştırır.

Dezavantajlar:

Overfitting Riski: Karmaşık modellerde ve küçük veri setlerinde aşırı öğrenme riski taşır.

Hiperparametre Ayarlama: Çok sayıda hiperparametre ayarının yapılması gerekebilir, bu da modeli karmaşık hale getirebilir.

4. LightGBM

Avantajlar:

Hız ve Verimlilik: LightGBM, XGBoost'a göre daha hızlıdır ve büyük veri setlerinde daha verimli çalışır.

Özellik Seçimi: Katkısız yaprak büyüme stratejisi kullanarak, önemli özellikleri daha iyi öğrenir.

Düşük Bellek Kullanımı: Daha az bellek kullanarak çalışır, bu da büyük veri setleriyle çalışmayı kolaylaştırır.

Dezavantajlar:

Özelleştirilebilirlik: XGBoost kadar esnek değildir, özellikle hiperparametre ayarları daha sınırlıdır.

Dengesiz Veri Setlerinde Performans: Dengesiz veri setlerinde performans düşüşü yaşanabilir.

5. CatBoost

Avantajlar:

Kategorik Verilerle Çalışma: Kategorik özellikleri otomatik olarak işler ve bu tür verilerde yüksek performans gösterir.

Hız ve Performans: XGBoost ve LightGBM ile karşılaştırılabilir performansa sahiptir.

Overfitting Azaltma: Aşırı öğrenmeyi azaltmak için çeşitli yerleşik düzenleme teknikleri kullanır.

Dezavantajlar:

Karmaşıklık: Model yapısı karmaşıktır ve eğitim süresi diğer gradyan artırma modellerine kıyasla daha uzun olabilir.

Bellek Kullanımı: Büyük veri setlerinde daha fazla bellek kullanımı gerekebilir.

6. LSTM (Long Short-Term Memory)

Avantajlar:

Zaman Serisi Verileri İçin İdeal: LSTM, uzun süreli bağımlılıkları öğrenmede etkilidir ve bu nedenle zaman serisi analizlerinde mükemmel performans gösterir.

Sıralı Veri İşleme: Metin veya konuşma gibi sıralı verilerle çalışırken güçlüdür.

Dezavantajlar:

Eğitim Süresi: Geri yayılım ve diğer eğitim süreçleri nedeniyle eğitim süresi uzayabilir.

Hiperparametre Hassasiyeti: Hiperparametre ayarlarına oldukça duyarlıdır ve uygun olmayan ayarlar performansı olumsuz etkileyebilir.

7. Random Forest

Avantajlar:

Genel Performans: Birden fazla karar ağacının çoğunluk oyu ile karar verdiği için overfitting riskini azaltır.

Basitlik ve Yorumlanabilirlik: Göreceli olarak basit bir modeldir ve genellikle yorumlanabilir sonuçlar sağlar.

Dengeleme: Dengesiz veri setleriyle çalışırken iyi performans gösterir.

Dezavantajlar:

Hesaplama Maliyeti: Büyük veri setleri veya birçok ağaç kullanıldığında hesaplama maliyeti artabilir.

Esneklik Eksikliği: Daha karmaşık veri ilişkilerini öğrenmede sınırlıdır.

8. BiLSTM (Bidirectional LSTM)

Avantajlar:

Çift Yönlü Bağımlılık Öğrenme: Verinin hem geçmiş hem de gelecekteki bağlamını öğrenebilir, bu da daha doğru tahminler yapmasını sağlar.

Sıralı Veri İçin Uygun: Metin ve dil işleme görevlerinde güçlüdür.

Dezavantajlar:

Eğitim Zamanı: LSTM'den daha fazla hesaplama gücü ve eğitim süresi gerektirir.

Model Karmaşıklığı: Daha karmaşık ve ayar gerektirir, bu da modelin yapılandırılmasını zorlaştırabilir.

9. TensorFlow Decision Forests

Avantajlar:

Esneklik ve Entegrasyon: TensorFlow platformuyla entegre çalışır ve diğer derin öğrenme yöntemleriyle birlikte kullanılabilir.

Çoklu Model Desteği: Karar ormanı, gradyan artırma ağaçları gibi birden fazla model türünü destekler.

Skalabilite: Büyük veri setleri üzerinde iyi çalışır ve TensorFlow'un GPU desteğinden yararlanabilir.

Dezavantajlar:

Karmaşıklık: Kullanımı daha karmaşıktır ve TensorFlow ekosistemine hâkim olmayı gerektirir.

Yüksek Öğrenme Eğrisi: TensorFlow kullanımı yeni başlayanlar için zor olabilir.

Bu modellerin her biri, farklı veri tipleri ve problem türleri için avantajlar ve dezavantajlar sunar. Model seçimi, genellikle veri setinin yapısına, problemi çözme gereksinimlerine ve eldeki kaynaklara bağlı olarak yapılmalıdır.

IMDb Score Estimation using Movie Dialogues: A Text Mining and Machine Learning Hybrid Approach

Yunus EROĞLU*

Industrial Engineering / Engineering Faculty, Gaziantep University, Gaziantep, Türkiye

Abstract

IMDb scores in the movie industry provide viewers crucial information about the quality and watchability of movies. This is why viewers often refer to the IMDb score, which reflects the opinions of others about the movie, before deciding to watch it, allowing them to assess the quality of the movie. In this paper, the estimation of IMDb scores is modeled using English subtitles containing dialogues of several movies from different genres and release times with IMDb scores. In the beginning, 500 different movies from the IMDb database were organized in a separate file along with the movie ID, title, release date, number of votes, duration, and IMDb rating. The English subtitles of these movies were gathered from various websites and collected in a separate directory for text mining analysis. The python-based Orange 3 package program, which allows for easy modeling of data mining and machine learning tools, was selected for the methods applied in this study. In the proposed model, movie subtitles are first subjected to text preprocessing where unwanted symbols, numbers and phrases are removed. Text embedding is then performed to convert textual expressions into numeric vectors. Subsequently, each movie subtitle was converted into a 300-dimensional numeric vector and the dataset was transformed into a 500x300 matrix. Afterwards, the IMDb score for each movie was added to the newly obtained matrix as the target parameter to be estimated. The dataset, divided into 70% training and 30% test sets, was trained using AdaBoost and kNN methods and the obtained models were tested on the test sets. The proposed machine learning methods perform well on the training set with almost 100% accuracy, while their performance on the test set is acceptable with about 70% consistency. The most notable aspect of this study is that IMDb scores were predicted without considering any information related to actors, directors, budget, film platform, or origin, demonstrating the model's applicability in their absence. In this regard, it can be suggested that the approximate IMDb score of a film can be predicted solely based on its dialogues, even before the film is produced, when details such as actors, directors, producers, and budget are not yet available. This allows both producers and screenwriters to use the estimated IMDb score when making critical decisions about the movie such as the film's budget, casting choices, or whether the script requires further development.

Key words: IMDb score, Text mining, Machine learning, Estimation

INTRODUCTION

The Internet Movie Database (IMDb) is one of the most well-known digital platforms globally, offering a vast knowledge base on the movie and television industry. IMDb provides access to information on various elements of cinema, such as movies, TV series, actors, and directors, as well as a system that allows users to rate this content. The platform was founded in 1990 and acquired by Amazon in 1998, becoming an important reference source in both academic and commercial circles. IMDb's one of the most remarkable features is that users rate the content they watch by scoring it between 1 and 10. These scores represent the audience's collective opinion about a movie or TV series and form the basis for various analyses. Ratings are an important indicator for both audiences and industry professionals. Especially film producers, distributors and investors take IMDb scores into account when evaluating the success potential of films (Lin, 2018). Viewers consider high-rated movies as a more reliable option in terms of watchability, mostly taking IMDb scores into account. Movies with an IMDb score of 7 and above generally achieve high ratings and remain popular for a long time. IMDb scores are also an invaluable piece of information for movie producers and distributors. Movies with high IMDb scores tend to have a wider distribution network and a larger marketing budget. These scores are considered as an important factor in the selection of films to be released in cinemas and in determining the content to be featured on digital platforms.

The primary goal of this paper is to examine the impact of movie dialogues in predicting IMDb scores. In the existing literature, as far as the authors have been able to investigate, factors such as cast, director, budget and

production quality are often used to predict the quality of a movie. The effects of different factors affecting movie success were investigated by Bhave et al. (2015) and modeled by considering both movie characteristics such as cast, producer, director and user reviews on different online platforms (Bhave et al., 2015). Dhir and Raj (2018) developed a method that combines data already available in the IMDb database (director, gross, budget, genre, etc.) with social media analytics (number of votes, number of reviews, number of Facebook likes, etc.) to predict movie success. Using machine learning models, they inferred the factors affecting movie success (Dhir & Raj, 2018). Bristi et al. (2019) proposed machine learning algorithms that predict the success of movies using the movie's own features such as Director, Screenplay, Actors/Actress, Country, and Genre (Bristi et al., 2019). Singh et al. (2024) tried to predict the success of a movie with kNN, Random Forest and Logistic Regression methods using such parameters as title, genre, director, country, language, and year of release. (Singh et al., 2024).

However, this study aims to investigate the potential of predicting IMDb scores as a quality indicator by focusing only on movie dialogues. Dialogues are one of the most important components of a movie's impact on the viewer, and this study will seek to answer the question of whether IMDb scores can be successfully predicted based on dialogues alone. Such an approach has the potential to provide producers and screenwriters with a tool that can help them make strategic decisions during the production phase of a movie. These early insights can be instrumental in making strategic decisions such as the production budget, casting, and script development.

The next section provides information about the data set and methods used. Then, the findings of the model are presented the results section. The last part of this paper summarizes the conclusions, discussions, limitations, and possible future work.

MATERIAL AND METHODS

Material

The data for this study were obtained from the IMDb database and various movie dialog texts. The data used in this study includes basic information about the movies from the IMDb platform (IMDb, 2024). IMDb is a database that provides information on millions of movies, TV series, actors, and filmmakers worldwide. In this study, the data obtained from IMDb includes Movie ID, Movie Name, Release Date, IMDb Score, Vote Count and Duration information.

Data based on movie dialogues are the English subtitles of the relevant movies collected from various online sources. These subtitles were analyzed with natural language processing (NLP) methods because they present the dialogue content of the films as text. The collected dialogues were used as one of the main components to predict the IMDb ratings of the movies.

Methods

The Collection of the Data

The data is collected from IMDb platform and various online subtitle sources. The collected dialogue data was analyzed using text mining and machine learning techniques to predict the IMDb scores of the films. To understand the relationship between dialogues and IMDb scores, text preprocessing and embedding methods were used. The study covers the process of creating predictive models with Python-based Orange 3, particularly natural language processing, text preprocessing and machine learning algorithms.

Research Question

The main research question of this study can be expressed as follows:

"Can IMDb scores be predicted accurately with machine learning methods based only on movie dialogues, without taking into account traditional factors such as movie cast, director, budget, genre?"

This question aims to understand the impact of dialogues on IMDb scores. It also aims to examine scoring systems using a non-traditional approach. The study will investigate how text mining and machine learning techniques

answer this question.

Viewers generally give their IMDb scores after watching the movie, considering many factors such as the cast, director, effects used, etc., and these scores often provide clues about the success and quality of the movie. However, this study examines the effect of film dialogues alone on predicting IMDb scores, independent of these traditional factors.

Data Processing

In this study, models were developed using text mining method, which has been successfully applied in many different fields (Atay et al., 2021, Eroglu & Seçkiner, 2016, and Eroglu, 2023), to predict IMDb ratings from subtitles. Subtitles were converted into numerical data using natural language processing techniques and then analyzed with machine learning algorithms. This process involved a series of steps, and each step transformed the subtitles from text format into numerical data that could be analyzed by machine learning models.

In the text preprocessing phase, the dialogues were divided into sentences and words (tokenization) and broken into smaller pieces. Then, all words were converted to lower case to ensure consistency in the texts. Frequently used words such as “the”, “is”, “and” that do not make sense in English were removed from the data set. Unnecessary punctuation marks, numbers, and special characters were also cleaned. Finally, words were reduced to their roots with stemming/lemmatization; for example, words such as “playing”, “played”, and “plays” were converted to the root “play”.

As a result of these steps, the dialogues were transformed into a more organized and suitable format for analysis.

Document Embedding

Document embedding refers to the process of converting entire documents (or larger units of text) into digital vectors of fixed length. These vectors capture the semantic content of the text. Unlike simpler methods such as Bag of Words (BoW) or TF-IDF (Term Frequency-Inverse Document Frequency), document embedding methods preserve the context and relationships between words within the text.

fastText is a document embedding model developed by Facebook's Artificial Intelligence Research (FAIR) team. It considers sub-word information (character-level n-grams), extending traditional word embedding methods such as Word2Vec. This feature allows fastText to handle rare words, misspellings, and grammatically rich languages more effectively than traditional word embedding methods. In this study, because of the fastText method, each subtitle file was converted into a 300-dimensional digital vector, thus facilitating the use of machine learning algorithms.

Machine Learning Algorithms

In this study, k-Nearest Neighbor (kNN) and AdaBoost algorithms are used to predict IMDb rating based on movie dialogues.

The kNN algorithm has an easy-to-understand and simple implementation and does not require rule-based or complex modeling. At the beginning of the study, a simple algorithm was chosen to test the predictability of IMDb ratings based on dialogues. Being a nonparametric method, kNN does not make specific assumptions about the data distribution, which allows for flexible exploration of complexities and relationships within the data.

AdaBoost combines simple models to create a powerful model. Weak learners are models that make very simple predictions, but AdaBoost's iterative structure corrects these learners' errors incrementally. AdaBoost's ability to improve the errors of a simple model makes it an ideal choice for complex data structures, such as movie dialogue.

MSE (Mean Squared Error): It measures the mean squared difference between the predicted values and the actual values. A lower MSE indicates better performance.

- *RMSE (Root Mean Squared Error)*: The square root of the MSE and makes the error more interpretable by putting it in the same units as the target variable.

- *MAE (Mean Absolute Error)*: Measures the average absolute difference between predicted and actual values. It is more robust to outliers than MSE.
- *MAPE (Mean Absolute Percentage Error)*: Shows the percentage difference between predicted and actual values, useful for understanding relative errors.
- *R² (R-Square)*: Represents the proportion of variance in the target variable explained by the model. An R² value closer to 1 indicates a better fit.

Table 1 shows the performance of the models on the training set. Both models performed exceptionally well on the training dataset, achieving nearly 100% accuracy. This indicates that the models learned the patterns in the training data effectively.

Table 1. Performance comparisons of proposed models on training dataset

Model	MSE	RMSE	MAE	MAPE	R ²
kNN	0.000	0.000	0.000	0.000	1.000
AdaBoost	0.003	0.052	0.016	0.002	0.999

Table 2 presents the performance of the models on the test set. AdaBoost gave better results than kNN in all metrics, making more accurate predictions with lower MSE, RMSE, MAE and MAPE values. Additionally, it had a higher R² value and better fit the data. These results confirm that AdaBoost is a more effective model for predicting IMDb ratings based on movie dialogues.

Table 2. Performance comparisons of proposed models on test dataset

Model	MSE	RMSE	MAE	MAPE	R ²
kNN	2.154	1.468	1.028	0.252	0.552
AdaBoost	1.357	1.165	0.845	0.219	0.699

CONCLUSION AND DISCUSSION

This study has successfully demonstrated that IMDb scores can be predicted using only movie dialogue, without relying on traditional factors such as cast, budget, or director. By applying k-Nearest Neighbor (kNN) and AdaBoost machine learning models to a dataset consisting of 500 movies, it was possible to predict IMDb scores using only text data obtained from movie subtitles. The results revealed that the AdaBoost model performed better, achieving lower error rates (MSE, RMSE, MAE, MAPE) and higher R² value on the test set. AdaBoost explains approximately 69.9% of the variance in IMDb scores, while kNN explains 52.2%. This study proves that movie dialogues alone can be a useful indicator for predicting IMDb ratings, and this potential is especially strong when advanced machine learning techniques such as AdaBoost are used.

The ability to predict IMDb scores based on movie dialogue opens exciting possibilities for decision-making in the early stages of the movie production process. Screenwriters and producers can use dialogue-based models to evaluate a film's potential success before making significant investments in casting, budgeting, and production. This can make risk management and resource allocation more strategic in the movie industry. This could help optimize scripts or gauge whether further revisions are needed to improve the projected reception of the movie. Additionally, this approach could benefit platforms like Exxen, tabii, GAIN, Netflix, or Amazon, which are increasingly investing in original content and may want to evaluate scripts before committing to full-scale production. Therefore, the findings suggest a new and potentially powerful tool for early-stage decision-making in the movie industry.

The study has some limitations. First, the dataset used consists of only 500 movies, which may not fully reflect the wide diversity of the film industry, especially the diversity between different genres and languages. This limited dataset may affect the generalizability of the results, and it is thought that working with a larger dataset would provide more reliable results. In addition, the AdaBoost model used in the study achieved almost 100% accuracy on the training set, which brings with it the risk of overfitting. Overfitting the model to the training data means that

it may not show the same success when faced with previously unseen datasets.

Another limitation is that no sentiment analysis or sentiment analysis was performed on the dialogues. However, the emotional impact of movies can have a strong relationship with audience scores. Modeling emotional content could have predicted IMDb ratings more accurately, especially for dramatic or emotionally intense movies. In addition, although the kNN and AdaBoost models used in the study are effective in text-based predictions, more advanced natural language processing models are available today. Advanced models such as transformer models (e.g. BERT) could better understand complex relationships in texts and provide more accurate predictions. Therefore, the fact that more advanced models were not used in the study may limit the performance of the results obtained.

The study also focused only on predicting IMDb scores. However, other metrics are also important for evaluating the success of a movie, such as box office revenue or critic reviews. Considering such additional metrics could have allowed for a more comprehensive evaluation of movie success. As a result, the model did not consider temporal or cultural factors. Factors such as the year the movie was released, or the cultural preferences of the audience may affect IMDb scores. Including these variables in the model could have contributed to obtaining more accurate results.

Possible future research topics require in-depth studies in different areas. First, sentiment analysis and sentiment detection can be added to dialogue-based models. In this way, it may be possible to predict IMDb scores of movies more accurately with emotional intensity. In addition, the use of advanced natural language processing (NLP) models such as BERT, GPT, and Transformer is another research opportunity. Such deep learning models can improve prediction performance by understanding more complex relationships within dialogues.

Another important area of research is expanding the dataset. Increasing the size and diversity of the dataset by including more movies from different genres, languages, and regions can strengthen the generalization ability of the model. In addition, other success metrics can be estimated beyond IMDb scores. For example, a more comprehensive assessment of film success can be made using metrics such as box office revenue, critical reviews, or audience retention rates. The use of dialogue-based models in early-stage script evaluations can help writers and producers predict potential success before a film begins production. It is also suggested that this model be applied to non-English films. Similar predictive patterns can be explored in different languages and cultural contexts. Finally, analyzing time-based trends may be important to understand how the impact of dialogue on IMDb ratings has changed over time.

References

- Atay, M., Eroğlu, Y., & Ulusam Seçkiner, S., 2021. Investigation of Breaking Points in the Airline Industry with Airline Optimization Studies Through Text Mining before the COVID-19 Pandemic. *Transportation Research Record*, 2675(5), 301–313. <https://doi.org/10.1177/0361198120987238>
- Bhave, A., Kulkarni, H., Biramane, V., & Kosamkar, P., 2015. Role of different factors in predicting movie success. *2015 International Conference on Pervasive Computing (ICPC)*, 1–4. <https://doi.org/10.1109/PERVASIVE.2015.7087152>
- Bristi, W. R., Zaman, Z., & Sultana, N., 2019. Predicting IMDb Rating of Movies by Machine Learning Techniques. *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 1–5. <https://doi.org/10.1109/ICCCNT45670.2019.8944604>
- Dhir, R., & Raj, A. (2018). Movie Success Prediction using Machine Learning Algorithms and their Comparison. *2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC)*, 385–390. <https://doi.org/10.1109/ICSCCC.2018.8703320>
- Eroglu, Y. (2023). Text Mining Approach for Trend Tracking in Scientific Research: A Case Study on Forest Fire. *Fire*, 6(1), 33. <https://doi.org/10.3390/fire6010033>

Eroglu, Y., & Seçkiner, S. U. (2016). Trend Topic Analysis for Wind Energy Researches: A Data Mining Approach Using Text Mining. *Journal of Technology Innovations in Renewable Energy*, 5(2), 44–58. <https://doi.org/10.6000/1929-6002.2016.05.02.2>

IMDb. (2024). *IMDb Data Files Download*. <https://datasets.imdbws.com/>

Lin, Q. (2018). *Prediction of a film's profit and its IMDB score*. Stockholmes Universitet.

Singh, K. K., Makhania, J., & Mahapatra, M. (2024). Impact of ratings of content on OTT platforms and prediction of its success rate. *Multimedia Tools and Applications*, 83(2), 4791–4808. <https://doi.org/10.1007/s11042-023-15887-9>

Artificial Neural Networks in Marketing: A Comprehensive Review of Consumer Behavior Applications

İbrahim Halil Efendioğlu^{1*}

¹Business Administration/ Gaziantep University, Gaziantep, Turkey

Abstract

This study aims to provide a comprehensive review of the literature on how artificial neural networks (ANNs) have been applied in consumer behavior research in marketing. The potential of ANNs to enhance the accuracy and effectiveness of marketing strategies is linked to their ability to model complex data relationships. The study highlights how ANNs contribute to marketing processes such as customer segmentation, sales forecasting, and dynamic pricing strategies while also improving customer satisfaction and the effectiveness of loyalty programs. However, it also addresses limitations such as the lack of transparency, high data requirements, and ethical concerns associated with ANNs. By synthesizing the existing body of research, this study examines how ANNs have been used to predict consumer behavior, the challenges encountered, and the directions for future research. Furthermore, it identifies significant gaps in real-time data integration, model transparency, and data privacy. The findings emphasize the transformative potential of ANNs in marketing strategies while shedding light on critical areas that future research should address.

Keywords: Artificial Neural Networks, Consumer Behavior, Marketing

INTRODUCTION

The rapidly evolving intersection of artificial intelligence (AI) and consumer behavior in marketing has garnered significant interest both academically and practically (Abrardi et al., 2022). Among these AI technologies, artificial neural networks (ANNs) stand out as a unique and powerful tool for understanding and predicting consumer behavior. The utilization of artificial neural networks (ANNs) in consumer behavior offers significant advantages in the development of marketing strategies and the management of customer relationships. ANNs can analyze large and complex datasets, enabling more accurate predictions of consumer behavior. This allows for more effective targeting of marketing campaigns (Boozary, 2024). Additionally, ANNs can study consumer behavior to provide personalized recommendations tailored to individual consumers, thereby enhancing customer satisfaction and increasing the effectiveness of loyalty programs. Moreover, ANNs can predict future consumer trends based on past data and behaviors, giving businesses a strategic advantage by anticipating trends (Karim et al., 2024). Furthermore, ANNs help segment consumers into groups with similar characteristics, facilitating the development of optimal marketing strategies for each segment, leading to more efficient resource allocation. ANNs can also optimize customer service processes by learning from customer interactions and feedback, resulting in an improved customer experience (Dang et al., 2023). On the other hand, ANNs enhance the accuracy of sales forecasts through consumer behavior analysis and improve efficiency in inventory management, which contributes to cost reduction and increased profit margins. Therefore, the ability of ANNs to model complex patterns and relationships within data offers businesses unique opportunities to analyze consumer data more precisely and effectively, thereby improving their marketing strategies (Mkedder & Bakır, 2023). The use of ANNs and the research on this subject in predicting consumer behavior is increasing day by day.

In recent years, many studies have used the ANN approach to predict consumer preferences and decision-making processes (Albahri et al., 2022; Ashfaq et al., 2023; Boozary, 2024; Hajek et al., 2020; Kalinić et al., 2021; Klietnik et al., 2022; Lee et al., 2022; Leong et al., 2021; Ma et al., 2023; Mishra et al., 2023; Nica et al., 2022; Sharma et al., 2021; Wang et al., 2023). However, there is no comprehensive review that synthesizes the findings in these studies and evaluates the combined use of consumer behavior and the ANN approach. This gap in the literature underscores the need for a systematic review that consolidates existing knowledge and identifies the applications and limitations of ANNs in the marketing and consumer behavior context. Therefore, an up-to-date and comprehensive literature review on the various uses of ANNs in consumer behavior will enable the gaps in this area to be seen more clearly.

Accordingly, the aim of this study is to systematically review the literature on the use of ANNs in consumer behavior research, providing a comprehensive overview of the current state of knowledge in this area. The study also aims to identify critical areas in the existing literature, offer suggestions for future research, and provide guidance on how ANNs can be more effectively utilized in the marketing context. Thus, this study will provide valuable insights, aiming to contribute to the development of knowledge in this field.

This study is significant for several reasons: First, by deeply synthesizing research on ANNs and consumer behavior, it contributes to the marketing literature by highlighting key trends, challenges, and opportunities in this field. Second, it identifies gaps in the existing literature, offering valuable insights for future research and guiding new directions and methodologies in this area. Third, the study emphasizes the importance of integrating AI technologies like ANNs into marketing strategies and demonstrates how these tools can enhance the accuracy and effectiveness of marketing efforts.

Ultimately, the study aims to examine how ANNs have been applied in consumer behavior research, systematically review the existing literature in this field, and identify key trends, challenges, and opportunities. Thus, the study seeks to address the following research questions:

RQ1: What are the current trends in the literature regarding the use of ANNs in marketing and consumer behavior?

RQ2: What methodological approaches have been adopted for predicting consumer behavior using ANNs?

RQ3: What are the applications, advantages, limitations, and ethical challenges of ANNs in the literature?

RQ4: What gaps exist in the current literature, and how can future research address these gaps?

The remainder of this paper is organized as follows: First, the literature review section will explain ANNs and examine the existing literature on how ANNs have been used in consumer behavior research. Next, the methodology section will detail how the study was conducted, including the systematic literature review (SLR) method and the research process. The findings section will present the main findings from the literature review, and the use of ANNs together with consumer behavior will be examined. The discussion section will address the theoretical and practical contributions of the findings, discuss gaps in the literature, and offer suggestions for future research. Finally, in the conclusion section, the general findings of the study will be summarized and suggestions for future research will be presented.

LITERATURE REVIEW

Artificial Neural Networks (ANN)

Artificial Neural Networks (ANNs) are a powerful method of artificial intelligence (AI) that models the functioning of biological neural systems. They are particularly useful in data processing, where they process data through a series of interconnected nodes or 'neurons,' mimicking the learning and information processing capabilities of the human brain (Zou et al., 2009). First theoretically proposed by McCulloch and Pitts in 1943, ANNs have evolved over time into a powerful tool used to solve various complex problems, particularly in the field of data processing (Chakraverty et al., 2019).

The theoretical foundation of ANNs is based on how biological nerve cells interact with each other. A biological neuron receives signals from other cells through dendrites, processes them in the cell body, and transmits them to other cells through the axon. Similarly, ANNs receive input data, process this data, and produce an output (Rigby et al., 2019). Each neuron is associated with a specific weight, updated throughout the learning process (Sharma et al., 2017):

Core Components of ANNs:

Layers: ANNs typically consist of an input, hidden, and output layer. The input layer is where data is introduced to the network, the hidden layers process this data, and the output layer generates the final prediction or classification.

Weights: Each connection is associated with a weight that determines the strength of the signal. These weights, not to be underestimated, significantly influence how the network processes data and are optimized during the learning process.

Activation Functions: Activation functions determine how neurons process input data. Standard activation functions include sigmoid, ReLU (Rectified Linear Unit), and tanh.

Learning Process: ANNs learn using techniques such as the backpropagation algorithm. In this process, the difference between the network's prediction and the actual value is calculated, and this error is used to update the weights.

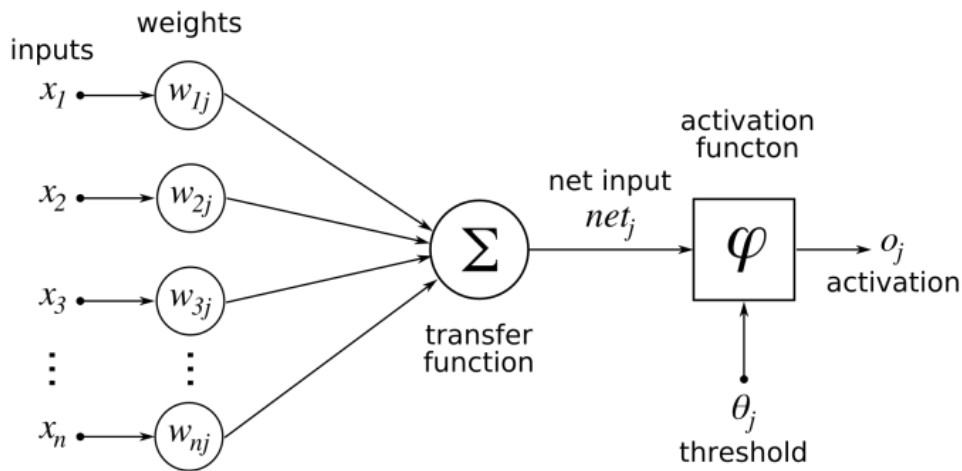


Figure 1. Artificial Neural Networks (ANN)
Source: Salah Alaloul & Hannan Qureshi (2020)

Use of ANNs in Predicting Consumer Behavior

Predicting consumer behavior plays a critical role in the development of marketing strategies. These predictions are used in various applications, such as customer segmentation, personalized marketing, demand forecasting, and enhancing customer loyalty (See Figure 2). ANNs are considered highly effective tools for making these predictions (Kasem et al., 2024):

Customer Segmentation: ANNs are used to analyze consumer data and identify customer groups with similar behaviors. These segments enable the creation of more targeted and effective marketing campaigns. For example, a retail company can analyze customers' purchasing habits to identify different consumer groups and offer customized promotions for each group.

Purchase Prediction: ANNs can analyze past purchasing data to predict future buying behaviors. This provides significant advantages in inventory management, promotional strategies, and supply chain optimization. For instance, an e-commerce platform can use users' previous purchase histories to predict when certain products are likely to be repurchased.

Optimization of Marketing Strategies: ANNs can optimize the effectiveness of marketing campaigns by analyzing consumer feedback, social media interactions, and other data. This process involves analyzing which messages are most effective for different consumer groups and determining which strategies best enhance customer loyalty.

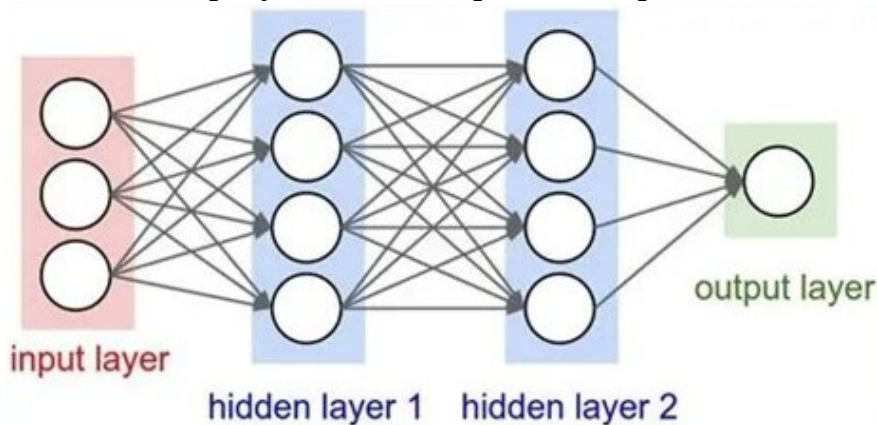


Figure 2. ANN Prediction Layers
Source: Akdagli (2021)

Advantages and Limitations of ANNs

ANNs are highly capable of handling vast and complex datasets. They excel in modeling non-linear relationships and making highly accurate predictions. These qualities make ANNs ideal for solving complex problems such as predicting consumer behavior. However, ANNs require large amounts of data and are computationally expensive techniques (Kurani et al., 2023). Additionally, the network's learning process can be challenging to understand, leading to what is known as the "black box" problem, where the internal workings of the network are not easily interpretable (Fraternali et al., 2023). ANNs have emerged as a powerful tool for predicting consumer behavior. With a theoretical background rooted in biological neural systems, this technology not only plays a critical role but also empowers in developing marketing strategies. However, the challenges and limitations associated with using ANNs must also be considered. Therefore, further research on the applications of ANNs in the field of marketing is necessary.

Methodology

This study employed the systematic literature review (SLR) method to delve into the relationship between ANNs and consumer behavior. The SLR, a method that systematically, transparently, and reproducibly reviews the existing literature in a specific research area, was structured according to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) protocol (Attaoui & Gaber, 2024). This protocol, which sets forth recommended reporting standards for systematic reviews and meta-analyses, is crucial in maintaining the rigor and transparency of the research process, thereby enhancing the credibility of our study. It encompasses four key steps: defining the scope of the literature review, data collection, setting criteria for article selection, and analyzing the findings.

The first step involved outlining the general framework of the literature to be reviewed and identifying relevant studies using specific keywords. In the second step, the data collection process was initiated, and articles meeting the defined criteria were meticulously screened, demonstrating the thoroughness of our research. The third step was dedicated to setting inclusion and exclusion criteria, and studies were meticulously selected from the literature based on these criteria. Finally, the collected data were meticulously analyzed to examine the impact of ANNs on consumer behavior in detail.

The literature review was conducted using the Scopus database, a reliable academic database that includes high-impact publications. Additionally, Scopus is one of the most comprehensive databases in fields such as marketing and business. It also has a large collection of articles from high-quality and peer-reviewed journals.

Thus, a literature search was performed using pre-determined keywords in the detailed database of Scopus. The search strategy employed the following keywords: (consumer behavior) AND (artificial neural networks), aimed at identifying studies that explore the relationship between ANNs and consumer behavior. After the search, only articles were selected, and the review was limited to publications from 2020 to 2024. This limitation was intended to focus on the most recent and high-quality studies evaluating the impact of ANNs on consumer behavior. Finally, articles related to the marketing field were selected, ensuring the credibility and relevance of the literature review.

Specific inclusion and exclusion criteria were used to select the studies from the literature review, ensuring a thorough and objective selection process (see Figure 3).

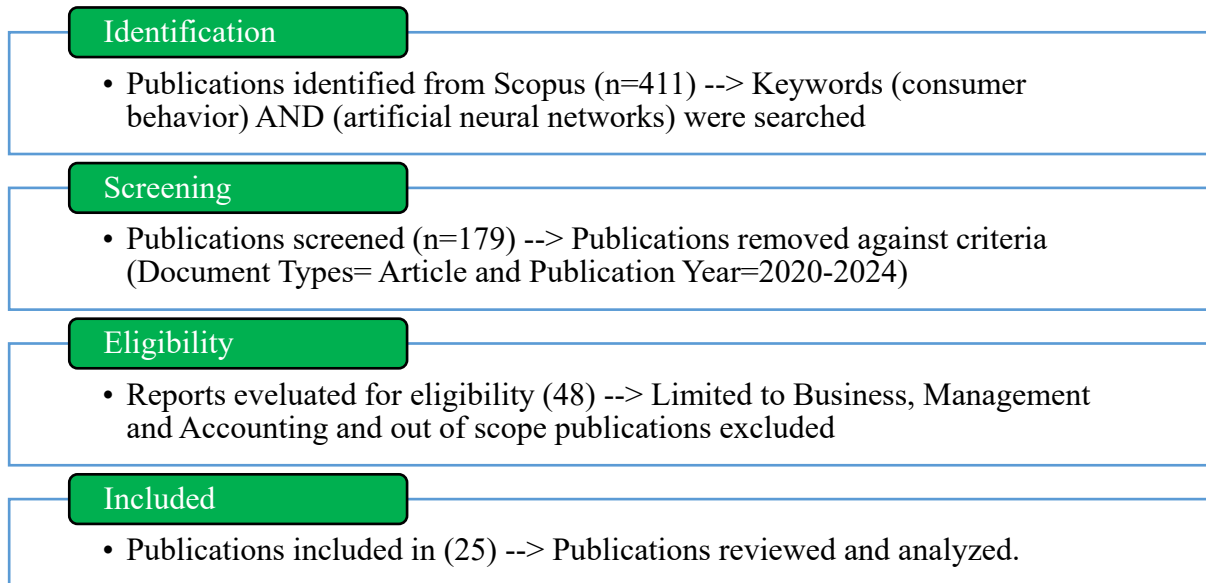


Figure 3. PRISMA Process

Inclusion Criteria

Time Frame: Publications from 2020 to 2024 in the Scopus database were considered. This selection aims to capture the most recent developments in applying ANNs in marketing and consumer behavior.

Type of Publication: Only articles published in peer-reviewed journals were included. Conference papers, book chapters, and other types of publications were excluded. This selection ensures a focus on high-quality research that has been recognized within the field.

Scope: Among the articles, only those related to marketing were selected. Studies that did not focus on consumer behavior or examined ANNs in contexts outside of marketing were excluded.

Exclusion Criteria

Conference Papers and Book Chapters: These publications were excluded, with only journal articles being evaluated.

Studies Not Related to Marketing: Research that did not directly pertain to marketing and consumer behavior was excluded.

Initial Screening Results: The initial search in the Scopus database identified 411 publications, but this number was reduced to 25 based on the exclusion criteria outlined above.

The selected articles were evaluated using the SLR analysis method. Each article was analyzed in terms of the main themes and findings related to the impact of ANNs on consumer behavior. The analysis process involved summarizing each article's key findings and methodological approaches and then categorizing these findings under general themes. Additionally, methodological differences between the articles and existing gaps in the literature were identified and discussed. This approach allowed for a thorough understanding of how ANNs influence consumer behavior and highlighted areas within the current literature that are open to further development..

Findings

According to the SLR study, the articles reviewed were published between 2020 and 2024. The number of publications has shown a noticeable increase over the years. In 2020, 2 articles were published, rising to 3 in 2021. The number of articles published in 2022 reached 4; by 2023, 8 articles had been published. For 2024, as of August, eight articles have already been published. Since 2024 is not yet complete, this number is expected to increase by the end of the year. These data indicate that research in this area has risen steadily each year and that academic interest in this topic rapidly grows over time (See Figure 4).

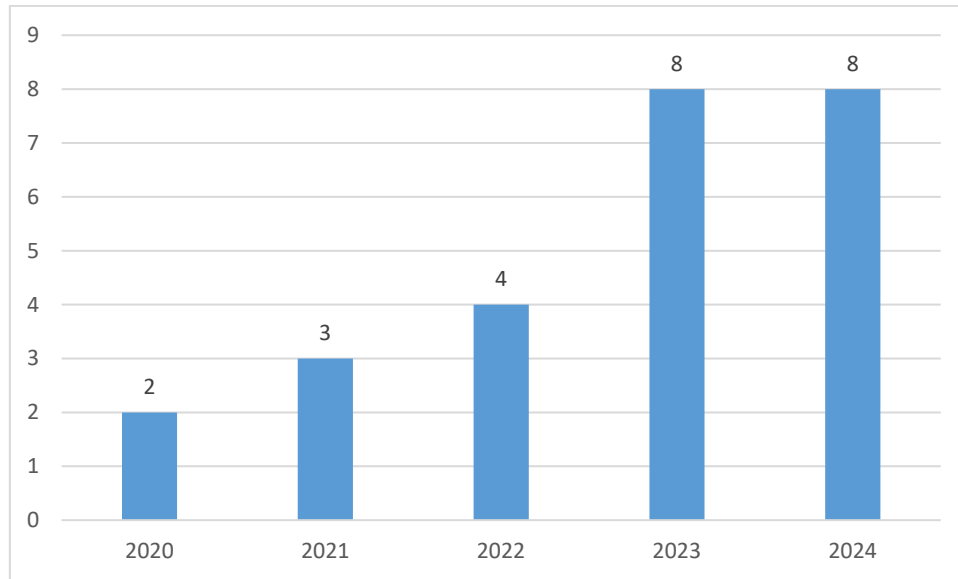


Figure 4. Number of Articles

Among the 25 articles reviewed, empirical studies (20 articles) are predominant. In contrast, literature reviews (4 articles) and case studies (1 article) are fewer in number, indicating a potential gap in terms of theoretical and methodological contributions (See Table 1).

Table 1. Summary of Reviewed Articles: Authors, Study Types, and Titles

Authors	Study Type	Article Title
Zhao et al. (2024)	Empirical	Decoding green food safety information dependency in the digital era: An intelligent validation using SEM-ANN framework
Srivastava and Bag (2024)	Literature Review	Modern-day marketing concepts based on face recognition and neuro-marketing: a review and future research directions
Singh and Ebana-Cabanillas (2024)	Empirical	An SEM-Neural Network Approach for Predicting Antecedents of Online Grocery Shopping Acceptance
Mkedder and Özata (2024)	Empirical	I will buy virtual goods if I like them: a hybrid PLS-SEM-artificial neural network (ANN) analytical approach
Matin et al. (2024)	Empirical	Determinants of Green Smartphone Application Adoption for Sustainable Food Consumption Among University Students
Jin et al. (2024)	Empirical	Forecasting fish prices with an artificial neural network model during the tuna fraud
Chen and Wu (2024)	Empirical	Would you be willing to purchase virtual gifts during esports live streams? Streamer characteristics and cultural traits
Azadravesh et al. (2024)	Empirical	Predicted consumer buying behavior in neural marketing based on convolutional neural network and short-term long-term memory
Zhou et al. (2023)	Empirical	Why do social media users follow tourism-related posts? Roles of bloggers and posts in trip planning

Zambrano-Asanza et al. (2023)	Empirical	Integrating artificial neural networks and cellular automata model for spatial-temporal load forecasting
Ye and Chai (2023)	Empirical	Research and application flow-based live-streaming shopping towards compulsive buying
Storm et al. (2023)	Empirical	Identifying farmers' response to changes in marginal and average subsidies using deep learning
Špicas et al. (2023)	Empirical	Estimating the Acceptance Probabilities of Consumer Loan Offers in an Online Loan Comparison and Brokerage Platform
Ganeshkumar et al. (2023)	Literature Review	Artificial intelligence in agricultural value chain: review and future directions
Arranz et al. (2023)	Empirical	The effect of consumption and production policies on circular economy business models: A machine learning approach
Alaminos et al. (2023)	Empirical	Neural networks for estimating Macro Asset Pricing model in football clubs
Moral-Cuadra et al. (2022)	Empirical	Discovering gastronomic tourists' profiles through artificial neural networks: analysis, opinions and attitudes
Jajić et al. (2022)	Empirical	Deployment of the Microeconomic Consumer Theory in the Artificial Neural Networks Modelling: Case of Organic Food Consumption
Giovanis et al. (2022)	Empirical	Mining the hidden seam of proximity m-payment adoption: A hybrid PLS-artificial neural network analytical approach
Duarte et al. (2022)	Literature Review	Machine Learning and Marketing: A Systematic Literature Review
Sobhanifard and Eshtiaghi (2021)	Empirical	Exploratory modelling and ranking of the trust factors of messages about organic foods in social networks
Senguler and Inel (2021)	Empirical	An Empirical Study Based on Artificial Intelligence for Determining Brand Value Based on Financial Data
Dias et al. (2021)	Empirical	Assessing the Effects of Delivery Attributes on E-Shopping Consumer Behaviour
Peng et al. (2020)	Literature Review	Artificial-Neural-Network-Based Consumer Behavior Prediction: A Survey
Annunziata et al. (2020)	Case Study	Household Food Waste: A Case Study in Southern Italy

Focused Topics

The focused topics include predicting consumer behavior, determining brand value, and developing e-commerce delivery preferences. The potential of ANNs to understand the dynamics and complexities of consumer behavior is particularly emphasized. Studies have shown that ANNs can more comprehensively and accurately model consumer behavior than traditional statistical methods. Additionally, some studies have concentrated on specific application areas such as consumer segmentation, product recommendation systems, and predicting purchase intentions using ANNs. In this context, the contribution of ANNs to making more effective decisions in marketing strategies has been highlighted.

For instance, Zhao et al. (2024) examined the dependency on green food safety information in the digital age and how consumers' psychological perceptions influence their information-seeking behavior and purchase decisions. The study collected data from 630 online users in Guangzhou using PLS-SEM and ANN. It was found that consumers' psychological perceptions significantly affect their information-seeking behavior and purchase decisions regarding green food safety, with reliable information sources playing a critical role in consumer behavior. Using ANNs, the authors could more accurately analyze the relationship between consumers' psychological perceptions, information-seeking behavior, and purchase decisions. The findings provide theoretical and empirical contributions to understanding consumer behavior related to green food safety. The study emphasizes that concerns about green food safety shape information-seeking behavior and purchase decisions, and personalized information delivery based on reliable sources can enhance consumer satisfaction. Future research is recommended to focus on cross-cultural comparisons, long-term effects, and the development of customized strategies to meet the needs of different consumer groups, with a deeper examination of consumer behavior through mixed-methods research.

Srivastava and Bag (2024) explored the potential of facial recognition and neuromarketing in modern marketing strategies, examining data-driven marketing strategies shaped by customer behavior data. The study conducted a systematic literature review on facial recognition marketing and neuromarketing, revealing that the modern market still has not sufficiently explored these areas. The study emphasized that facial recognition and neuromarketing could be significant in understanding consumer behavior and optimizing marketing strategies. ANNs were highlighted as critical in using these advanced marketing techniques more effectively, predicting consumer behavior, and optimizing marketing strategies. The conclusion was that these technologies offer significant opportunities to improve marketing strategies by providing cognitive insights into consumer behavior. Future research should focus on developing predictive models and further exploring the applications of these technologies in marketing strategies.

Singh and Ebana-Cabanillas (2024) examined the factors influencing the acceptance of online grocery shopping in India, such as visibility, economic values, habit, informativeness, and website trust. The study adopted a two-stage PLS-SEM and ANN approach to predict intentions to use online grocery shopping. The findings indicated that habit was the most influential factor affecting user intentions, followed by perceived usefulness, ease of use, and website trust. The ANN model provided flexible and highly accurate results by capturing nonlinear relationships more accurately, supporting the classification of significant predictors identified by SEM. This allowed for more precise predictions of customer behavior and enabled online grocery service providers to tailor their strategies more effectively to customer preferences. Future studies are recommended to cover different demographic regions for broader representation and to enhance the effectiveness of marketing strategies.

In their study, Mkedder and Özata (2024) investigated the functional, emotional, and social values influencing the intention to purchase virtual goods in Free-to-Play (F2P) games. Data from 352 participants were analyzed using PLS-SEM and ANN. The results showed that perceived value positively affects the intention to purchase virtual goods, with quality and social presence being critical factors. ANN supported the PLS-SEM results by capturing nonlinear relationships. The research suggests that F2P game developers focus on social features and high-quality graphics to enhance player experience and increase virtual goods sales.

Additionally, strategies such as time-limited promotions were emphasized as potentially increasing purchase intentions. The study theoretically extended the Theory of Consumption Values in the context of virtual goods and demonstrated the effectiveness of combining PLS-SEM with ANN in understanding consumer behavior. These findings can help F2P game providers optimize their strategies. The study suggests that the model should be tested in different demographics and regions in the future and that the effects of 3D virtual goods and other virtual environments on purchase intentions should be explored.

The study by Matin et al. (2024) investigated the factors influencing the adoption of green smartphone applications for sustainable food consumption among university students. The study employed an integrated model combining the Content Richness Model and the Unified Theory of Acceptance and Use of Technology (UTAUT) with external components such as personalization and environmental concerns. Data collected from 700 participants were analyzed using SEM and ANN, among three machine learning techniques. The results showed that UTAUT, personalization, and environmental concerns positively affected the adoption of green applications. ANN provided the highest accuracy in determining the importance of independent variables, such as performance expectancy, and supported the SEM results. The practical contributions of the study offer significant insights to developers and marketers on how to design and market applications that promote sustainable consumption more effectively. Theoretically, this research contributes to the literature on the factors influencing the adoption of green applications. Although the study's cross-sectional nature, limited to a specific time frame, presents some limitations in generalizing the findings,

it is suggested that further research be conducted in different regions and demographic groups. Additionally, future studies should focus on a more in-depth examination of the role of personalization in adopting green applications.

In their study, Jin et al. (2024) used an ANN model and a threshold vector autoregressive model (TVAR) to predict fish prices during a tuna fraud incident in Spain in 2017. The study found that TVAR offered better short-term predictions for tuna and salmon prices, while ANN performed better in the medium term. The ANN model was particularly effective in capturing nonlinear relationships and hidden structures in the data, making it more accurate in price predictions. This study significantly contributes to the literature on applying machine learning in price prediction during food safety incidents. The practical contributions of the study offer guidance to policymakers and retailers in better understanding market dynamics and consumer behavior during food safety events.

Additionally, the high accuracy rates provided by ANN in agricultural food price prediction present new opportunities for future research and applications. The study's limitations include the inability of social media data to fully represent public engagement and the findings limited to the tuna fraud incident in Spain. Future research should examine food safety incidents in different regions and include additional social media platforms. This study highlights the potential of ANNs in price prediction during food safety events.

In their research, Chen and Wu (2024) examined the factors influencing viewers' intentions to purchase virtual gifts during esports live streams in China. The study investigated how broadcasters' characteristics (expertise, trustworthiness, attractiveness, interaction, and closeness) and Chinese cultural traits shape viewers' social cognition and how this impacts their flow experiences, emotional attachment, and purchase intentions. Viewer behavior was analyzed using SEM, ANN, and Necessity Analysis. The results showed that broadcasters' attractiveness and cultural traits like face consciousness significantly influenced viewers' sense of presence and belonging. ANN provided a higher fit in determining these relationships than SEM. The study recommends developing effective marketing strategies to enhance viewer engagement and promote virtual gift purchases for esports livestream platforms and broadcasters. However, the study notes that the findings must be more generalizable to other cultural contexts due to its focus on Chinese culture. Long-term longitudinal studies are required to track changes in viewer behavior. Additionally, the use of neuroscience technologies to delve deeper into viewers' mental processes in future research is recommended.

In their study, Azadravesh et al. (2024) focused on predicting consumer purchase behavior in neuromarketing using EEG data and machine learning techniques. The study employed a combination of EEG indicators, discrete wavelet transformation (DWT) for feature extraction, and a CNN-LSTM model for classification, which outperformed traditional methods. The proposed CNN-LSTM model accurately predicted consumer preferences by capturing spatial and temporal dependencies from EEG data. The results showed that this model had higher accuracy rates than other classifiers, such as SVM and RF. The study demonstrated that EEG and machine learning techniques could effectively develop neuromarketing strategies, significantly contributing to understanding subconscious consumer behavior and decision-making patterns. Practically, this model has the potential to enhance the accuracy of neuromarketing tools by combining EEG data with biometric indicators such as galvanic skin response and heart rate variability. Future research should focus on improving the data set with a better headset, incorporating eye-tracking, and using more subjects for cross-validation. This study shows that using more advanced machine learning techniques in processing EEG-based neuromarketing data is more effective in predicting consumer behavior than traditional marketing methods.

In their study, Zhou et al. (2023) investigated the factors influencing social media users' intentions to follow tourism-related posts (TRPs) for travel planning. The study identified alignment between bloggers and users and interest in TRPs as critical elements influencing the intention to follow these posts through evaluation and attitude components. The relationships between bloggers' trustworthiness, content quality, attitudes toward posts, and the intention to follow TRPs were analyzed using PLS-SEM and ANN. The findings showed that alignment between bloggers and users and interest in TRPs significantly impacted the intention to follow these posts in travel planning. Strong R^2 coefficients supported these relationships. The study contributed to understanding how social media influences travel behavior and provided practical recommendations for bloggers and destination marketing organizations. The results offer valuable insights to social media platforms and bloggers on developing effective marketing strategies.

In the study by Zambrano-Asanza et al. (2023), a novel method was proposed for spatial-temporal load forecasting by integrating ANN and cellular automata (CA) models. This innovative model uses land use preferences, neighborhood conditions, spatial constraints, and random perturbations to predict consumer and electrical load development in small areas. The study emphasizes the importance of normalization and temporal factors in improving error rates in spatial-temporal load forecasting. The proposed method was validated in an electrical service area in

Ecuador, achieving acceptable metrics with spatial error pattern measurements during the training and testing phases. The research highlights the need for strategic expansion planning and understanding consumer behavior in spatial-temporal load forecasting for long-term distribution planning. The study aims to address deficiencies in spatial error prediction in existing forecasting models and shows how crucial spatial-temporal data analytics is in calibration and validation processes. The findings provide an innovative framework for energy demand forecasting with broad application potential in the energy sector.

Ye and Chai (2023) conducted a comprehensive investigation into customer behavior in live-stream e-commerce, focusing on the effects of flow state factors and post-flow state mediators on compulsive buying. The study, which gathered data from 517 individuals who frequently participated in live-stream e-commerce in China, tested hypotheses using the Stimulus-Organism-Response (SOR) model with ANN and SEM. The research identified trust and pleasure as significant flow factors, while loyalty and addiction were found to be post-flow mediators. The study's emphasis on the effects of trust, pleasure, and flow experience on compulsive buying and loyalty provides both theoretical and practical contributions to understanding consumer behavior in live-stream e-commerce. It sheds light on how social influence, trust, and social capital shape consumer experience, offering important insights into understanding and improving consumer behavior in live-stream e-commerce.

Storm et al. (2023) used deep learning methods to examine how farmers in Norway responded to changes in average and marginal subsidies. The study found that farmers responded more to average payments than marginal payments, challenging standard economic theories. The flexible RNN model and multi-year farm-level data used in the study analyzed how subsidies affected farmer decisions by considering cross-product effects on different agricultural activities. The findings highlight the potential contribution of deep learning tools to policymakers' evaluation and prediction processes of the impact of varying farm policy options. The results show the importance of cross-product effects in policy evaluation by illustrating how subsidies for one activity impact others.

Špicas et al. (2023) focused on developing statistical models to predict the acceptance probabilities of consumer credit offers on online credit comparison and brokerage platforms. The study created predictive models using various classifiers such as logistic regression, random forest, XGboost, ANN, and support vector machines, showing high success in predicting the acceptance probabilities of credit offers. The research emphasizes the importance of considering local regulations and information infrastructures in developing predictive models for online credit comparison platforms. The study's findings indicate that platform-specific features and data diversity significantly predict the acceptance probabilities of credit offers.

Ganeshkumar et al. (2023) presented a literature review examining the use of artificial intelligence (AI) in the agricultural value chain. The study underscores that the adoption and implementation of AI in farming can significantly enhance the efficiency and competitiveness of the industry. This has the potential to address the challenge of feeding a growing population by 2050. The research outlines potential applications of AI in the agricultural value chain (AVC), including the use of deep learning algorithms in areas such as water resource management, yield prediction, and personalized consulting. It suggests that increasing AI integration into the agricultural sector can provide benefits such as income growth, competitiveness, and cost reduction. The study concludes that further research in this area is crucial to fully realize the transformative potential of AI in agriculture.

In the study by Arranz et al. (2023), the effects of circular economy (CE) consumption policies on circular economy business models (CEBM) are examined. The study emphasizes the importance of consumption and production policies in developing CEBMs and highlights their complementary and synergistic effects. Using data from the EU, the article combines classical econometric methods with machine learning approaches, revealing that consumption policies directly impact CEBMs. Furthermore, the study concludes that the combined implementation of consumption and production policies contributes to the development of CEBMs. This research provides valuable insights for policymakers in the context of environmental policies by analyzing the interaction between consumption and production policies and their effects on the development of CEBMs.

Alaminos et al. (2023) present a study that uses ANNs to estimate macro asset pricing models for football clubs. The study highlights the importance of using nonlinear techniques for predicting asset pricing models and achieved over 90% accuracy in predictions with ANN techniques like DNDT and QNN. The research points out the significance of the increasing component, which accounts for approximately 40% of the variance in price-dividend ratios, emphasizing the importance of long-term risks and the habit component. The study aims to improve financial asset pricing predictions in football to prepare investors for unexpected crises like COVID-19 better.

Moral-Cuadra et al. (2022) analyze the profiles of gastronomic tourists using ANNs to examine the relationship between the social characteristics of gastronomic tourists and their views on gastronomic tourism. Using data collected from field research in Córdoba, Spain, the study developed an ANN model to create profiles of gastronomic tourists. The research highlights the importance of traditional gastronomy, gastronomic innovation, and the role of gastronomy in social interactions. The results provide helpful information for public institutions promoting gastronomic tourism and professionals in the gastronomy sector. The study found that as tourists age, their appreciation for diversity, desire to taste on-site, and willingness to return increase, while higher income levels enhance interest in gastronomy. The study concludes that traditional gastronomy significantly enriches the tourist experience, especially among those with higher education levels and cultural interests.

Jajić et al. (2022) use the Lancaster approach of microeconomic consumer theory, modeling it with ANNs to analyze organic food consumption. The study examines the relationship between organic food consumption, demographic characteristics, and health consciousness attitudes of individuals living in Croatia. The research highlights the impact of healthy food stores and consumer awareness of organic food consumption. The findings validate the Lancaster approach and show that consumer-defined attributes are significant in organic food consumption. The study successfully models the relationship between consumer awareness and organic food consumption using ANNs. The results reveal that consumers' awareness of health food stores significantly affects organic food consumption. Additionally, the research emphasizes the influence of the availability of health food stores and demographic factors on organic food consumption. It provides various recommendations to public health authorities for developing healthy eating habits.

Giovanis et al. (2022) extend the Theory of Planned Behavior (TPB) to investigate the adoption of proximity mobile payment services (PMPS). The study employs a two-stage hybrid analytical methodology considering linear and nonlinear relationships. This methodology combines partial least squares (PLS) regression with ANNs to identify the critical determinants of PMPS adoption and rank their relative impact. The research reveals that normative and control beliefs play a significant role in PMPS adoption, while some contradictions related to customer attitudes and behavioral intentions are identified. The findings offer important insights for marketing managers to strengthen strategies for promoting PMPS services.

Duarte et al. (2022) provide a systematic literature review examining adopting machine learning (ML) methods in marketing from 2008 to 2022. The study emphasizes the significant growth and maturation of ML techniques, such as deep learning, supervised learning, reinforcement learning, and hybrid methods, in solving various problems in marketing, including consumer behavior, recommendation systems, forecasting, market segmentation, and text analysis. The research shows that ML techniques have matured and developed in marketing, allowing non-experts to use ML effectively in complex data analysis. The study also suggests that ML should be applied more extensively in new and unexplored areas of marketing, highlighting the importance of correctly applying more straightforward ML techniques.

Sobhanifard and Eshtiaghi (2021) conducted a study to model and rank the trust factors of messages about organic foods on social networks. The study consists of four stages: literature review, hypothesis development, data collection, and ANN analysis. The research identifies 31 factors affecting the trust of messages about organic foods and uses exploratory factor analysis (EFA) and ANNs to rank these factors. The findings suggest that Iranian and international organic food producers can increase consumer trust and sales by focusing on these 31 factors. The EFA-based model reveals six factors that positively affect trust levels: valid experiences, content creation, providing guarantees, informing about product benefits, emphasizing naturalness, and communicating with previous buyers. The study also ranks the top five factors affecting the trust of messages about organic foods and significantly contributes to this field.

In the study by Senguler and İnel (2021), artificial intelligence applications were used to determine brand value based on financial data. The study examined how brand value is determined using financial techniques and consumer behavior-based models, evaluating the effectiveness of ANNs in this process. Various models were created and compared based on R-squared, MAPE (Mean Absolute Percentage Error), and RMSE (Root Mean Square Error) values. The research emphasizes the importance of ANNs in objectively determining brand value. The study's findings show that the brand values obtained using ANNs are consistent with those chosen by the 'Brand Finance' firm. The study contributes to the literature on how financial models can be used to determine brand value and demonstrates the effectiveness of ANNs in this area.

Dias et al. (2021) investigate the impact of delivery features on consumer behavior in e-commerce. This study, conducted on Brazilian consumers, analyzes the effects of delivery time, delivery cost, and delivery method on

consumer decisions. The study evaluates consumer behavior using logistic regression and ANNs and finds that these delivery features significantly influence consumer behavior depending on their sociodemographic characteristics. The research highlights that delivery cost plays a decisive role in the behavior of middle-aged consumers, emphasizing the critical importance of this feature in increasing customer loyalty on e-commerce platforms. Additionally, considering the growing importance of e-commerce during the COVID-19 pandemic, the study discusses the importance of delivery features in acquiring and retaining new customers. The study provides significant findings for developing sustainable delivery strategies in the e-commerce sector and suggests further investigation into the impact of delivery features on consumer behavior.

Peng et al. (2020) examined how consumer behavior prediction (CBP) is performed using ANNs. The research compares artificial intelligence (AI) with statistical systems, emphasizing ANN's superiority in CBP due to its high accuracy and learning capabilities. In this study, which reviewed 1,675 articles, ANN was widely used in predicting consumer behavior, particularly effective in analyzing shopping scan data and credit card expenditures. The research highlights that algorithms like ANN-LM achieved high accuracy rates of 94.7%, showing that ANN generally excels in classifying consumer behavior. The study also addresses challenges in ANN training processes, particularly the issue of forgetting older examples, which can negatively impact results. In conclusion, this study significantly contributes to consumer behavior prediction with AI and ANN applications.

The study by Annunziata et al. (2020) investigates household food waste in southern Italy, aiming to understand consumer behavior and reduce food waste. The study explores different consumer profiles that lead to food waste and the factors affecting this waste. Four consumer clusters were identified and classified based on their behavior. The research reveals that consumer awareness, concerns, and intentions influence food waste. These findings emphasize the need for interventions to be tailored based on consumer behavior to prevent food waste. The study highlights the importance of developing targeted anti-waste strategies at the EU and national levels. Increasing consumer awareness, providing education, and collaborating with stakeholders in the food supply chain can contribute to developing effective food waste reduction strategies.

The Techniques Used

The articles reviewed, covering 2020-2024, focus on studies utilizing ANNs in marketing and consumer behavior. These studies emphasize the potential of ANNs to predict consumer behavior and optimize marketing strategies, employing SEM as a critical technique. In addition, various algorithms such as logistic regression, clustering analyses, and backpropagation are used to enhance the accuracy and performance of ANNs. Furthermore, hybrid models and deep learning techniques have been employed to improve the accuracy and generalization capability of ANNs on complex datasets, highlighting the flexibility and strengths of ANNs in marketing applications. These studies demonstrate that ANNs are an effective tool with a broad range of marketing and consumer behavior applications.

Thematic Areas

The reviewed articles focus on various thematic areas in research conducted using ANN in marketing and consumer behavior. These thematic areas highlight the potential of ANN in enhancing marketing strategies and understanding consumer behavior. One of the primary thematic areas has been predicting and modeling consumer behavior. The studies address ANN's ability to predict consumer purchasing decisions, loyalty tendencies, behaviors toward product preferences, and overall consumer segmentation. Insights gained from analyzing consumer data in this area have enabled the development of strategies to optimize marketing campaigns and increase customer satisfaction.

Another significant thematic area is the personalization of marketing strategies and targeting processes. ANN has been used to create targeted marketing messages and campaigns tailored to consumer segments, resulting in more effective targeting. These studies demonstrate that ANN can be used to develop personalized offers based on individual customer profiles and past purchasing behaviors.

Pricing strategies and revenue management have also emerged as prominent thematic areas in the reviewed articles. ANN has been mainly utilized to develop and optimize dynamic pricing models. The studies have explored how different pricing strategies influence consumer reactions and how these reactions can be predicted using ANN. The

findings in this area have allowed companies to implement flexible and dynamic pricing strategies to maximize revenue.

Additionally, customer loyalty and Customer Lifetime Value (CLV) prediction have been common research topics. ANN has been employed to predict customers' long-term value and develop strategies to enhance loyalty. This thematic area contributes to companies' ability to plan more effective strategies for minimizing customer churn and increasing customer loyalty.

Finally, market segmentation and target audience analysis have also been identified as critical thematic areas. ANN's capability to process large and complex data sets has helped companies more accurately and precisely define market segments. This segmentation has allowed for the organization of more specific and effective campaigns aimed at target audiences to enhance the effectiveness of marketing activities.

These thematic areas provide a broad perspective on how ANN can be utilized in marketing and consumer behavior, contributing to deepening research in this field. The studies reveal that ANN offers significant advantages in critical areas such as customizing marketing strategies, customer loyalty, pricing, and segmentation.

Trends

Artificial Neural Networks (ANN) and Consumer Behavior

Prediction of Consumer Behavior: A significant portion of the studies emphasizes the superiority of ANN in predicting consumer behavior. ANN has been used to forecast consumer purchasing habits, spending tendencies, and product preferences.

Comparison of ANN and Traditional Methods: Many studies reveal that ANN provides higher accuracy and sensitivity than traditional statistical methods. It is noted that conventional techniques like regression and logistic regression have limited capacity in modeling complex consumer behaviors compared to ANN.

Application Areas of ANN

Brand Equity: Some studies have demonstrated the use of ANNs in determining brand equity. ANNs have been used in conjunction with financial data to predict brand equity, and these methods have been found to have high accuracy rates.

Consumer Behavior: ANNs have been utilized to analyze consumer behavior and identify tendencies. These studies indicate that ANNs are an effective consumer segmentation and behavior modeling tool.

Methodological Approaches

Mixed Methods: In the literature reviewed, it was observed that mixed methods (for example, the use of statistical techniques such as ANN and SEM) are a common trend. These methods provide the capacity to model linear and nonlinear relationships, allowing for more in-depth analysis.

Data Variety: The data sets used in the studies usually show a wide variety. Different data sources such as e-commerce data, survey results, financial reports, and social media data have been used to model consumer behavior through ANNs.

Future Research Directions

Dynamic Consumer Segmentation: Future research emphasizes that dynamic consumer segmentation can be conducted more effectively using ANNs. In this context, attention is drawn to the potential of ANNs to capture changing trends in consumer behavior over time.

Sustainability and Ethical Usage: There is a growing recognition of the need to focus more on ethics and sustainability when using ANNs in marketing and consumer behavior analysis. In this regard, the transparency and fairness of ANN-based models are emerging as significant research topics for future studies.

Prediction Performance of ANNs

Accuracy: Studies highlight that ANNs provide high marketing and consumer behavior prediction accuracy rates. However, some studies also note that the performance of ANNs is highly dependent on the dataset used and the model parameters.

Reliability: ANNs can provide reliable predictions when used correctly. However, reliability is directly related to data quality, protection against overfitting, transparency, and generalization ability. Therefore, these factors must be carefully considered when implementing ANN models. Additionally, the reliability of ANN outputs can be enhanced by supporting them with other prediction methods and continuously validating the results.

Advantages and Challenges of ANNs

Advantages of ANNs

High Accuracy and Predictive Power: ANNs can make more accurate predictions than traditional statistical methods, especially because they can process large and complex datasets. Various studies highlight the superior performance of ANNs in predicting consumer behaviors, purchasing tendencies, and the outcomes of marketing strategies.

Ability to Model Complex Relationships: The multi-layered structure of ANNs allows them to model complex and nonlinear relationships between data. This capability enables a more precise understanding and analysis of dynamic and multidimensional processes, such as consumer behavior.

Comprehensive Data Analysis: ANNs can analyze large datasets quickly and efficiently, providing significant advantages, particularly in personalized marketing strategies and customer segmentation. Companies can use ANNs to analyze vast amounts of customer data and extract meaningful and strategic insights from it.

Flexibility and Adaptation: ANNs exhibit flexibility in various marketing and consumer behavior applications. These systems can be adapted to different types of data and problems, enabling them to adapt to constantly changing market conditions quickly.

Learning Capability: ANNs can continuously learn, allowing them to improve their performance over time as they are fed more data. This learning process is crucial for optimizing marketing strategies and detecting changes in customer behavior.

Challenges of ANNs

Model Complexity: The multi-layered and complex structure of ANNs can present significant challenges during model development and optimization. An overly complex model increases the risk of overfitting, negatively impacting the model's general performance on new data.

High Data Requirement: ANNs require large amounts of high-quality data to function effectively. When working with small datasets, the model's predictive power can diminish, making it more challenging to achieve accurate results.

Explainability and Transparency: ANNs are often called a “black box,” which limits the explainability of the model’s decision-making process. In marketing, understanding how and why results are obtained is essential, and this lack of transparency can hinder the adoption of ANNs.

High Computational Cost: ANNs require substantial computational power to process large datasets. This can be a barrier, especially for companies with limited resources. Providing the necessary resources for training and implementing ANNs can be challenging for small and medium-sized enterprises.

Sensitivity to Data Quality: ANNs are highly sensitive to data quality, significantly affecting the model’s training and output quality. Missing, incorrect, or noisy data can negatively impact the accuracy and reliability of the model. Therefore, the preprocessing and cleaning of data present a significant challenge.

These challenges highlight the potential of ANNs in marketing and consumer behavior and the factors that must be considered for their successful implementation. While ANNs offer high accuracy and the ability to model complex relationships, contributing significantly to personalized marketing strategies and customer segmentation, challenges such as model complexity, data requirements, and explainability necessitate a careful approach when applying this technology.

Gaps Identified in the Literature

The 25 articles reviewed significantly contribute to understanding marketing and consumer behavior using ANNs. However, they also reveal existing gaps in the literature and unresolved issues. These gaps present significant opportunities for future research and indicate areas necessary for a deeper understanding of the impact of ANNs on marketing strategies.

Impact of ANNs on Consumer Behavior

While many studies highlight the high accuracy rates of ANNs in predicting consumer behavior, more research is needed on how these techniques can be adapted to all areas of consumer behavior. It has been noted that the performance of ANNs in different demographic and cultural contexts needs to be sufficiently explored.

Explainability and Transparency of ANNs

The "black box" nature of ANNs makes understanding how model outputs are generated difficult. This is a barrier, particularly in marketing, where knowing how and why results are obtained is crucial. The literature reveals a significant research gap in methods to enhance the explainability of ANNs. More studies are needed to focus on making the decision-making processes of ANNs more transparent.

Data Quality and Sufficiency for ANNs

ANNs require large amounts of high-quality data to function effectively. However, it has been noted that model performance can decline when working with small or incomplete datasets. There must be more techniques or strategies in the literature to overcome such challenges. Future research should focus on developing methods to improve the performance of ANNs when working with low-quality or limited datasets.

Applications of ANNs in Different Marketing Strategies

More studies are needed on the impact of ANNs on marketing strategies and how they can be integrated with different marketing techniques. Specifically, more research is required on the applications of ANNs in areas such as digital marketing, customer segmentation, and product recommendation systems. These gaps indicate the need for more studies to guide the integration of ANNs into marketing strategies.

Use of ANNs with Dynamic and Real-Time Data

Many studies focus on the use of ANNs with static data. However, dynamic marketing and consumer behavior may require real-time data analysis. The literature needs more research on how ANNs operate with dynamic and real-time data. Research in this area could enhance the use of ANNs as a more flexible and responsive tool in marketing strategies.

Ethical and Data Privacy Issues of ANNs

Using ANNs to predict consumer behavior raises concerns about data privacy and ethical issues. However, it has been observed that there needs to be more in-depth discussion on these topics in the literature. More research is needed on the ethical challenges and data privacy issues encountered when using ANNs in marketing. Addressing this gap would contribute to developing policies necessary for the safe and ethical application of ANNs.

In summary, the gaps in the literature include the transparency and comprehensibility of ANNs in predicting consumer behavior, the performance differences of ANN models across different demographic groups, the ethical use of ANNs and challenges in data privacy, and the methodological challenges encountered when integrating ANNs with traditional methods. Additionally, more research is needed on the long-term predictive performance of ANNs and their impact on evolving consumer behavior over time. These gaps stand out as important areas for future research.

Discussion

This systematic literature review aims to assess the current state of the literature on ANNs and consumer behavior by comprehensively examining research conducted in the Scopus database over the past five years (2020-2024). The study has obtained significant findings in this field by profoundly exploring the potential of ANNs in predicting marketing and consumer behaviors. It demonstrates that ANNs are a unique tool, particularly in modeling consumer behavior and optimizing marketing strategies. However, specific challenges and limitations must be considered to use this technology effectively.

While this study highlights the wide range of applications that ANNs offer in marketing and consumer behavior, it also thoroughly examines the trends, challenges, and future research areas identified in the current literature. The research conducted by Zhao and colleagues (2024) emphasizes the potential of ANNs to more accurately predict consumer behaviors, while Srivastava and Bag (2024) demonstrate that the integration of innovative marketing techniques, such as facial recognition and neuromarketing with ANNs can enhance the effectiveness of marketing strategies. Singh and Ebana-Cabanillas (2024) provide evidence of ANNs' accuracy in predicting user intentions in

online shopping. Mkedder and Özata (2024) successfully investigated the impact of ANNs on the intention to purchase virtual products.

The findings reveal the high accuracy of ANNs in making predictions in the marketing field while highlighting the challenges encountered in using these techniques. In particular, the study by Chen and Wu (2024) underscores the importance of ANNs working in harmony with variables such as cultural characteristics and audience interactions. Moreover, Jin and colleagues (2024) demonstrate how ANNs contribute to understanding market dynamics during food security events by providing high accuracy rates in predicting agricultural food prices.

However, the "black box" nature of ANNs complicates understanding the model's decision-making processes, which poses a significant barrier in situations where transparency is required in the marketing field. The literature emphasizes the need for more research on methods to enhance the explainability of ANNs (Zhao et al., 2024). Additionally, sensitivity to data quality and high data requirements are other key challenges encountered in ANN applications, necessitating the development of more innovative methods to overcome these challenges (Singh & Ebana-Cabanillas, 2024).

Future research should conduct more comprehensive studies in different demographic and cultural contexts to ensure the more effective integration of ANNs into marketing strategies. It should also focus on in-depth research on ethical and data privacy issues. Furthermore, increasing studies that examine how ANNs work dynamically and in real-time with data could allow this technology to be used as a more flexible and responsive tool in the marketing field (Srivastava & Bag, 2024). In this context, to fully realize the potential of ANNs in marketing, more transparent, understandable, and ethical solutions must be developed.

Theoretical Contributions

The high accuracy rates provided by ANNs in predicting consumer behavior compared to traditional methods have led to this technology gaining significant prominence in marketing literature. The findings from this study demonstrate that ANNs can make more accurate predictions in areas such as consumer segmentation, pricing strategies, customer loyalty, and product recommendation systems compared to traditional statistical methods. This supports the increasing preference for ANNs in the field of marketing.

Moreover, ANNs' ability to model complex and nonlinear relationships within data provides a significant advantage in personalizing marketing strategies and more accurately understanding dynamic consumer behaviors. This allows marketing professionals and researchers better to comprehend consumers' purchasing tendencies and decision-making processes. As a result, the theoretical contributions of ANNs in the marketing field facilitate the expansion of their application areas.

Practical Implications

Integrating ANNs into marketing strategies allows businesses to understand customer behavior better and direct their marketing efforts more effectively. This study found that ANNs perform highly in customer segmentation, dynamic pricing, and customer loyalty analysis applications. Particularly in analyses conducted on large datasets, the ability of ANNs to make fast and accurate predictions helps businesses optimize their marketing strategies more efficiently.

Furthermore, using ANNs in marketing strategies offers significant opportunities for companies to increase customer loyalty and minimize customer attrition. For example, customer loyalty models developed using ANNs can proactively predict customer behavior, enabling businesses to devise strategies that enhance customer satisfaction. These findings demonstrate the practical contributions of ANNs in marketing and support the broader acceptance of this technology.

Conclusion

This study systematically examines how ANNs are utilized in marketing and consumer behavior, providing an in-depth understanding of this area's current state of knowledge. The findings indicate that ANNs are a powerful tool for predicting consumer behavior and optimizing marketing strategies. Thanks to their ability to model complex and nonlinear relationships, ANNs enable the development of more accurate and effective marketing strategies.

However, challenges such as ANNs' "black box" nature, their sensitivity to data quality, and high data requirements limit the widespread adoption of this technology in the marketing field. Therefore, future research should focus on developing methods to enhance ANNs' explainability, improving the quality of datasets, and exploring their applicability in different demographic and cultural contexts.

Additionally, increasing research on the ethical use of ANNs and data privacy issues will contribute to this technology's more reliable and sustainable use in marketing. To fully realize the potential of ANNs, it is crucial to ensure transparency, reliability, and adherence to ethical principles. In this context, expanding the application range of ANNs in marketing and using them as a powerful tool to understand better consumer behavior will be possible through more profound research in this area and the development of practical applications.

Limitations and future research

While this study provides significant findings by systematically examining the applications of ANNs in marketing and consumer behavior, it also has some limitations. First, the study only examines articles from a specific time frame (2020-2024), and this limited period may only partially reflect the long-term effects of ANNs in the marketing field. Additionally, the articles included in the review were selected only from the Scopus database, which means that other essential studies in the literature may have been overlooked.

A second limitation of the study is that it focuses on synthesizing the existing literature rather than providing an in-depth theoretical framework for ANNs' applications in marketing. Therefore, more research is needed on the impact of ANNs on marketing theory and how this technology can be integrated into marketing literature. Furthermore, the study needs to sufficiently examine the technical and methodological challenges associated with using ANNs, which may highlight potential limitations in practical applications.

Several important directions are suggested for future research. First, studies spanning broader time frames should be conducted to understand the longer-term effects of ANNs in marketing. Such studies could reveal the lasting impacts of ANNs on marketing strategies and how they have evolved. Additionally, research should be conducted on how ANNs perform in different demographic and cultural contexts and how this technology can be used more effectively in global markets.

Moreover, research focusing on ANNs' explainability and transparency could contribute to the broader adoption of this technology in marketing. Specifically, new methods and tools should be developed to go beyond ANNs' "black box" nature and enable marketing professionals to better understand this technology. More research is needed on the ethical use of ANNs and data privacy issues. Research could contribute to developing policies and practices necessary for the responsible and sustainable use of ANNs in marketing.

Finally, studies should be conducted on how ANNs can be used in dynamic marketing strategies with real-time data. Closely monitoring and responding to constantly changing consumer behavior in the marketing field is critical to maximizing the potential of ANNs. In this context, integrating ANNs with real-time data analysis could be crucial in achieving higher accuracy and effectiveness in marketing strategies.

References

- Abrardi, L., Cambini, C., & Rondi, L. (2022). Artificial intelligence, firms and consumer behavior: A survey. *Journal of Economic Surveys*, 36(4), 969-991.
- Akdagli, E. (2021). Yapay Sinir Ağları (Artificial Neural Network (ANN)).
<https://eceakdagli.medium.com/yapay-sinir-a%C4%9Flar%C4%B1-artificial-neural-network-annd7846dfcb022>.
 Date of access: 20.08.2024
- Alaminos, D., Esteban, I., & Salas, M. B. (2023). Neural networks for estimating Macro Asset Pricing model in football clubs. *Intelligent Systems in Accounting, Finance and Management*, 30(2), 57-75.
- Albahri, A. S., Alnoor, A., Zaidan, A. A., Albahri, O. S., Hameed, H., Zaidan, B. B., ... & Yass, A.A. (2022). Hybrid artificial neural network and structural equation modelling techniques: a survey. *Complex & Intelligent Systems*, 8(2), 1781-1801.
- Al-Otaibi, Y. D. (2024). Enhancing e-Commerce Strategies: A Deep Learning Framework for Customer Behavior Prediction. *Engineering, Technology & Applied Science Research*, 14(4), 15656-15664.
- Annunziata, A., Agovino, M., Ferraro, A., & Mariani, A. (2020). Household food waste: a case study in Southern Italy. *Sustainability*, 12(4), 1495.
- Arranz, C. F., Kwong, C., & Sena, V. (2023). The effect of consumption and production policies on circular economy business models: A machine learning approach. *Journal of Industrial Ecology*, 27(4), 1089-1104.

- Ashfaq, M., Tandon, A., Zhang, Q., Jabeen, F., & Dhir, A. (2023). Doing good for society! How purchasing green technology stimulates consumers toward green behavior: A structural equation modeling–artificial neural network approach. *Business Strategy and the Environment*, 32(4), 1274–1291.
- Attaoui, M., & Gaber, H. (2024). Online Customer Engagement: A Systematic Literature Review using PRISMA protocol. *Valley International Journal Digital Library*, 6488-6515.
- Azadravesh, H., Sheibani, R., & Forghani, Y. (2024). Predicted consumer buying behavior in neural marketing based on convolutional neural network and short-term long-term memory. *Multimedia Tools and Applications*, 1-17.
- Badea, L. M. (2014). Predicting consumer behavior with artificial neural networks. *Procedia Economics and Finance*, 15, 238-246.
- Boozary, P. (2024). The impact of marketing automation on consumer buying behavior in the digital space via artificial intelligence. *Power System Technology*, 48(1), 1008-1021.
- Borimnejad, V., & Eshraghi Samani, R. (2016). Modeling consumer’s behavior for packed vegetable in “Mayadin management organization of Tehran” using artificial neural network. *Cogent Business & Management*, 3(1), 1208898.
- Briesch, R., & Rajagopal, P. (2010). Neural network applications in consumer behavior. *Journal of Consumer Psychology*, 20(3), 381-389.
- Çağil, G., & Erdem, M. B. (2012). An intelligent simulation model of online consumer behavior. *Journal of Intelligent Manufacturing*, 23, 1015-1022.
- Chattoraj, A. (2024). Trends in Consumer Behavior. *Contemporary Is Accounting*
- Chakraverty, S., Sahoo, D. M., Mahato, N. R., Chakraverty, S., Sahoo, D. M., & Mahato, N. R. (2019). McCulloch–Pitts neural network model. *Concepts of soft computing: fuzzy and ANN with programming*, 167-173.
- Chaudhary, R., Chaudhary, S., Singh, A., Bhanu, A., Chandela, S., & Singh, B. (2024, March). Artificial Intelligence-Based Digital Marketing for Discovering Shopping Possibilities and Enhancing Customer Experience. In *2024 2nd International Conference on Disruptive Technologies (ICDT)* (pp. 413-417). IEEE.
- Chen, J., & Wu, Y. (2024). Would you be willing to purchase virtual gifts during esports live streams? Streamer characteristics and cultural traits. *Computers in Human Behavior*, 152, 108075.
- Chowdhury, P., & Samuel, M. S. (2014). Artificial neural networks: a tool for understanding green consumer behavior. *Marketing Intelligence & Planning*, 32(5), 552-566.
- Dang, T. Q., Tan, G. W. H., Aw, E. C. X., Ooi, K. B., Metri, B., & Dwivedi, Y. K. (2023). How to generate loyalty in mobile payment services? An integrative dual SEM-ANN analysis. *International Journal of Bank Marketing*, 41(6), 1177-1206.
- Dias, E. G., Oliveira, L. K. D., & Isler, C. A. (2021). Assessing the effects of delivery attributes one shopping consumer behaviour. *Sustainability*, 14(1), 13.
- Duarte, V., Zuniga-Jara, S., & Contreras, S. (2022). Machine learning and marketing: A systematic literature review. *IEEE Access*, 10, 93273-93288.
- Fraternali, P., Milani, F., Torres, R. N., & Zangrando, N. (2023). Black-box error diagnosis in Deep Neural Networks for computer vision: a survey of tools. *Neural Computing and Applications*, 35(4), 3041-3062.
- Ganeshkumar, C., Jena, S. K., Sivakumar, A., & Nambirajan, T. (2023). Artificial intelligence in agricultural value chain: review and future directions. *Journal of Agribusiness in Developing and Emerging Economies*, 13(3), 379-398.
- Giovanis, A., Rizomyliotis, I., Konstantoulaki, K., & Magrizos, S. (2022). Mining the hidden seam of proximity m-payment adoption: A hybrid PLS-artificial neural network analytical approach. *European Management Journal*, 40(4), 618-631.
- Greene, M. N., Morgan, P. H., & Foxall, G. R. (2017). NEURAL networks and consumer behavior: NEURAL models, logistic regression, and the behavioral perspective model. *The Behavior Analyst*, 40, 393-418.

- Hajek, P., Barushka, A., & Munk, M. (2020). Fake consumer review detection using deep neural networks integrating word embeddings and emotion mining. *Neural Computing and Applications*, 32(23), 17259-17274.
- Haryanto, J. O., Silva, M., & Moutinho, L. (2015). Neural network approach to understanding the children's market. *European Journal of Marketing*, 49(3/4), 372-397.
- Herrera, M., García-Díaz, J. C., Izquierdo, J., & Pérez-García, R. (2011). Municipal water demand forecasting: tools for intervention time series. *Stochastic Analysis and Applications*, 29(6), 998-1007.
- Jajić, I., Herceg, T., & Pejić Bach, M. (2022). Deployment of the Microeconomic Consumer Theory in the Artificial Neural Networks Modelling: Case of Organic Food Consumption. *Mathematics*, 10(17), 3215.
- Jin, Y., Li, W., & Gil, J. M. (2024). Forecasting Fish Prices with an Artificial Neural Network Model during the Tuna Fraud. *Journal of Agriculture and Food Research*, 101340.
- Kalinić, Z., Marinković, V., Kalinić, L., & Liébana-Cabanillas, F. (2021). Neural network modeling of consumer satisfaction in mobile commerce: An empirical analysis. *Expert Systems with Applications*, 175, 114803.
- Karim, K. S., Islam, M. E., Ibrahim, A. M., Pan, S. H., & Rahman, M. M. (2023). Online marketing trends and purchasing intent: advances in customer satisfaction through PLS-SEM and ANN approach. *Advances in Decision Sciences*, 27(4), 1-31.
- Kasem, M. S., Hamada, M., & Taj-Eddin, I. (2024). Customer profiling, segmentation, and sales prediction using AI in direct marketing. *Neural Computing and Applications*, 36(9), 4995-5005.
- Kazemi, S. M. R., Hadavandi, E., Mehmanpazir, F., & Nakhostin, M. M. (2013). A hybrid intelligent approach for modeling brand choice and constructing a market response simulator. *Knowledge Based Systems*, 40, 101-110.
- Kliestik, T., Zvarikova, K., & Lázároiu, G. (2022). Data-driven machine learning and neural network algorithms in the retailing environment: Consumer engagement, experience, and purchase behaviors. *Economics, Management and Financial Markets*, 17(1), 57-69.
- Kurani, A., Doshi, P., Vakharia, A., & Shah, M. (2023). A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting. *Annals of Data Science*, 10(1), 183-208.
- Lee, C. T., Pan, L. Y., & Hsieh, S. H. (2022). Artificial intelligent chatbots as brand promoters: a two stage structural equation modeling-artificial neural network approach. *Internet Research*, 32(4), 1329-1356.
- Leong, L. Y., Hew, T. S., Ooi, K. B., & Dwivedi, Y. K. (2020). Predicting trust in online advertising with an SEM-artificial neural network approach. *Expert Systems with Applications*, 162, 113849.
- Lin, M., Afshari, A., & Azar, E. (2018). A data-driven analysis of building energy use with emphasis on operation and maintenance: A case study from the UAE. *Journal of Cleaner Production*, 192, 169-178.
- Luckyn, B. J., Alabere, I., & Ogra, O. (2024). Predictive Modeling for Diaper Sales in Retail: An Artificial Neural Network Approach. *Iconic Research and Engineering Journals*, 7(8).
- Ma, C., Wu, J., Sun, H., Zhou, X., & Sun, X. (2023). Enhancing user experience in digital payments: A hybrid approach using SEM and neural networks. *Finance Research Letters*, 58, 104376.
- Matin, A., Khoshtaria, T., Todua, N., Bareja-Wawryszak, O., Pajewski, T., & Todua, N. (2024). Determinants of Green Smartphone Application Adoption for Sustainable Food Consumption Among University Students. *International Journal of Marketing, Communication and New Media*, 11(21).
- Mazurek, M., Hałas, M., Sikora, J., Wiśniewska-Vistula, A., Wróblewska, D., & Zupok, S. (2024). Analysis of consumer behavior using an intelligent multi-source system.
- Mishra, A. K., Bansal, R., Maurya, P. K., Kar, S. K., & Bakshi, P. K. (2023). Predicting the antecedents of consumers' intention toward purchase of mutual funds: A hybrid PLS-SEM-neural network approach. *International Journal of Consumer Studies*, 47(2), 563-587.
- Mkedder, N., & Bakır, M. (2023). A hybrid analysis of consumer preference for domestic products: Combining PLS-SEM and ANN approaches. *Journal of Global Marketing*, 36(5), 372-395.
- Mkedder, N., & Özata, F. Z. (2024). I will buy virtual goods if I like them: a hybrid PLS-SEM-artificial neural network (ANN) analytical approach. *Journal of Marketing Analytics*, 12(1), 42-70.

- Moussa, F. A., & El-Sherbini, N. N. (2024). Prediction of Consumer-Oriented Sales Promotion Technique Using Artificial Intelligence. *Journal of Public Relations Research Middle East/Magallat Bhut Al-Laquat Al-Amh-Al-Srq Al-Aust*, (50).
- Nica, E., Sabie, O. M., Mascu, S., & Luțan, A. G. (2022). Artificial intelligence decision-making in shopping patterns: Consumer values, cognition, and attitudes. *Economics, management and financial markets*, 17(1), 31-43.
- Oztekin, A. (2018). Creating a marketing strategy in healthcare industry: a holistic data analytic approach. *Annals of Operations Research*, 270(1), 361-382.
- Peng, C. C., Wang, Y. Z., & Huang, C. W. (2020, May). Artificial-neural-network-based consumer behavior prediction: a survey. In *2020 IEEE 2nd Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability (ECBIOS)* (pp. 134-136). IEEE.
- Rigby, M., Anthonisen, M., Chua, X. Y., Kaplan, A., Fournier, A. E., & Grütter, P. (2019). Building an artificial neural network with neurons. *AIP Advances*, 9(7).
- Romanenko, N., Sharma, K., & Verma, S. (2024). Prediction of financial customer buying behavior based on machine learning. *Journal of Artificial Intelligence General science (JAIGS) ISSN: 3006 4023*, 5(1), 125-131.
- Salah Alaloul, W., & Hannan Qureshi, A. (2020). *Data Processing Using Artificial Neural Networks*. IntechOpen.
- Sharma, S., Sharma, S., & Athaiya, A. (2017). Activation functions in neural networks. *Towards Data Sci*, 6(12), 310-316.
- Sharma, A., Dwivedi, Y. K., Arya, V., & Siddiqui, M. Q. (2021). Does SMS advertising still have relevance to increase consumer purchase intention? A hybrid PLS-SEM-neural network modelling approach. *Computers in Human Behavior*, 124, 106919.
- Singh, A. K., & Liébana-Cabanillas, F. (2024). An SEM-neural network approach for predicting antecedents of online grocery shopping acceptance. *International Journal of Human-Computer Interaction*, 40(7), 1723-1745.
- Sobhanifard, Y., & Eshtiaghi, K. (2021). Exploratory modelling and ranking of the trust factors of messages about organic foods in social networks. *British Food Journal*, 123(2), 594-609.
- Špicas, R., Neifaltas, A., Kanapickienė, R., Keliuotytė-Staniulienė, G., & Vasiliauskaitė, D. (2023). Estimating the Acceptance Probabilities of Consumer Loan Offers in an Online Loan Comparison and Brokerage Platform. *Risks*, 11(7), 138.
- Srivastava, G., & Bag, S. (2024). Modern-day marketing concepts based on face recognition and neuro marketing: a review and future research directions. *Benchmarking: An International Journal*, 31(2), 410-438.
- Storm, H., Heckelei, T., Baylis, K., & Mittenzwei, K. (2023). Identifying farmers' response to changes in marginal and average subsidies using deep learning. *American Journal of Agricultural Economics*.
- Şengüler, H., & İnel, M. N. (2021). An empirical study based on artificial intelligence for determining brand value based on financial data. *Sosyoekonomi*,(30), 53, 395-424.
- Vinothkumar, K. R., Reka, D. M. P., & Janani, E. S. V. (2024). Predictions Of Consumer Behaviour And Their Impact On Visual Merchandising Using Combined Machine Learning Concept. *Educational Administration: Theory and Practice*, 30(4), 2865-2878.
- Wang, L., Li, X., Zhu, H., & Zhao, Y. (2023). Influencing factors of livestream selling of fresh food based on a push-pull model: A two-stage approach combining structural equation modeling (SEM) and artificial neural network (ANN). *Expert Systems with Applications*, 212, 118799.
- Ye, M., & Ching, T. C. (2023). Research and application flow-based live-streaming shopping towards compulsive buying. *Annals of Operations Research*, 1-29.
- Zambrano-Asanza, S., Morales, R. E., Montalvan, J. A., & Franco, J. F. (2023). Integrating artificial neural networks and cellular automata model for spatial-temporal load forecasting. *International Journal of Electrical Power & Energy Systems*, 148, 108906.
- Zhao, T., Shi, Q., Zhang, X., & Zhang, T. (2024). Decoding green food safety information dependency in the digital era: An intelligent validation using SEM-ANN framework. *Journal of Retailing and Consumer Services*, 79, 103886.

- Zhou, Z., Li, T., Liu, C., Zhou, Y., Li, P., & Wen, S. (2023). Why do social media users follow tourism related posts? Roles of bloggers and posts in trip planning. *Industrial Management & Data Systems*, 123(12), 3080-3108.
- Zou, J., Han, Y., & So, S. S. (2009). Overview of artificial neural networks. *Artificial neural networks: methods and applications*, 14-22.

Comparative Analysis of Monthly Average Rainfall in Amasya Province with Exponential Smoothing Methods

Keziban Kılıç Topal^{1*}, Vedide Rezan Uslu²

¹Department of Educational Sciences, Amasya University, Amasya, Türkiye

² Department of Statistics, Ondokuz Mayıs University, Samsun, Türkiye

Abstract

The decrease in water resources both globally and in our country necessitates the proper use of water resources and the accurate prediction of the parameters that constitute these resources. Rainfall is one of the most important parameters influencing water resources, and its accurate prediction is crucial. Since rainfall data changes over time, it can be defined as a time series. Furthermore, an analysis of rainfall data reveals a seasonal pattern. Therefore, in recent years, seasonal time series analyses have been increasingly used to forecast rainfall data. In this study, the monthly average rainfall in the region of the Gökhöyük Agricultural Enterprise Directorate in Amasya province was predicted using the Holt-Winters exponential smoothing method with double seasonality, and the prediction results were compared using the root mean square error (RMSE). The smoothing parameters and initial values for the Double Seasonal Holt-Winters (DSHW) method were obtained using a Genetic Algorithm (GA). The results show that this approach provides more accurate predictions compared to other methods.

Key words: Rainfall, Rainfall Forecasting, Genetic Algorithm, Time Series, Double Seasonal Holt Winte

INTRODUCTION

In today's world, the impact of climate change has highlighted the importance of effective water resource management. Water, an essential resource for the survival of all living beings, forms the foundation of life. Water resources play a vital role in meeting human needs across agriculture, industry, energy production, and daily life. In agriculture, water required for irrigation is a fundamental element in food production. Additionally, water used in industrial processes supports economic growth. Moreover, water plays a critical role in maintaining the sustainability of natural ecosystems. Preserving and efficiently utilizing water resources is essential to ensure future generations' access to water. The demand for reliable rainfall forecasts is increasing, particularly to ensure water supply and to prevent natural disasters. Rainfall forecasting models have been developed to predict future rainfall events based on historical climate data.

In this study, rainfall forecasting was carried out by combining seasonal time series analysis methods, such as double seasonal Holt-Winters, with Genetic Algorithm (GA) techniques in a hybrid approach. Seasonal methods are highly effective in identifying periodic patterns within the data, and the double seasonal Holt-Winters model, in particular, enhances model accuracy by considering two seasonal cycles in the data. However, the precise determination of model parameters plays a crucial role in the reliability of results. In this context, optimization techniques like Genetic Algorithms serve as powerful tools to enhance prediction performance by identifying the model's optimal parameters. In this study, a hybrid approach that combines the seasonal Holt-Winters model with Genetic Algorithm techniques is proposed for rainfall forecasting. This approach aims to produce more precise and reliable forecasts by optimizing both the periodic patterns and the parameters.

Time series data analysis is used to analyze data that takes into account both time and data characteristics to forecast future events. Exponential smoothing methods, which continuously update forecasts by considering changes in time series data, first emerged through the work of Brown. These methods are known for their three main variations: simple (Brown, 1959), double (Holt, 1957), and the Holt-Winters exponential smoothing method (Winters, 1960). In recent years, the combined use of time series and artificial intelligence methods has produced better results for time series forecasting. Especially in recent years, the Holt-Winters method has been applied for rainfall prediction using time series methods by researchers such as Momani and Naill (2009), Pertiwi (2020), Wiguna et al. (2023), Joseph (2024), and Hendri and Fadhli (2024).

In this study, the Double Seasonal Holt-Winters (DSHW) exponential smoothing method was used to forecast the monthly average rainfall amount in the Gökhöyük Agricultural Enterprise Directorate region in Amasya province. The forecast results were compared using the root mean square error. In the DSHW method, smoothing parameters and initial values were obtained through the use of a Genetic Algorithm (GA). The results showed that the proposed approach yielded more accurate predictions compared to other methods.

Genetic Algorithm (GA) is a search-based optimization technique grounded in the principles of genetics and natural selection. It is frequently used to find optimal or near-optimal solutions to complex problems that are difficult to solve. Utilizing the evolutionary methods found in nature, GA is based on the principle that good generations thrive while poor generations perish. Rather than producing a single solution, it generates clusters of good solutions. This allows for the simultaneous evaluation of multiple points, thereby increasing the likelihood of reaching a holistic solution.

METHOD

In exponential smoothing methods, when there is a trend and only one seasonal behavior in the data, good forecasts are obtained using the Holt-Winters (HW) exponential smoothing method. When a time series exhibits two seasonal patterns simultaneously, Taylor (2003) conducted the first study using the Double Seasonal Holt-Winters (DSHW) method to handle both seasonal movements at the same time.

In the Double Seasonal Holt-Winters (DSHW) method, forecasts are obtained through the simultaneous updating of the following equations. Equation 1 provides the smoothing equation for the average level, Equation 2 is for the trend smoothing, Equation 3 is for the smoothing equations of the first seasonality, and Equation 4 is for the smoothing equation of the second seasonality. The forecast for m periods ahead is given in Equation 5.

$$S_t = \alpha \frac{y_t}{I_{t-L_1} W_{t-L_2}} + (1 - \alpha)(S_{t-1} + b_{t-1}) \quad (1)$$

$$b_t = \gamma(S_t - S_{t-1}) + (1 - \gamma)b_{t-1} \quad (2)$$

$$I_t = \beta \frac{y_t}{S_t W_{t-L_2}} + (1 - \beta)I_{t-L_1} \quad (3)$$

$$W_t = \delta \frac{y_t}{S_t I_{t-L_1}} + (1 - \delta) W_{t-L_2} \quad (4)$$

$$F_{t+m} = (S_t + mb_t)I_{t-L_1+m_1} \times W_{t-L_2+m_2} \quad (5)$$

$$m_1 = [(m - 1) \bmod L_1] + 1 \text{ and } m_2 = [(m - 1) \bmod L_2] + 1$$

The proposed method, defined as DSHW-GA, has been applied to the monthly average rainfall data of the Gökhöyük Agricultural Enterprise Directorate in Amasya Province. The analysis results obtained from HW, DSHW, and DSHW-GA have been compared. The rainfall data represents the average monthly rainfall amount per square meter. Figure 1 shows the time series graph of the data. Upon examining the graph, it can be observed that there is a seasonal behavior, and this seasonality repeats itself monthly and seasonally. The method that yields the smallest RMSE as a performance measure is considered the best solution.

Upon examining the time series graph of the data, it can be seen that there is a seasonal pattern in the data, and this seasonality repeats itself both seasonally and monthly. In this case, the seasonal periods for the application data are as follows: the first seasonal period is daily with a length of 4, and the second seasonal period is weekly with a length of 12.

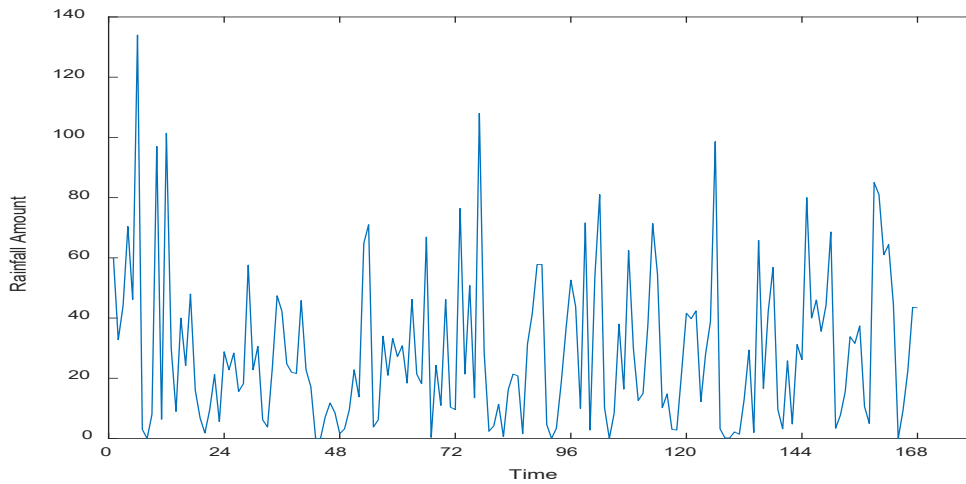


Figure 1. Time series graph of the application data

APPLICATION

The dataset has been analyzed in three separate ways: for one year, two years, and three years. RMSE results for all methods have been compared for the one-year test set, and the results for both the training and test sets are presented. RMSE results for all methods have been compared for the two-year test set, and the results for both the training and test sets are presented. Similarly, RMSE results for all methods have been compared for the three-year test set, with results for the training and test sets provided. Table 1 shows the RMSE results for the training set, while Table 2 shows the RMSE results for the test set. Here, n_t represents the number of data points in the test set. HW(12) indicates that the seasonal period in the Holt-Winters method is set to 12.

Results for the training set:

Table 1. When $n_t=12$, $n_t=24$, $n_t=36$, RMSE results for the training set

	$n_t=12$	$n_t=24$	$n_t=36$
MODEL	RMSE	RMSE	RMSE
DSHW-GA	20.88	22.63	21.79
DSHW	21.02	22.70	22.53
HW (12)	21.40	22.80	22.68

Results for the test set:

Table 2 When $n_t=12$, $n_t=24$, $n_t=36$, RMSE results for the training set

	$n_t=12$	$n_t=24$	$n_t=36$
MODEL	RMSE	RMSE	RMSE
DSHW-GA	24.79	18.68	19.44
DSHW	24.85	19.25	23.89
HW (12)	24.92	22.79	25.97

It can be seen in Table 1 and Table 2 that the method with the lowest RMSE among the compared methods is the proposed method. The graphs provided below are plotted for the DSHW-GA method, which achieved the best results for three different test sets.

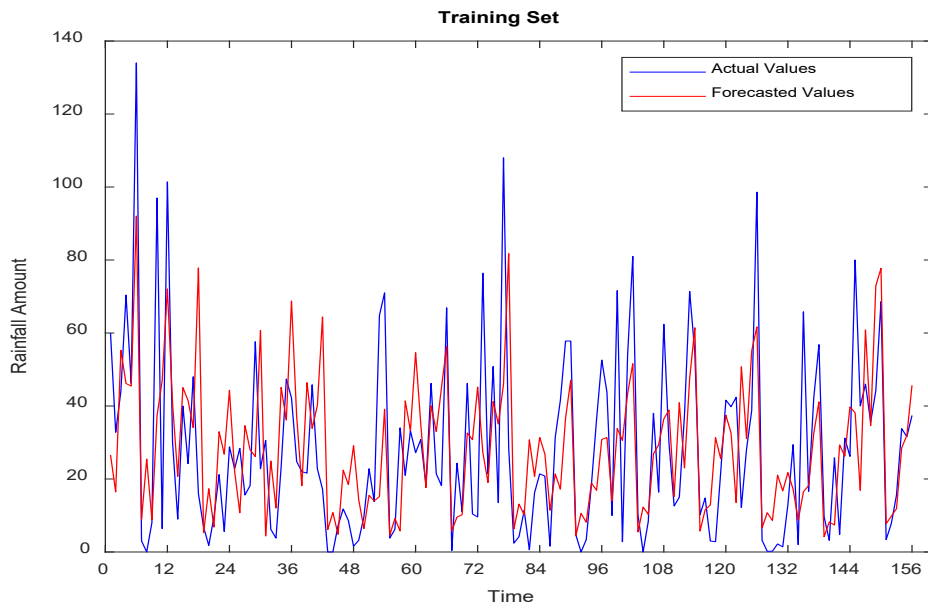


Figure 2. When $n_t=12$, the graph of the training set.

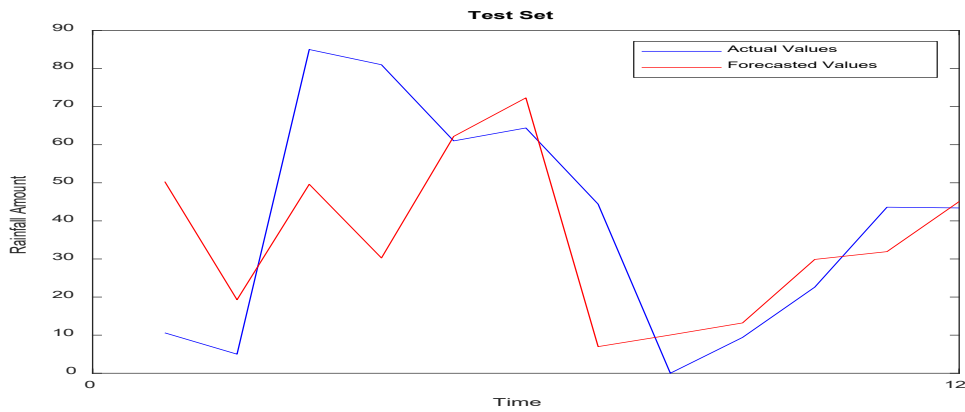


Figure 3. When $n_t = 12$, the graph of the test set

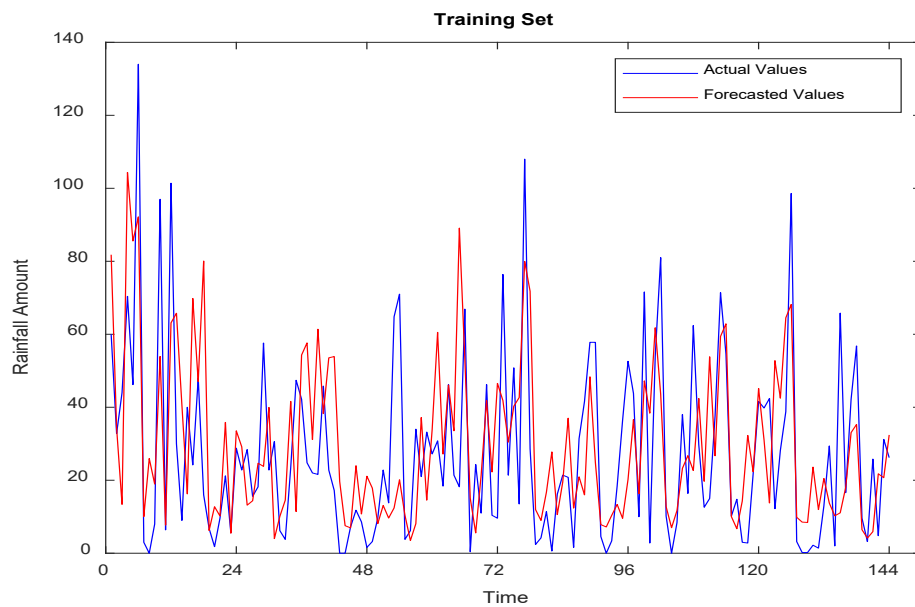


Figure 4. When $n_t = 24$, the graph of the training set

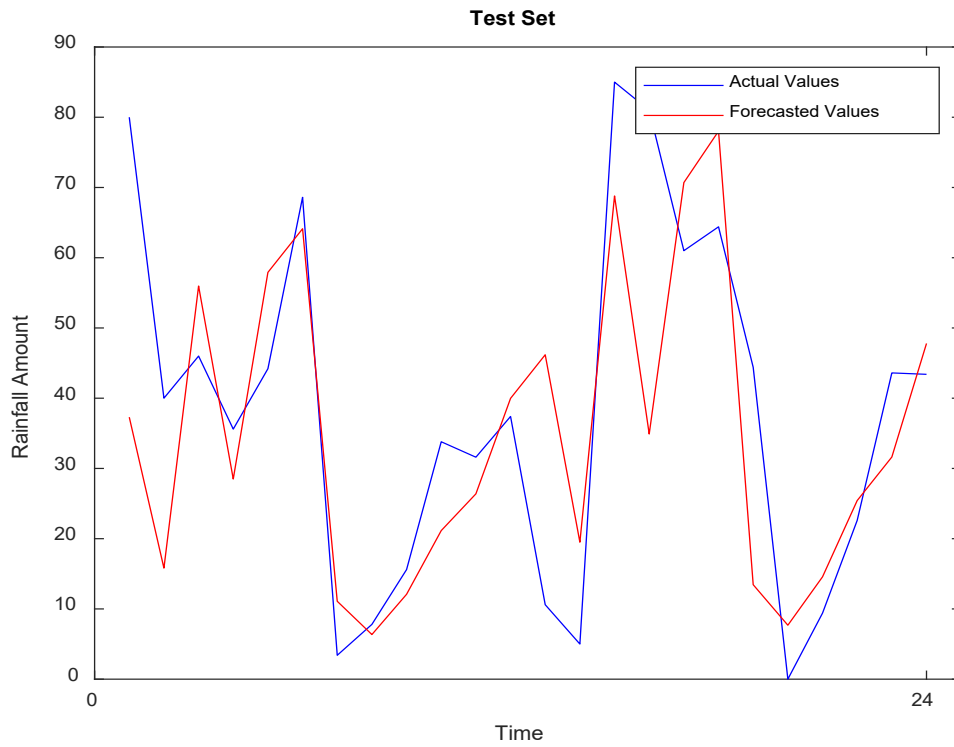


Figure 5. When $n_t = 24$, the graph of the test set

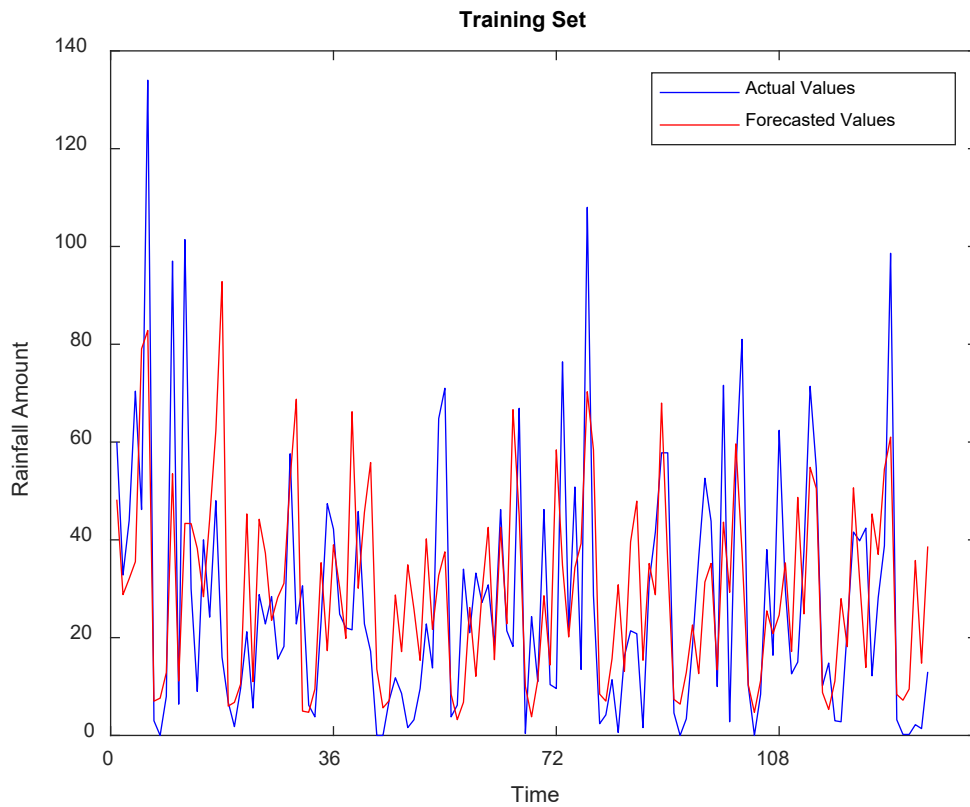


Figure 6. When $n_t = 36$, the graph of the training set

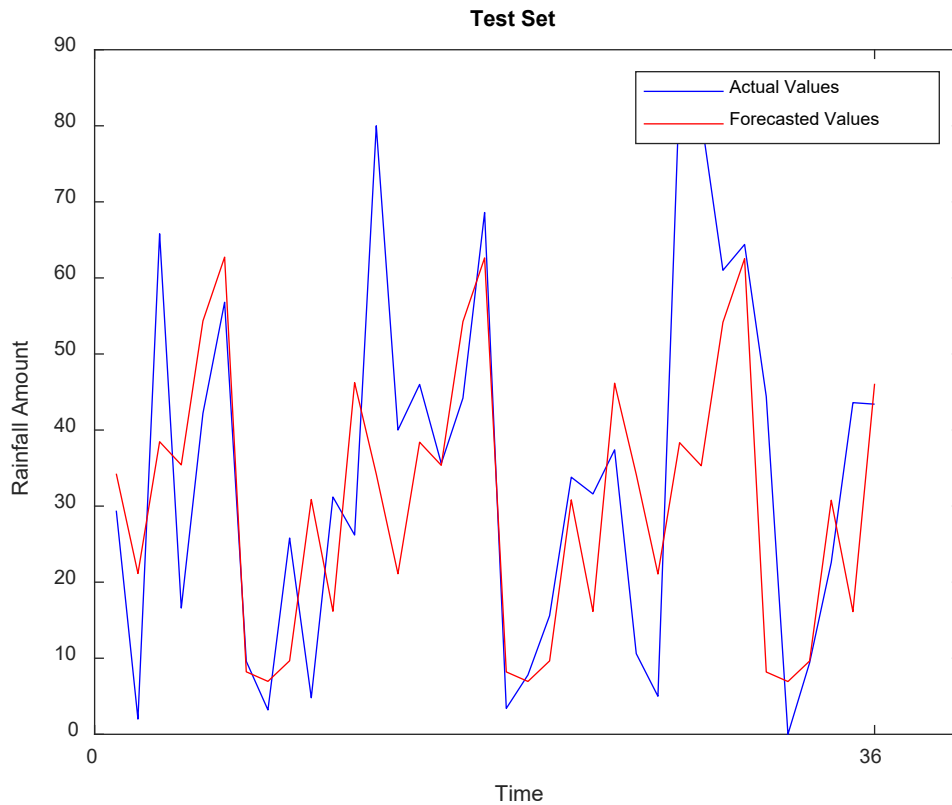


Figure 7. When $n_t = 36$, the graph of the training set

DISCUSSION AND CONCLUSION

In this rainfall prediction study, the proposed DSHW-GA method has achieved the best results compared to other methods. The results indicate that the DSHW-GA method is an effective forecasting tool in terms of both accuracy and overall performance. The predictions that account for seasonal components have provided higher reliability, especially in long-term forecasts. In light of these findings, it is recommended to encourage the use of the DSHW-GA method in future rainfall predictions. Particularly under the influence of dynamic factors such as climate change, adapting and improving this method will contribute to making more precise and reliable forecasts. Additionally, further comprehensive testing and performance enhancement with various datasets will expand its applicability across different climatic regions. In conclusion, it can be said that the DSHW-GA method offers a solid alternative for meteorological forecasting applications and holds potential for further research and development in this field.

REFERENCES

- Brown, R. G., 1959. Statistical forecasting for inventory control. New York McGraw-Hill.
- Hendri, E. P., & Fadhlia, S. 2024. Times series data analysis: The Holt-Winters model for rainfall prediction In West Java. *International Journal of Applied Mathematics, Sciences, and Technology for National Defense*, 2(1), 1-8.
- Holt, C. C. 1957. Forecasting seasonals and trends by exponentially weighted averages. O.N.R. Memorandum 52/1957, Carnegie Institute of Technology. Reprinted with discussion in 2004. *International Journal of Forecasting*, 20, 5 – 13.
- Joseph, Y. A. M. E. O. G. O. 2024. Trends and forecasts of extreme rainfall in small, medium and large towns in Burkina Faso: a non-parametric statistical analysis and Holt-Winters Smoothing Method (West Africa).
- Momani, P. E. N. M., & Naill, P. E. 2009. Time series analysis model for rainfall data in Jordan: Case study for using time series analysis. *American Journal of Environmental Sciences*, 5(5), 599.
- Pertiwi, D. D. 2020. Applied exponential smoothing Holt-Winter method for predict rainfall in Mataram City. *Journal of Intelligent Computing & Health Informatics*, 1(2), 46-49.

Taylor, J. W. 2003. Short-term electricity demand forecasting using double seasonal exponential smoothing. *Journal of Operational Research Society*. 54: 799–805.

Wiguna, I. K. A. G., Utami, N. L. P. A. C., Parwita, W. G. S., Udayana, I. P. A. E. D., & Sudipa, I. G. I. (2023). Rainfall forecasting using the Holt-Winters exponential smoothing method. *Jurnal Info Sains: Informatika dan Sains*, 13(01), 15-23.

Winters, P. R. 1960. Forecasting sales by exponentially weighted moving averages. *Management Science*, 6, 324–342.

SARIMA-YSA Hibrit Model Yaklaşımı: Türkiye’de Turizm Üzerine Bir Uygulama

Kadriye Nurdanay ÖZTÜRK^{1*}, Serpil TÜRKYILMAZ¹

¹İstatistik ve Bilgisayar Bilimleri Bölümü, Bilecik Şeyh Edebali Üniversitesi, Bilecik, Türkiye

Özet

Bu çalışmada, Mevsimsel Otoresif Bütünleşik Hareketli Ortalama (SARIMA) modeli ve Yapay Sinir Ağları (YSA) modeli birleştirilerek hibrit bir SARIMA-YSA modeli tahmin edilmiş, modellerin zaman serisi öngörü performansları karşılaştırılmıştır. Bu amaçla; Türkiye için Ocak-2012 ve Haziran-2024 dönemini kapsayan aylık bazda ziyaretçi sayıları alınarak ilk 120 gözlem eğitim verisi son 30 gözlem ise test verisi olarak kullanılmıştır. Yapay Sinir Ağları (YSA) ve SARIMA-YSA hibrit modelinin öngörü performansları Ortalama Kare Hata (MSE), Ortalama Mutlak Yüzde Hata (MAPE) ve Ortalama Mutlak Hata (MAE) seçim kriterlerine göre değerlendirilmiştir. Çalışma bulguları SARIMA-YSA hibrit modelinin öngörü performansı açısından en uygun model olduğunu göstermiştir.

Anahtar Kelimeler: SARIMA, YSA, Hibrit Model, Zaman Serisi.

1. GİRİŞ

Turizm, küresel ekonomiye sağladığı katkılarla pek çok ülkenin kalkınma stratejilerinde kritik bir rol oynamakta ve hızla gelişim göstermektedir. Ekonomik faydaların yanı sıra uluslararası ilişkileri güçlendirme, kültürel etkileşimi artırma, çevresel sürdürülebilirliği sağlama gibi önemli avantajlar da sunmaktadır. Türkiye, sahip olduğu kültürel ve doğal zenginlikler sayesinde her yıl milyonlarca turisti ağırlamaktadır. Konumu itibarıyla hem Asya hem de Avrupa kıtalarından gelen ziyaretçilere yönelik bir merkez olup, uluslararası turizmin önemli bir ekonomik faaliyet olduğu gelişmekte olan ülkeler arasındadır (Gaberli ve Akdeniz, 2024:299). Özellikle son yıllarda Türkiye’ye gelen turist sayısında düzenli bir artış görülmüş ve turizm sektörü, dış finansman kaynağı olarak ülke ekonomisi için kilit bir rol üstlenmiştir (Karagöz, 2008:154). Bu nedenle turist sayılarının doğru bir şekilde tahmin edilmesi, hem turizm sektöründe faaliyet gösteren kurumlar ve işletmeler hem de ekonomik planlamalar açısından büyük bir öneme sahiptir. Ancak, bu tahminlerin güvenilir olabilmesi için doğru yöntemlerin kullanılması elzemdir. Bu doğrultuda, çalışmamızda üç farklı yöntem ele alınmıştır: SARIMA, YSA ve SARIMA-YSA hibrit modeli.

SARIMA modeli, mevsimsel zaman serisi tahminleri için en etkili doğrusal modellerden biri olarak kabul edilmektedir. Temel avantajı, sınırlı veri setleri ile trendleri, mevsimsel kalıpları ve kısa vadeli ilişkileri etkili bir şekilde yakalayabilmesidir (Deretić vd., 2022:6). YSA, karmaşık ilişkileri öğrenme kabiliyetiyle öne çıkmaktadır. Hibrit model ise, doğrusal özellikleri yakalayan SARIMA ile doğrusal olmayan ilişkileri modelleyen YSA'nın birleştirilmesiyle oluşturularak, hem doğrusal hem de doğrusal olmayan yeteneklere sahip bir model sunar (Moeeni ve Bonakdari, 2017:2002). Bu çalışmanın amacı, SARIMA, YSA ve SARIMA-YSA hibrit modelinin performanslarını karşılaştırarak, turizm sektörü için en uygun tahmin yöntemini belirlemektir. Bu doğrultuda elde edilen sonuçlar, güvenilir turizm tahminleri sağlayarak karar vericilere ve sektör paydaşlarına rehberlik etmeyi amaçlamaktadır.

2. LİTERATÜR

Zaman serisi tahminlerinde hibrit modellerin etkinliği üzerine yapılan çalışmalar önemli bir artış göstermektedir. Díaz-Robles vd. (2008), hava kalitesi tahmini için çoklu doğrusal regresyon, ARIMA, YSA modelleri ile ARIMA ve YSA'yı birleştiren hibrit model kullanmıştır. Bulgular incelendiğinde hibrit modelin hava kalitesini tahmin etmede diğer üç modelden daha başarılı olduğu gözlemlenmiştir. Shafaei vd. (2016) tarafından yağış tahmini için yapılan çalışmada wavelet, SARIMA, YSA ve wavelet-SARIMA-YSA hibrit modeli kullanılmış ve bu yaklaşımların etkinliği test edilmiştir. İran'ın Nahavand meteoroloji istasyonundan elde edilen aylık verilerle gerçekleştirilen analizler sonucunda, wavelet-SARIMA-YSA modelinin yağış tahmininde diğer modellerden daha yüksek bir doğruluk sağladığı tespit edilmiştir. Moeeni ve Bonakdari (2017), Batı İran'daki Jamishan baraj rezervuarına gelen aylık akışı tahmin etmek amacıyla SARIMA ve YSA modellerinin hibrit bir versiyonunu uygulamışlardır. Hibrit model pik taşkın akışlarını daha iyi tahmin ederken, temel akış tahmininde SARIMA modelinin daha iyi performans

gösterdiği bulunmuştur. Sonuçlar, hibrit modelin tahmin hatasını azaltarak daha doğru sonuçlar sunduğunu ve uzun vadede kullanılabilirliğini göstermiştir. Mukaram ve Yusof (2017), Malezya'daki güneş radyasyonu verilerinin tahmini için SARIMA, YSA ve SARIMA-YSA hibrit modeli kullanmış ve çeşitli metrikler ile üç modelin performansını karşılaştırmışlardır. Sonuçlar SARIMA-YSA hibrit modelin güneş radyasyonu verilerini tahmin etmede diğer modellere göre üstün performans sergilediğini göstermiştir. Alencar vd. (2018) çok adımlı rüzgar hızı tahmini için SARIMA ve YSA tabanlı hibrit bir yaklaşım kullanmıştır. Sonuçlar, hibrit modelin yaygın tekniklerle karşılaştırıldığında daha yüksek doğruluk sağladığını göstermiştir. Soares vd. (2020), Sao Paulo'nun farklı bölgelerindeki suç oranlarını tahmin etmek amacıyla YSA, SARIMA, rastgele orman modelleri ile SARIMA-YSA modellerinden oluşan hibrit bir model kullanmış ve tüm sonuçları karşılaştırmıştır. Hibrit model yüksek doğruluk oranları ve düşük kök ortalama kare sapma değerleri ile diğer modellere kıyasla daha başarılı tahminler elde etmiştir. Nontapa vd. (2021), SARIMA ve YSA modellerine dayalı hibrit tahmin modelini Tayland'ın elektrik tüketimi ve SET endeksi verilerine uygulayarak tüm modelleri karşılaştırmış ve daha düşük hata oranları ile üstün tahmin performansı sergileyen hibrit modelin diğer tahmin modellerinden daha doğru sonuçlar üretebileceğini göstermiştir. Azad vd. (2022), Red Hills Rezervuarı'nın su seviyesinin tahmini için SARIMA- YSA hibrit modelini kullanmış ve tüm modelleri karşılaştırarak hibrit modelin hem doğrusal hem de doğrusal olmayan varyasyonları doğru tahmin etme yeteneği sayesinde, diğer modellerden daha iyi performans gösterdiğini ortaya koymuştur. Çağlayan-Akay ve Topal (2024) çalışmalarında, Türkiye'nin elektrik tüketimini tahmin etmek amacıyla tekli ve hibrit modellerin performanslarını karşılaştırmıştır. Elde edilen bulgular, hibrit modellerin tekli modellere kıyasla daha yüksek tahmin doğruluğuna sahip olduğunu, ayrıca hibrit modeller arasında Khashei ve Bijari modelinin en iyi performansı sergilediğini göstermiştir.

Literatür incelendiğinde hibrit modelin özellikle mevsimsel verilerde ve karmaşık dinamikler içeren zaman serilerinde geleneksel yöntemlere kıyasla daha iyi tahmin performansı sağladığı görülmektedir.

3. YÖNTEM

3.1. Hibrit Model

Toplamsal Hibrit Model: Zaman serileri doğrusal ve doğrusal olmayan bileşenleri içerebilir. Bu sebeple hibrit yaklaşımlar doğrusal ve doğrusal olmayan modelleri birlikte kullanarak zaman serisi öngörüsü için daha iyi sonuçlar verebilmektedir. Literatürdeki ARIMA-YSA hibrit modellerde genel yaklaşım şu şekildedir: İlgili zaman serisi için ARIMA modeli tahmin edilir. En uygun ARIMA modelinin hata terimleri elde edilir ve doğrusal olmayan kısım olarak değerlendirilerek bu doğrusal olmayan kısım farklı Yapay Sinir Ağı (YSA) modelleri ile modellenir. Zhang(2003)' de önerilen toplamsal hibrit modelde y_t zaman serisi doğrusal ve doğrusal olmayan bileşenlerin toplamı olarak (1)' deki gibi ifade edilmektedir (Alsuwaylimi, 2023; Erturan ve Merdivenci, 2022; Alrweili ve Fawzy 2022).

$$y_t = L_t + N_t \quad (1)$$

(1) nolu eşitlikte ifade edilen modelde; y_t zaman serisinin L_t ile gösterilen doğrusal kısmı için ARIMA modeli kullanılmaktadır ve elde edilen öngörüler \hat{L}_t için model (2) deki gibi gösterilmektedir.

$$\hat{L}_t = c + \sum_{i=1}^p \varphi_i y'_i + \sum_{j=1}^q \theta_j \varepsilon_{t-j} \quad (2)$$

Zhang (2003)' de y_t zaman serisinden ARIMA modelinin (2)' deki öngörü değerleri çıkartılarak serideki doğrusal olmayan değişimler için bu fark serisi ya da hata serisine odaklanılmıştır.

$$N_t = y_t - \hat{L}_t \quad (3)$$

N_t hata serileri Yapay Sinir Ağı (YSA) yaklaşımı ile elde edilen \hat{N}_t tahmin değerleri

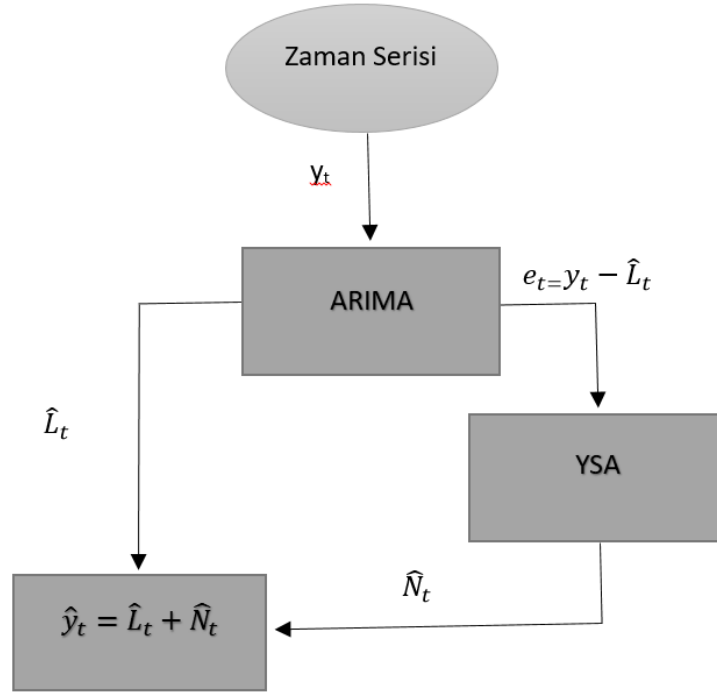
$$\hat{N}_t = f(n_1, n_{t-1}, \dots, n_{t-n}) + v_t \quad (4)$$

Burada \hat{N}_t öngörülen doğrusal olmayan hata serisidir, f doğrusal olmayan fonksiyonudur. v_t ise YSA modelinin beyaz gürültü bileşenidir.

Zhang (2003)' e göre toplamsal hibrit model tahminleri ise ARIMA modeli öngörülerini ile YSA modeli öngörülerinin toplamı olarak elde edilmektedir.

$$\hat{y}_t = \hat{L}_t + \hat{N}_t \quad (5)$$

ARIMA-YSA toplamsal hibrit modeli özet olarak Şekil 1' de gösterilmektedir.



Şekil 1. ARIMA-YSA Hibrit Modeli

3.2. Model Değerlendirme Kriterleri

Çalışmada en uygun modelin seçiminde model performanslarını değerlendirmek için kullanılan hata metrikleri aşağıda verilmiştir:

Ortalama Hata Kare (Mean Squared Error (MSE))

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (6)$$

Kök Ortalama Hata Kare (Root Mean Squared Error (RMSE))

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (7)$$

Mean Absolute Error (Ortalama Mutlak Hata (MAE))

$$MAE = \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{n} \quad (8)$$

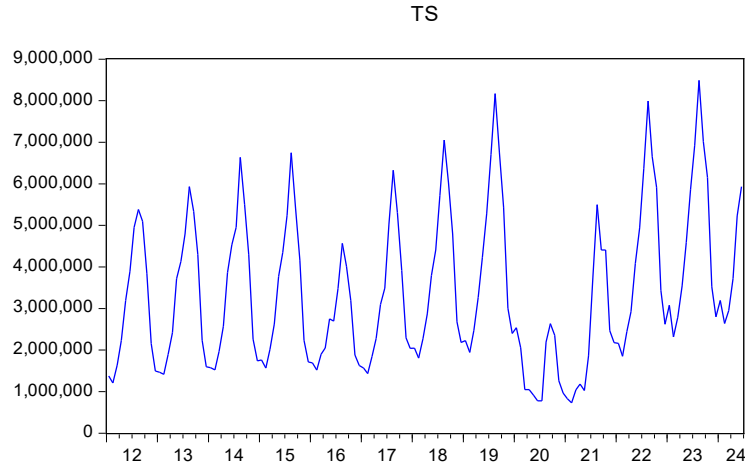
Mean Absolute Percentage Error (Ortalama Mutlak Yüzde Hata(MAPE))

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \times 100, \quad y_i \neq 0 \quad (9)$$

4. VERİ VE MODEL DEĞERLENDİRME

4.1. Veri

Çalışmada kullanılan veri seti TÜİK veri portalı (<https://data.tuik.gov.tr/>) adresinden 2012-2024 yılları arasında aylık bazda Türkiye' ye gelen turist sayıları kullanılmıştır. Toplam 150 gözlemden oluşan çalışma serisi için (TS) grafik Şekil 1' de ve tanımlayıcı istatistikler Tablo 1' de verilmiştir.



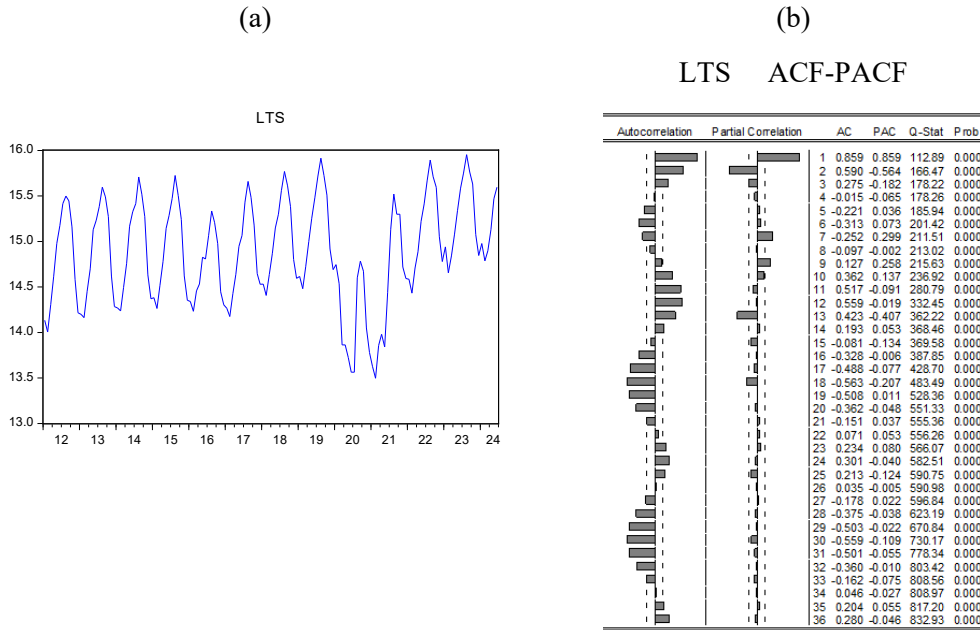
Şekil 2. Aylık Turist Sayıları (TS)' nin Zaman Serisi Grafiği

Şekil 2' de verilen seri grafiğine göre 2012-2024 yılları arasında aylık bazda Türkiye' ye gelen turist sayılarının mevsimsellik sergilediği görülmektedir. 2019 yılı sonu ve 2021 yılı ortaları itibariyle Covid-19 pandemisi dönemindeki düşüş dışında seri benzer eğilim sergilemektedir.

Tablo 1. Tanımlayıcı İstatistikler

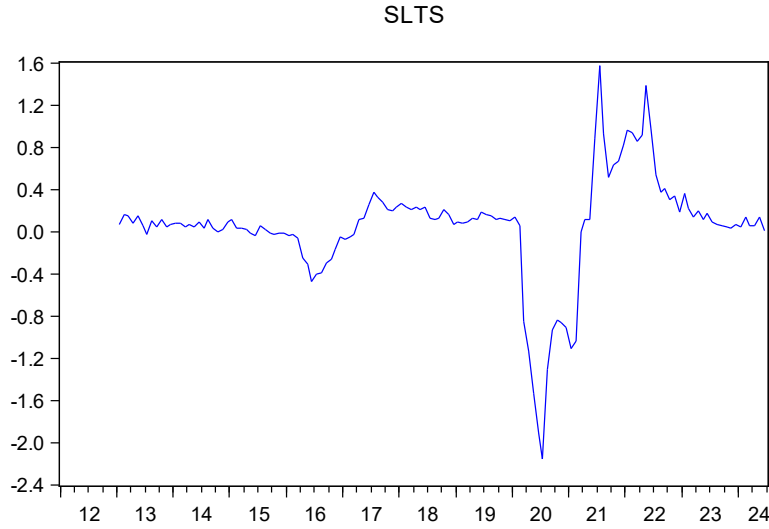
Değişken	Ortalama	Medyan	Mak.	Min.	Std.Sapma	Çarpıklık	Basıklık	Jarque-Bera
TS	3374851	2835425	8488026	727125	1812367	0.683992	2.617084	12.61254(0.002)

Tablo 1' de TS serisine ilişkin tanımlayıcı istatistikler verilmiştir.



Şekil 3. Logaritmik Dönüşümlü (LTS) Serisi (a) ve ACF-PACF Grafiği (b)

Serinin ACF ve PACF fonksiyonu incelendiğinde mevsimsellik sergilediği görülmektedir. Şekil 2' de 12' şerli mevsimlik farkı alınmış LTS serisinin (SLTS) grafiği verilmiştir.



Şekil 4. Mevsimsel Farkı Alınmış (SLTS) Seri Grafiği

Tablo 2’ de LST ve SLTS serileri için Birim Kök Testi Sonuçları verilmektedir.

Tablo 2. Birim Kök Testi Sonuçları

Değişken	ADF	PP	KPSS
LTS	-3.307117(0.0164)**	-4.254455(0.0008)**	0.116206**
Canova Hansen (CH) Mevsimsel Birim Kök Testi Sonuçları			
Joint LM İstatistiği	$L_f=3,000279$		(%5) Kritik Değer 2.750**
Değişken	ADF	PP	KPSS
SLST	-3.189662(0.0016)**	-2.844269(0.0047)**	0.088869**
**, % 5 İstatistiksel anlamlılığı ifade etmektedir, *Kwiatkowski-Phillips-Schmidt-Shin(1992, Tablo 1) %5 kritik değeri (0.463000).			

Canova-Hansen (CH) mevsimsel birim kök testi “mevsimsel frekanslarda durağanlık” sıfır hipotezini temel almaktadır. CH testinde tüm frekanslarda mevsimsel durağanlık hipotezi, ortak test(joint test) ile test edilirken ayrıca her bir frekans için de durağanlık ayrı ayrı test edilebilmektedir. Mevsimsel birim kökün varlığı mevsimselliğin stokastik bir yapıda olduğunu göstermektedir. Bu tür serinin analizinde mevsimsellikten arındırma ya da mevsimsel kukla değişkenler kullanmak yerine mevsimsel fark alınması daha uygun olmaktadır. Literatürde yaygın kullanılan ve aylık verilere uygulanabilen Canova-Hansen (CH) mevsimsel birim kök testi için L_f istatistiği bulgularına Tablo 2’ de yer verilmiştir.

Buna göre $L_f=3,000279$ değeri %5 anlamlılık düzeyinde 2.750 kritik değerinden büyük olduğu için “ H_0 : Mevsimsel Birim Kök Yoktur” hipotezi reddedilmektedir. Mevsimselliğin stokastik bir yapıda olduğu görülmektedir. Bu nedenle serinin 12’ şerli mevsimlik farkları alınarak mevsimselliği giderilmiş seriye uygulanan ADF, PP ve KPSS birim kök testleri sonuçları Tablo 2’ de gösterilmiştir. Bulgulara göre serinin durağan olduğu görülmektedir.

4.2. SARIMA Modeli

Tablo 2’ ye göre SLTS serisinin durağan olduğu görülmektedir. Durağan SLTS serisi için önerilen farklı P ve Q gecikmelerinde, mevsimsel ARIMA modelleri tahmin edilmiş ve AIC (Akaike information criterion), Schwarz information criterion (SIC), Hannan–Quinn information criterion (HQ)’ ne göre model karşılaştırma sonuçları Tablo 3’ de verilmiştir.

Tablo 3. SARIMA Model Tahminleri

Model	R ²	AIC	SIC	HQ
SARIMA(1,0,1)(1,1,0) ₁₂	0,882	-0,5770	-0,4925	-0,5429
SARIMA(1,0,1)(0,1,1)₁₂	0,912	-0,7852	-0,7004	-0,7508
SARIMA(1,0,0)(0,1,1) ₁₂	0,906	-0,7287	-0,6651	-0,7029
SARIMA(2,0,0)(0,1,1) ₁₂	0,777	0,1227	0,1863	0,1486
SARIMA(3,0,0)(0,1,1) ₁₂	0,662	0,5208	0,5840	0,5466
SARIMA(0,0,3)(0,1,1) ₁₂	0,611	0,7066	0,7703	0,7324
SARIMA(2,0,2)(0,1,1) ₁₂	0,786	0,1154	0,2003	0,1499
SARIMA(3,0,3)(0,1,1) ₁₂	0,667	0,5258	0,6107	0,5603
SARIMA(3,0,3)(1,1,0) ₁₂	0,567	0,7176	0,8024	0,7521
SARIMA(2,0,1)(1,1,0) ₁₂	0,873	-0,5000	-0,4152	-0,4656

Tablo 3' deki model seçim kriterlerine göre en uygun model SARIMA(1,0,1)(0,1,1)₁₂ modeli seçilmiştir.

Modelin genel ifadesi:

$$\varphi(B)\nabla_{12}^1 y_t = \Theta_1(B^{12})\theta(B)\varepsilon_t \quad (10)$$

$$(1 - \varphi_1 B)(1 - B^{12})y_t = (1 + \Theta_1 B^{12})(1 + \theta_1 B)\varepsilon_t \quad (11)$$

$$y_t = \varphi_1 y_{t-1} + y_{t-12} - \varphi_1 y_{t-13} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_1 \varepsilon_{t-12} + \Theta_1 \theta_1 \varepsilon_{t-13} \quad (12)$$

olarak yazılabilir.

Model tahmin sonuçları ise Tablo 4' de verilmiştir.

Tablo 4. SARIMA(1,0,1)(0,1,1)₁₂ Modeli Tahmin Sonuçları

Değişken	Katsayı	Std. Hata	p-Değ.
AR(1)	0,901614	0,021992	0,0000**
MA(1)	0,270026	0,066590	0,0001**
SMA(12)	-0,852587	0,074588	0,0000**

** , %5 istatistiksel anlamlılığı ifade etmektedir.

SARIMA (1,0,1)(0,1,1)₁₂ modeli ile turist sayılarının öngörü değerleri Tablo 5' de verilmiştir.

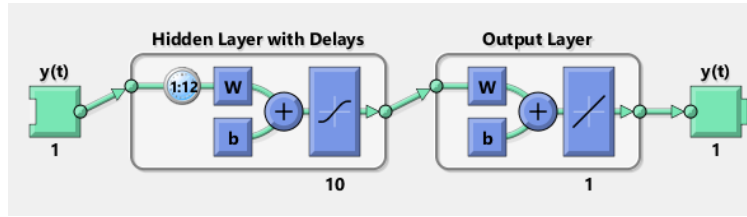
Tablo 5. SARIMA(1,0,1)(0,1,1)₁₂ Modeli ile Tahmin Değerleri

Tarih	TS-Tahmin	TS-Gerçek
2023M06	5266807	5869766
2023M07	7429484	6926346
2023M08	9045256	8488026
2023M09	7031830	7013861
2023M10	5670898	6151397
2023M11	3369848	3520467
2023M12	2719623	2800281
2024M01	2728917	3194819
2024M02	2716789	2640392
2024M03	2941994	2949590
2024M04	3419674	3720878
2024M05	4502524	5224931
2024M06	6156515	5929709

4.3. YSA Modeli

Doğrusal olmayan otoregresif modeller (NAR) farklı geçmiş dönem değerlerini kullanarak gelecek dönem değerlerini tahmin etmektedir. Doğrusal olmayan bir zaman serisi için oldukça iyi uyum yeteneğine sahip olan ve üç katmandan (Giriş, Gizli ve Çıkış) oluşan NAR sinir ağı çalışmada turist sayısı tahmini için kullanılmıştır. Modelde gecikmeli aylık turist sayıları girdi değişkeni olarak modele girilmiş ve NAR ağı MATLAB programı nstool modülü ile oluşturulmuştur. Verinin %80' i eğitim, %10' u doğrulama ve %10' u test için ağı tanıtılmıştır ve eğitim algoritması Levenberg-Marquardt seçilmiştir. 12 gecikmeli aylık turist sayıları için 1 Girdi katmanı, 10 nöronlu bir gizli katman ve mevcut turist sayılarının oluşturduğu 1 Çıktı katmanı ile NAR(1:10:1) modeli tahmin edilmiştir.

Şekil 5, aylık turist sayıları için Yapay Sinir Ağı NAR ağ yapısını göstermektedir.



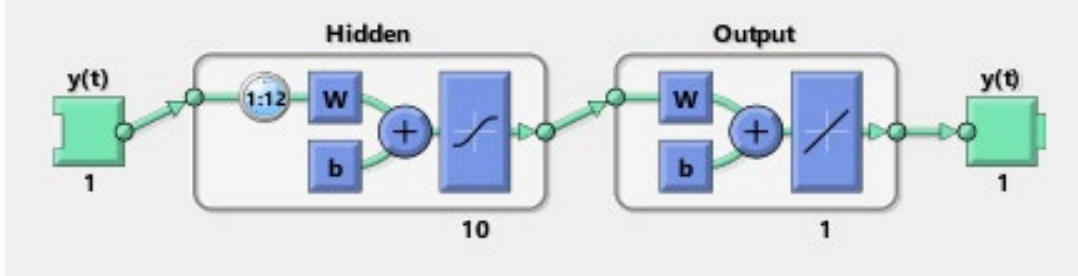
Şekil 5. NAR Sinir Ağı Modeli Yapısı

Tablo 6. YSA Modeli ile Tahmin Değerleri

Tarih	TS-Tahmin	TS-Gerçek
2023M06	6032057,47	5869766
2023M07	5817315,463	6926346
2023M08	6452369,996	8488026
2023M09	6272140,737	7013861
2023M10	4741767,491	6151397
2023M11	3094400,367	3520467
2023M12	2785457,317	2800281
2024M01	2281031,148	3194819
2024M02	1910906,16	2640392
2024M03	2489668,187	2949590
2024M04	1688795,663	3720878
2024M05	2513810,138	5224931
2024M06	2893252,893	5929709

4.4. SARIMA-YSA Hibrit Modeli

SARIMA-YSA hibrit modeli tahmininde; aylık turist sayıları için tahmin edilen en uygun SARIMA (1,0,1)(0,1,1)₁₂ modelinin hata terimleri 3 katmanlı Yapay Sinir Ağı NAR modeli için girdi olarak kullanılmıştır. Bu hibrit modelde de gecikme sayısı 12, gizli katman nöron sayısı 10' dur. Modelin ağ yapısı benzer şekilde Şekil 4' de verilmiştir.



Şekil 6. NAR Sinir Ağı Model Yapısı

SARIMA-YSA hibrit modeli tahminleri, girdi olarak kullanılan SARIMA (1,0,1)(0,1,1)₁₂ modelinin hata terimleri tahminleridir. Bu nedenle denklem (5) ile verilen Zhang (2003)' in toplamsal hibrit modeli yaklaşımına göre turist sayısı tahmin değerleri; SARIMA-YSA hibrit modeli tahminleri ile SARIMA (1,0,1)(0,1,1)₁₂ modeli tahminleri toplanarak elde edilmiştir. SARIMA-YSA Hibrit Modeli' ne göre turist sayısı tahmin değerleri Tablo 7' de verilmiştir.

Tablo 7. SARIMA-YSA Hibrit Modeli ile Tahmin Değerleri

Tarih	TS-Tahmin	TS-Gerçek
2023M06	5266807,057	5869766
2023M07	7429483,910	6926346
2023M08	9045256,035	8488026
2023M09	7031829,660	7013861
2023M10	5670897,687	6151397
2023M11	3369848,500	3520467
2023M12	2719622,700	2800281
2024M01	2728916,900	3194819
2024M02	2716789,272	2640392
2024M03	2941993,611	2949590
2024M04	3419673,715	3720878
2024M05	4502524,290	5224931
2024M06	6156514,692	5929709

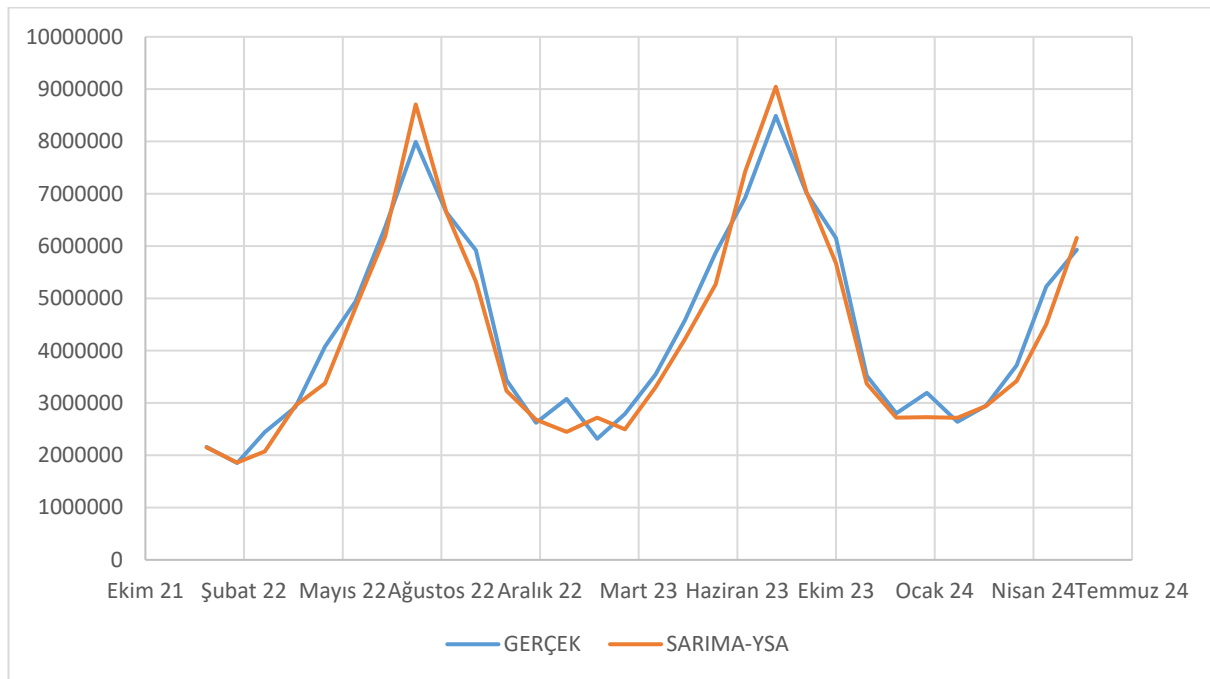
4.5. Model Performans Karşılaştırması

Çalışmada kullanılan modellerin performans karşılaştırmalarına ilişkin göstergeler ise Tablo 8' de verilmiştir.

Tablo 8. Model Performans Karşılaştırması

MODEL	MSE	RMSE	MAE	MAPE
SARIMA	2614950851917,44	1617080,9664	1508036,8349	35,9388
YSA	15896642791671,60	3987059,4166	3656609,2149	85,0792
SARIMA-YSA	150094859075,05	387420,7778	303848,6522	7,0860

Tablo 8’ de çalışma verileri için tahmin edilen SARIMA modeli, YSA modeli ve SARIMA-YSA Hibrit modeli öngörü performansı için elde edilen MSE, RMSE, MAE ve MAPE kriterlerine göre hibrit modelin diğer modellere göre oldukça düşük hata değerlerine sahip olduğunu ve turist sayıları serisinin yapısını diğer modellere göre daha iyi yakaladığını söylemek mümkündür. Şekil 7’ de verilen SARIMA-YSA hibrit modelin öngörü değerleri-gerçek değer karşılaştırmasına ilişkin grafikten de uyum görülmektedir.



Şekil 7. SARIMA-YSA Hibrit Modelin Öngörü Değerleri-Gerçek Değer Karşılaştırması

5. SONUÇ VE ÖNERİLER

Bu çalışmada, 2012-2024 yılları arasında aylık frekansta Türkiye’ ye gelen turist sayıları için bir hibrit model önerilerek mevsimsel ARIMA(SARIMA), yapay sinir ağı modeli (YSA) ve SARIMA-YSA hibrit modellerinin tahmin performansları hata metrikleri ile karşılaştırılmıştır. Hibrit model için Zhang(2003)’ ün toplamsal hibrit model yaklaşımı değerlendirilmiştir. Elde edilen sonuçlar SARIMA-YSA hibrit modelinin diğer modellere kıyasla daha düşük hata değerleri ile turist sayıları serisinin yapısını daha iyi yakaladığını ortaya koymuştur. Sonuçlar, turizm sektöründe ileriye dönük planlamalarda hibrit modellerin etkili bir şekilde kullanılabileceğini göstermektedir. Çalışmanın geliştirilebilir yönü olarak devamı için farklı hibrit yaklaşımlar olan Çarpımsal Hibrit Yaklaşım, Khashei ve Bijari Hibrit (2010)’ un hibrit yaklaşımı gb. ve makine öğrenmesi, derin öğrenme yöntemleri ile hibrit yaklaşımlar gibi farklı tahmin yöntemlerinin performanslarının karşılaştırılması hedeflenmektedir.

Referanslar

Alencar, D. B., Affonso, C. M., Oliveira, R. C., Jose Filho, C. R. 2018. Hybrid approach combining SARIMA and neural networks for multi-step ahead wind speed forecasting in Brazil. IEEE Access, 6, 55986-55994.

- Alrweili, H., Fawzy, H. 2022. Forecasting crude oil prices using an ARIMA-ANN hybrid model. *J Stat Appl Probab*, 11(3), 845-855.
- Alsawaylimi, A. A. 2023. Comparison of ARIMA, ANN and Hybrid ARIMA-ANN models for time series forecasting. *Inf. Sci. Lett*, 12(2), 1003-1016.
- Azad, A. S., Sokkalingam, R., Daud, H., Adhikary, S. K., Khurshid, H., Mazlan, S. N. A., Rabbani, M. B. A. 2022. Water level prediction through hybrid SARIMA and ANN models based on time series analysis: Red hills reservoir case study. *Sustainability*, 14(3), 1843.
- Çağlayan-Akay, E., Topal, K. H. 2024. Forecasting Turkish electricity consumption: A critical analysis of single and hybrid models. *Energy*, 305, 132115.
- Deretić, N., Stanimirović, D., Awadh, M. A., Vujanović, N., Djukić, A. 2022. SARIMA modelling approach for forecasting of traffic accidents. *Sustainability*, 14(8), 4403.
- Díaz-Robles, L. A., Ortega, J. C., Fu, J. S., Reed, G. D., Chow, J. C., Watson, J. G., Moncada-Herrera, J. A. 2008. A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile. *Atmospheric Environment*, 42(35), 8331-8340.
- Erturan, M., Merdivenci, F. 2022. Zaman serileri analizi için optimize ARIMA-YSA melez modeli. *Journal of the Faculty of Engineering and Architecture of Gazi University*, 37(2).
- Gaberli, Ü., Akdeniz, A. 2024. Türkiye’de Uluslararası Turizm Talebinin Belirleyicileri 2002-2022 Dönemi Panel Veri Analizi. *Trakya Üniversitesi Sosyal Bilimler Dergisi*, 26(1), 297-316.
- Karagöz, K. 2008. Türkiye'nin turizm potansiyeli: çekim modeli yaklaşımı. *Anatolia: Turizm Araştırmaları Dergisi*, 19(2), 149-156.
- Moeeni, H., Bonakdari, H. 2017. Forecasting monthly inflow with extreme seasonal variation using the hybrid SARIMA-ANN model. *Stochastic environmental research and risk assessment*, 31, 1997-2010.
- Mukaram, M. Z., Yusof, F. 2017. Solar radiation forecast using hybrid SARIMA and ANN model: A case study at several locations in Peninsular Malaysia. *Malays J Fundam Appl Sci Spec Issue Some Adv Ind Appl Math*, 13, 346-350.
- Nontapa, C., Kesamoon, C., Kaewhawong, N., Intrapai boon, P. 2021. A Comparative of a New Hybrid Based on Neural Networks and SARIMA Models for Time Series Forecasting. In *International Conference on Multi-disciplinary Trends in Artificial Intelligence* (pp. 94-105). Cham: Springer International Publishing.
- Shafaei, M., Adamowski, J., Fakheri-Fard, A., Dinpashoh, Y., Adamowski, K. 2016. A wavelet-SARIMA-ANN hybrid model for precipitation forecasting. *Journal of Water and Land Development*, 28(1), 27.
- Soares, F. A., Silveira, T. B., Freitas, H. C. 2020. Hybrid Approach based on SARIMA and Artificial Neural Networks for Knowledge Discovery Applied to Crime Rates Prediction. In *ICEIS (1)* (pp. 407-415).

Analyzing and Improving the Impact of Software Changes on IT Operations: A Process Mining and Forecasting Approach

Ayşegül Kaçar*¹, Erkan Işıklı²

¹Endüstri Mühendisliği Bölümü/ Lisansüstü Eğitim Enstitüsü, İstanbul Teknik Üniversitesi, İstanbul, Türkiye

²Endüstri Mühendisliği Bölümü/ İşletme Fakültesi, İstanbul Teknik Üniversitesi, İstanbul, Türkiye

Abstract

This study presents a novel methodology that integrates process mining with time series forecasting to analyze and enhance the impact of software changes on IT Operations and Service Desk processes. Employing process mining techniques, we extracted key features from change, incident, and interaction logs to uncover relationships and identify critical points within these processes. These features served as exogenous variables in an ARIMAX model, enabling us to accurately predict operational metrics such as incident frequency and interaction counts. The enhanced model achieved a significant improvement in predictive accuracy, reducing the Mean Absolute Percentage Error (MAPE) to approximately 12% for both incidents and interactions. The forecasting results provide decision-makers with reliable, data-driven insights that aid in optimizing process improvements and resource allocation. Enabling early identification of potential risks associated with software changes, this methodology supports proactive intervention planning, thereby enhancing operational efficiency and service quality. The flexible and data-driven nature of this approach demonstrates broad applicability for process optimization across various sectors, including banking, healthcare, and manufacturing.

Keywords: Software changes, Process mining, Forecasting, Operational efficiency

INTRODUCTION

In modern IT environments, frequent software changes can lead to incidents and interactions that do not appear immediately but emerge after a delay, typically over the following week (Rahman & Devanbu, 2013). Predicting these future impacts based on past processes is essential for IT managers, as it allows them to allocate resources efficiently and mitigate potential risks. In the context of ITIL (Information Technology Infrastructure Library), an *incident* is defined as any unplanned disruption to an IT service or a reduction in its quality (van Bon, 2007). The process begins when a client encounters a problem with an ICT service and interacts with the service desk (Iden & Eikebrokk, 2013). In most cases, the service desk resolves the issue during the initial interaction. However, if the problem requires specialized knowledge, an incident is formally recorded, necessitating expert intervention (Potgieter, Botha, & Lew, 2005). Repeated occurrences of the same issue lead to a problem and improvement analysis aimed at identifying the root cause and preventing future disruptions (Mesquida et al., 2012). Following this analysis, a change process is initiated to modify the configuration item responsible for the problem (Pollard & Cater-Steel, 2009). This structured workflow—connecting interaction management, incident management, and change management—ensures a continuous improvement cycle within IT service delivery (see Figure 1).

Process mining has emerged as a significant tool in today's business environment, demonstrating substantial potential for creating business value by enabling organizations to analyze and visualize process data. By increasing transparency through data-driven decisions, process mining enhances process efficiency, reduces costs, and improves customer satisfaction (Badakhshan et al., 2022). In the healthcare sector, for instance, it contributes significantly by analyzing processes, ensuring compliance, managing resources, and reducing costs, thereby improving patient experiences and enabling evidence-based process improvements (Muñoz-Gama et al., 2022). The applicability and benefits of process mining in software development processes have also been demonstrated through case studies conducted in SMEs, revealing that these techniques are effective tools for process improvement (Urrea-Contreras et al., 2021; Urrea-Contreras et al., 2024). Despite extensive literature on process mining applications, gaps remain regarding the integration of these techniques with prediction and simulation methods (vom Brocke et al., 2021; Martin et al., 2021).

This study aims to address this gap by developing a predictive model that accounts for the delay between process-generated features and the subsequent emergence of incidents and interactions. By applying process mining techniques to change, incident, and interaction logs, we extract key features from the processes occurring within a specific timeframe to reflect how they influence operational outcomes in the following period. The goal is to forecast future incidents and interactions using an Autoregressive Integrated Moving Average with Exogenous variables

(ARIMAX) model enriched with features derived from process mining. This approach allows for more accurate predictions of future process performance and more effective planning of process improvement strategies, thereby contributing to more efficient IT service management.

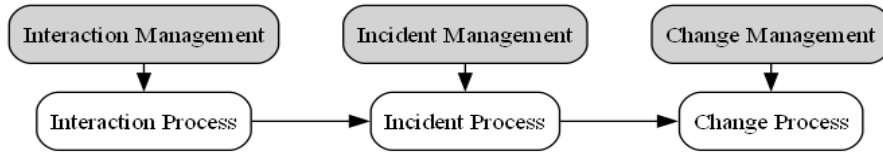


Figure 1. Overview of the ITIL process.

MATERIAL AND METHODS

The Collection of the Data

The dataset utilized in this study originates from Rabobank, a global Dutch banking and financial services corporation headquartered in Utrecht, Netherlands (Van Dongen, 2014; Cacciola et al., 2014). Covering the period from October 1, 2013 to March 31, 2014, the dataset comprises four distinct logs: Interaction records (147,004 events), Incident records (46,606 events), Incident Activity records (466,639 events), and Change records (30,236 events). This comprehensive dataset includes information on software changes, client interactions with the service desk, incidents logged within IT operations, and detailed incident management activities. It enables a robust analysis of IT service management processes and facilitates the extraction of meaningful features for predictive modeling. Previous studies utilizing similar datasets have commonly aimed to analyze business processes to enhance performance, identify bottlenecks, and predict future events (Lopes & Ferreira, 2019; Mehdiyev, Evermann, & Fettke, 2017, 2020; Impedovo, Pirlo, & Semeraro, 2023). Many of these works compare machine learning and deep learning techniques for tasks such as predicting the next step in a business process or estimating completion times. Notably, some research has shown that simpler algorithms—such as Decision Trees and Gradient Boosting—can outperform complex deep learning models, achieving higher accuracy rates. In contrast, our work differentiates itself by integrating process mining techniques to generate meaningful features and employing a time series forecasting approach with the ARIMAX model. This combination allows us to more effectively incorporate the underlying process dynamics into the analysis and achieve high predictive accuracy.

Method

In this study, we employed the ARIMAX (AutoRegressive Integrated Moving Average with Exogenous variables) model for time series forecasting. The ARIMA model is a widely recognized statistical technique used to analyze and predict future points in a time series by accounting for its own past values (autoregressive terms), the integration of past values (differencing to achieve stationarity), and past forecast errors (moving average terms). The general form of an ARIMA (p, d, q) model is expressed as (Box et al., 2015):

$$\phi_t(B)(1 - B)^d y_t = \theta_t(B)\epsilon_t$$

where

- y_t is the value of the time series at time t ,
- B is the backward shift operator ($B y_t = y_{t-1}$),
- $\phi_t(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$ represents the autoregressive (AR) terms of order p ,
- $\theta_t(B) = 1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q$ represents the moving average (MA) terms of order q ,
- d is the order of differencing required to make the series stationary,
- ϵ_t is the white noise error term.

The ARIMAX model extends the ARIMA model by incorporating exogenous variables (X_t) that are believed to influence the dependent time series y_t . The general form of the ARIMAX model is:

$$y_t = c + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + \beta_1 X_{t,1} + \dots + \beta_k X_{t,k} + \theta_1 \epsilon_{t-1} + \dots + \theta_q \epsilon_{t-q} + \epsilon_t$$

where

- c is a constant term,
- β_k are the coefficients of the exogenous variables,
- $X_{t,k}$ are the exogenous variables at time t

By incorporating exogenous variables, the ARIMAX model accounts not only for the internal dynamics of the time series but also for the impact of external factors. This makes it particularly suitable for forecasting in contexts where external predictors are significant, such as predicting incidents and interactions based on changes implemented.

The model parameters p , d , and q were determined through iterative testing and selection based on the Akaike Information Criterion (AIC), which balances model fit and complexity. The inclusion of relevant exogenous variables was guided by process mining insights, ensuring that the model effectively captures the underlying process dynamics influencing the time series.

ARIMAX Model Implementation

As illustrated in Figure 2, the proposed methodology employs time series analysis using the ARIMAX (AutoRegressive Integrated Moving Average with Exogenous variables) model to predict future incidents and interactions in IT service management. The process begins by collecting and preprocessing datasets such as Change, Interaction, and Incident logs from week t . Key features are extracted to encapsulate the effects of software changes on subsequent service disruptions.

Table 1. Features of the prediction model.

Feature	Description
Change Frequency in Week t	Number of changes implemented in week t
Change Complexity in Week t	Complexity of changes made in week t
Configuration Item Sensitivity in Week t	Correlation between changes and incidents after one week
Cycle Time in Week t	Duration of processes before incidents or interactions occur
Incident Severity in Week t	Severity of incidents in week t that affect week $t+1$
Interaction Complexity in Week t	Steps or escalations needed to resolve interactions in week t
Reassignments in Week t	Number of reassignments in week t before resolution
Escalations in Week t	Incidents escalated in week t that affect week $t+1$
Bottlenecks in Week t	Stages with delays in week t that affect future incidents
Activity Frequency in Week t	Frequency of activities in week t that influence week $t+1$

Figure 3 demonstrates the Directly-Follow Graph (DFG) of incident activities, which provides a visual representation of the most frequent activity transitions within the process. By filtering transitions based on a minimum frequency threshold, the figure highlights major steps such as "Open," "Assignment," and "Operator Update," enabling the identification of high-impact activities and transitions. These insights help in extracting key features, such as activity frequencies, sequence dependencies, and reassignment patterns, which directly feed into the feature engineering process. For example, the frequent transitions between "Open" and "Assignment" suggest the importance of understanding reassignment behaviors, while the transitions between "Operator Update" and "Closed" activities highlight the need to account for operator intervention frequency in future incident predictions.

These extracted features serve as inputs for the prediction model development phase. Initially, the model utilizes the number of changes completed in week t as the primary exogenous variable to forecast incidents and interactions in week $t + 1$. This alignment ensures that the predictive model captures the temporal relationship between changes implemented and their immediate impacts on service operations. To enhance predictive accuracy, additional exogenous variables were incorporated based on their relevance to incident and interaction occurrences. Table 1 outlines these features.

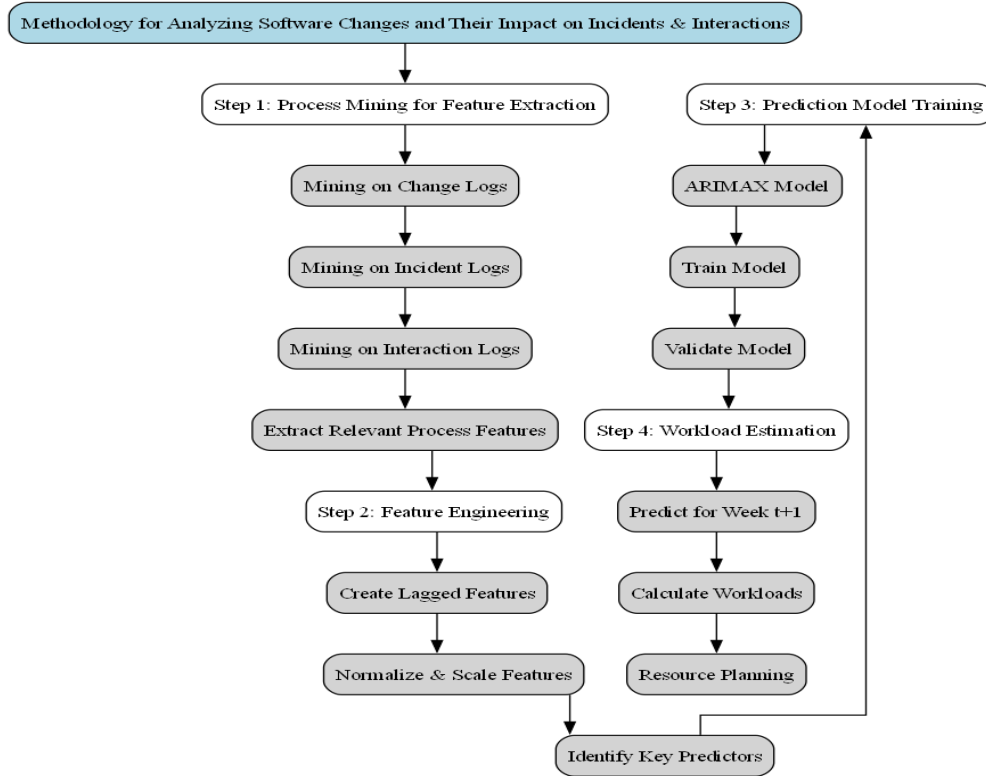


Figure 2. The proposed methodology.

The ARIMAX model leverages these features to capture various aspects of the change management process, formalized as $y_{t+h} = f(y_t, X_t)$ where y_t represents historical incident and interaction counts, and X_t denotes the exogenous variables from week t . The model parameters (p, d, q) were optimized using the Akaike Information Criterion (AIC) to balance model complexity and goodness of fit. This systematic approach, depicted in Figure 2, facilitates accurate anticipation of service demands, enabling effective workload planning and resource allocation for improved operational efficiency.

RESULTS

In this study, the ARIMAX modeling methodology was implemented in Python using statistical libraries such as statsmodels. The analysis commenced with data preprocessing, including handling missing values and ensuring the time series data was stationary or appropriately differenced. The dataset was split into training and testing sets to evaluate the model's performance on unseen data. Specifically, the last 8 weeks of the dataset were reserved for testing, while the preceding data was used for training the models. This approach allowed for a robust assessment of the model's predictive capabilities over the most recent period.

The initial ARIMAX model incorporated only the number of changes completed in week t as the exogenous variable. This model yielded a Mean Absolute Percentage Error (MAPE) of approximately 24.31% for weekly interactions and 20.77% for weekly incidents, as shown in Table 2.

To improve predictive accuracy, the model was enhanced by incorporating the additional exogenous variables listed in Table 1. The inclusion of these variables led to significant improvements in the model's performance. The optimal ARIMAX order changed to $(1,0,1)$, and the AIC values decreased, indicating a better fit. The MAPE values were substantially reduced to 15.12% for weekly interactions and 11.87% for weekly incidents, as summarized in Table 3.

Table 2. Initial ARIMAX model performance.

Model	Exogenous Variables	Best ARIMAX Order	AIC	MAPE (%)
Interactions	Change frequency	$(0, 1, 1)$	446.84	24.31

Incidents	Change frequency	(0, 1, 1)	391.39	20.77
-----------	------------------	-----------	--------	-------

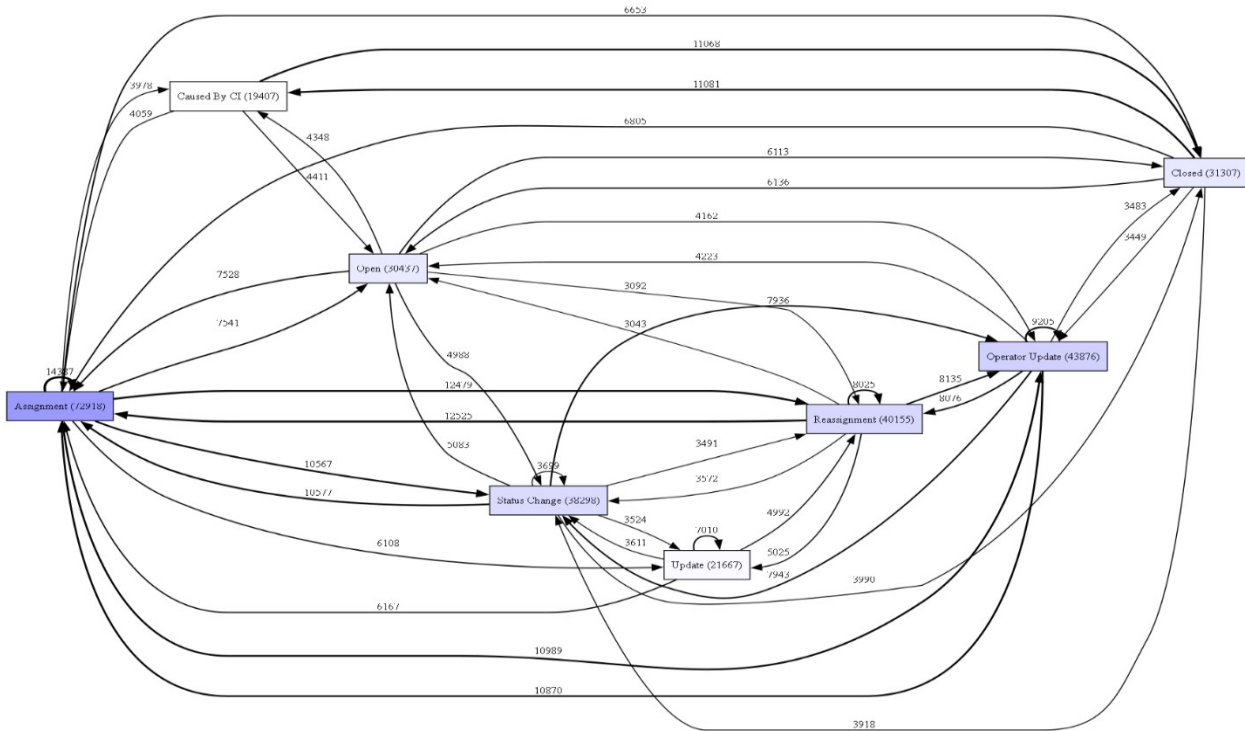


Figure 3. Frequency diagram of incident activities.

Table 3. Enhanced ARIMAX model performance with additional exogenous variables.

Model	Exogenous Variables	Best ARIMAX Order	AIC	MAPE (%)
Interactions	Change frequency+generated features	(1, 0, 1)	320.45	15.12
Incidents	Change frequency+generated features	(1, 0, 1)	315.78	11.87

Figures 4 and 5 illustrate the improved predictions of the enhanced ARIMAX models compared to the actual values. The significant reduction in MAPE indicates that the model’s predictions closely align with observed incidents and interactions, enhancing its utility for operational planning.

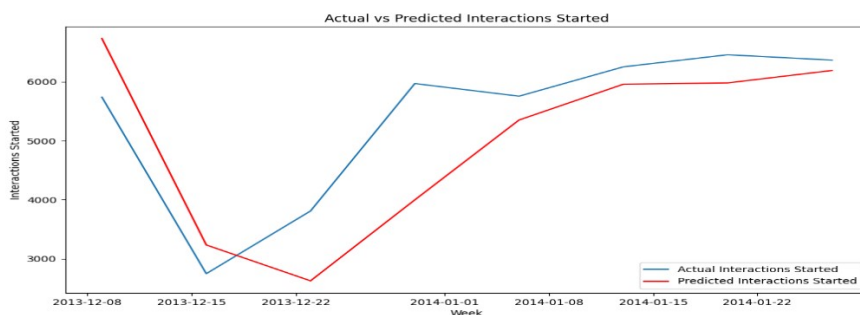


Figure 4. Actual and forecasted weekly interactions.

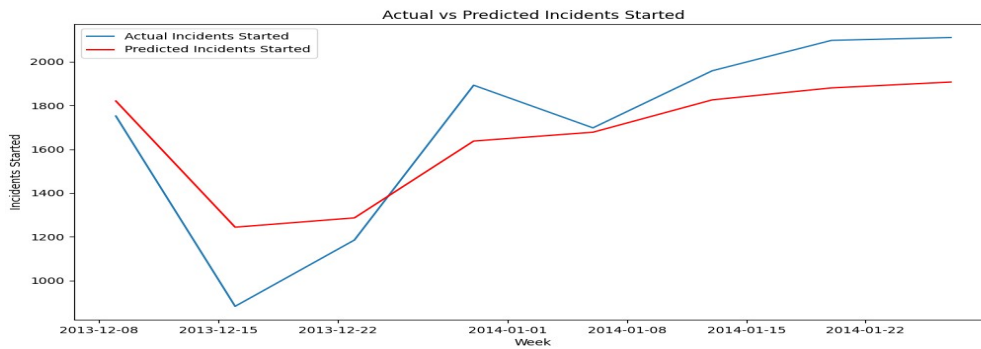


Figure 5. Actual and forecasted weekly incidents.

DISCUSSION AND CONCLUSION

This study successfully developed a predictive model using time series analysis to forecast future incidents and interactions resulting from frequent software changes in modern IT environments. By integrating process mining techniques to extract key features from change, incident, and interaction logs, we enriched the ARIMAX model with relevant exogenous variables. The enhanced model significantly improved predictive accuracy, reducing the Mean Absolute Percentage Error (MAPE) to approximately 12% for both incidents and interactions. This high level of accuracy demonstrates the model's effectiveness in capturing the complex relationships between software changes and their operational impacts.

The inclusion of additional features—such as change complexity, configuration item sensitivity, cycle time, and activity frequencies—proved instrumental in enhancing the model's performance. The Directly-

Follows Graph (DFG) of incident activities was particularly valuable in identifying critical transitions and bottlenecks within the processes. By focusing on high-impact activities and transitions, we were able to extract features that directly influenced future incidents and interactions, thereby improving the model's predictive capabilities.

Despite the positive outcomes, some limitations remain. The dynamic nature of IT environments and potential data quality issues necessitate continuous model updates and validations. External factors not captured in the current dataset may also influence incident and interaction rates. Future studies should consider integrating advanced analytical techniques, such as machine learning algorithms and simulation modeling, to capture nonlinear patterns and further enhance prediction accuracy.

In conclusion, the methodology presented offers a robust framework for anticipating service demands in IT service management. By leveraging time series analysis enriched with process mining-derived features, organizations can proactively manage risks, optimize resource allocation, and improve service quality. Continued refinement of the model and incorporation of additional data sources will further strengthen its utility, supporting informed decision-making in increasingly complex IT landscapes.

References

- Badakhshan, P., Wurm, B., Grisold, T., Geyer-Klingenberg, J., Mendling, J., & Vom Brocke, J. (2022). Creating business value with process mining. *The Journal of Strategic Information Systems*, 31(4), 101745.
- Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2015). *Time series analysis: forecasting and control*. John Wiley & Sons.
- Cacciola, G., Conforti, R., & Nguyen, H. (2014). Rabobank: a process mining case study BPI challenge 2014 report. Fourth International Business Process Intelligence Challenge (BPIC'14).
- Iden, J., & Eikebrokk, T. R. (2013). Implementing IT Service Management: A systematic literature review. *International Journal of Information Management*, 33(3), 512–523.
- Impedovo, D., Pirlo, G., & Semeraro, G. (2023). Next Activity Prediction: An application of shallow learning techniques against deep learning over the BPI Challenge 2020. *IEEE Access*.

- Lopes, I. F., & Ferreira, D. R. (2019). A survey of process mining competitions: the BPI challenges 2011–2018. In *Business Process Management Workshops: BPM 2019 International Workshops*, Vienna, Austria, September 1–6, 2019, Revised Selected Papers 17 (pp. 263-274). Springer International Publishing.
- Martin, N., Fischer, D. A., Kerpedzhiev, G. D., Goel, K., Leemans, S. J., Röglinger, M., ... & Wynn, M. T. (2021). Opportunities and challenges for process mining in organizations: results of a Delphi study. *Business & Information Systems Engineering*, 63, 511-527.
- Mehdiyev, N., Evermann, J., & Fettke, P. (2017, July). A multi-stage deep learning approach for business process event prediction. In *2017 IEEE 19th conference on business informatics (CBI)* (Vol. 1, pp. 119-128). IEEE.
- Mehdiyev, N., Evermann, J., & Fettke, P. (2020). A novel business process prediction model using a deep learning method. *Business & information systems engineering*, 62, 143-157.
- Mesquida, A. L., Mas, A., Amengual, E., & Calvo-Manzano, J. A. (2012). IT Service Management Process Improvement based on ISO/IEC 20000 and ISO/IEC 15504. *Journal of Software: Evolution and Process*, 24(6), 605–618.
- Munoz-Gama, J., Martin, N., Fernandez-Llatas, C., Johnson, O. A., Sepúlveda, M., Helm, E., ... & Zerbato, F. (2022). Process mining for healthcare: Characteristics and challenges. *Journal of Biomedical Informatics*, 127, 103994.
- Pollard, C., & Cater-Steel, A. (2009). Justifications, strategies, and critical success factors in successful ITIL implementations in US and Australian companies: An exploratory study. *Information Systems Management*, 26(2), 164–175.
- Potgieter, B. C., Botha, J. H., & Lew, C. (2005). Evidence that use of the ITIL framework is effective. *Proceedings of the 18th Annual Conference of the National Advisory Committee on Computing Qualifications*, 157–161.
- Urrea-Contreras, S. J., Flores-Rios, B. L., Astorga-Vargas, M. A., & Ibarra-Esquer, J. E. (2021, August). Process mining perspectives in software engineering: A systematic literature review. In *2021 Mexican International Conference on Computer Science (ENC)* (pp. 1-8). IEEE.
- Urrea-Contreras, S. J., Astorga-Vargas, M. A., Flores-Rios, B. L., Ibarra-Esquer, J. E., Gonzalez-Navarro, F. F., Garcia Pacheco, I., & Pacheco Agüero, C. L. (2024). Applying Process Mining: The Reality of a Software Development SME. *Applied Sciences*, 14(4), 1402.
- van Bon, J. (Ed.). (2007). *Foundations of IT Service Management Based on ITIL V3*. Van Haren Publishing.
- Van Dongen, B. F. (2014). BPI challenge 2014. In *4th International Business Process Intelligence Challenge (BPIC'14)*.
- Vom Brocke, J., Jans, M., Mendling, J., & Reijers, H. A. (2021). A five-level framework for research on process mining. *Business & Information Systems Engineering*, 1-8.

Zaman Serisi Tahmin Kombinasyonları İçin Çok Kriterli Karar Verme Yöntemine Dayalı Yeni Bir Yaklaşım

Tuğba Yasemin Karagöz^{1*}, Erkan Işıklı²

¹İşletme Mühendisliği Bölümü/İşletme Fakültesi, İstanbul Teknik Üniversitesi, İstanbul, Türkiye

²Endüstri Mühendisliği Bölümü/İşletme Fakültesi, İstanbul Teknik Üniversitesi, İstanbul, Türkiye

Özet

Zaman serilerinde tahmin kombinasyonları, farklı tahmin modellerinin güçlü taraflarını bütünleştirerek doğruluk performansı daha yüksek modeller elde etmeyi amaçlayan nicel tahmin yöntemleridir. Bu çalışmada, çok kriterli karar verme (ÇKKV) yöntemleri kullanılarak hem tekil tahmin modellerinden hem de ilgili literatürde en çok kabul gören bazı tahmin kombinasyon modellerinden daha iyi performans gösteren, çok boyutlu ve nesnel yeni bir yaklaşımın geliştirilmesi amaçlanmıştır. Bu bağlamda, tahmin modellerinin ağırlıklarının CRITIC ve TOPSIS yöntemlerini içeren hibrit bir ÇKKV yaklaşımıyla belirlendiği yeni bir kombinasyon modeli geliştirilmiştir. Geliştirilen model, 2001 yılı Ocak ile 2024 yılı Mayıs ayları arasını kapsayan, Amerika Birleşik Devletlerine ait aylık bazdaki elektrik enerjisi tüketicilerine doğal gaz teslimatı (MMcf) veri kümesi üstünde çalışılmıştır. Önerilen modelin göreceli performansı; geleneksel tekil tahmin modelleri (Basit Hareketli Ortalama, ETS, THETAM, STL), makine öğrenmesi modelleri (RandomForest, SVM, xgboost, NNAR) ve literatürde birçok kaynakta en etkili kombinasyon yöntemi olarak atfedilen basit aritmetik ortalama yöntemi de dahil olmak üzere regresyon ve optimizasyon temelli kombinasyon modellerini de kapsayacak şekilde halihazırda bulunan farklı zaman serisi tahmin kombinasyonu modelleri ile karşılaştırılmıştır. Test kümesinden elde edilen bulgular, önerilen modelin tekil tahmin modellerinden çoğunlukla daha iyi sonuç verdiğini, ele alınan kombinasyon modellerinden de daha başarılı olduğunu ortaya koymuştur.

Anahtar Kelimeler: Tahmin Kombinasyonları, Zaman Serileri, Çok Kriterli Karar Verme, TOPSIS, CRITIC.

GİRİŞ

Nicel tahminleme teknikleri; geçmiş verilere ve bu verilerin altta yatan istatistiksel eğilimlerine dayalı olarak gelecekteki belirli bir periyodun sonuçlarını öngörmek için yıllardır çalışılarak ve geliştirilerek kaynak tahsisi, üretim çizelgeleme, lojistik, risk yönetimi ve kısa/uzun vadeli stratejik planlama gibi karar alma, yönetim ve planlama faaliyetleri için önemli içgörüler sağlama potansiyelinde olmuştur. ARIMA, mevsimsel ayrıştırma ve üstel düzleştirme gibi geleneksel zaman serisi modelleri, tahminleme literatüründe yıllardır yaygın olarak kullanılmaktadır. Son yıllarda, sinir ağları, rastgele orman ve destek vektör makinesi gibi doğrusal olmayan makine öğrenimi modellerini içeren modern tahmin teknikleri de zaman serisi analizi için kullanılmakta olup büyük ve karmaşık veri kümelerinde oldukça ümit verici deneysel sonuçlar vermektedir (Masini vd., 2023).

Tahminleme için ayrı ayrı modeller kullanmanın yanı sıra, birden fazla tekil yöntemi bütünleştirmek de mümkündür. Literatürde, özellikle son birkaç on yıldır zaman serisi tahmin kombinasyonları hakkında göz ardı edilemeyecek sayıda araştırma ve uygulama yapılmıştır. Çeşitli alanları kapsayan bu çalışmalar, en yalın tahmin tekniklerinin birleştirilmesiyle tahmin ve öngörü performanslarının artırılabilirliğini ortaya koymuştur (Yang, 2004).

Hem tekil hem de kombine edilmiş tahmin modellerinin amacı, daha iyi doğruluk oranlarına ulaşmaktır. Literatürdeki çalışmaların çoğu, tahmin modellerinin performansını önceden belirlenmiş tek bir doğruluk ölçütüne göre veya farklı birkaç ölçütün “uzlaştırıldığı” noktayı esas alarak bir arada değerlendirmektedir. Kurulan modellerin performanslarının farklı yönlerini yakalayabilen birbirinden farklı doğruluk ölçütlerinden uygun olan(lar)ının belirlenmesi, bu yüzden önemli bir konu olarak karşımıza çıkmaktadır (Mehdiyev vd., 2016). Çeşitli performans ölçümlerinin eş zamanlı olarak dikkate alınması en iyi tahmin modelinin seçimini geliştirebilir (Xu & Ouenniche, 2012).

Literatürde, bir kombinasyon modelinin bileşenlerinin ağırlıkları belirlenirken, birbirleriyle muhtemel korelasyon ilişkisi olan farklı doğruluk ölçütlerini ve kombine edilen farklı tahmin modellerini aynı matris üzerinde dikkate alan nesnel bir metodolojiye ihtiyaç olduğu açıktır. Bu çalışmada, seçilen tekil tahmin modelleri “alternatifler,” doğruluk

ölçütleri ise “kriterler” olarak değerlendirilerek bir karar matrisine uyarlanmakta, ardından bu karar matrisine CRITIC ve TOPSIS yöntemlerini birleştiren bir ÇKKV yaklaşımı uygulanmaktadır. Böylece performans ölçütlerinin aralarındaki ilişki nesnel bir şekilde göz önünde bulundurulurken, tekil tahmin modellerinin ideal çözüme yakınlığına göre performans puanının elde edilmesi mümkün olmaktadır. Bu nihai performans puanlarının normalleştirilmiş değerleri, kombine edilecek her tekil tahmin modelinin ağırlıklarını temsil eder. Bu araştırma kapsamında, hibrit bir ÇKKV yönteminden elde edilen yeni bir kombinasyon modelinin performansı bulunarak hem tekil tahmin modelleriyle hem de literatürde halihazırda var olan en yaygın kombinasyon yöntemleriyle göreceli performansı karşılaştırılmıştır. Çalışmanın temel amacı, literatürdeki mevcut modellere kıyasla çok boyutlu ve nesnel bir zemin üzerinde, doğruluk performansı daha yüksek yeni bir zaman serisi tahmin modeli yaklaşımı oluşturmaktır. Bu amaçla, tek değişkenli bir zaman serisi olan ABD Elektrik Enerjisi Tüketicilerine Doğal Gaz Teslimatı (MMcf) veri kümesi üstünde çalışılarak, önerilen yeni yaklaşım test edilmektedir.

LİTERATÜR

Zaman Serileriyle Tahmin Kombinasyonları

Literatürde, birden fazla farklı tahmin modelinin bir araya getirilmesi, tek bir model tarafından tanımlanamayan farklı istatistiksel eğilimlerin, farklı modellerce yakalanabilmesi açısından değerli bir yaklaşım olarak kabul edilmektedir (Armstrong, 2007; Clemen, 1989; De Menezes vd., 2000; Hendry & Clements, 2002; Makridakis vd., 2018; Mariñas-Collado vd., 2022). Zaman serilerinin tahmin kombinasyonları, yepyeni bir çalışma alanı değildir; geçmiş John Bates ve Clive Granger’ın yaklaşık 60 yıl önceki bir çalışmasına kadar uzanmaktadır. Bates ve Granger (1969), ortalama kareler tahmin hatalarına ait matrisin köşegen elemanlarını kullanarak, tahmin modelleri için kombinasyon ağırlıkları atamışlar ve tekil tahmin modelleriyle karşılaştırıldığında daha doğru sonuçlar elde etmişlerdir. İlerleyen yıllarda Clemen (1989) ve Timmermann (2006) da bu alana önemli katkılarda bulunmuştur. Clemen (1989), tahmin kombinasyonlarına ilişkin o zamana kadarki en geniş araştırmayı yapmıştır. Basit ortalama alma, regresyon tabanlı modeller, Bayesgil teknikler ve bireysel modeller için ağırlıkların belirlenmesine yönelik matematiksel programlama yaklaşımlarını içeren hem teorik hem de deneysel çalışmalarını paylaşmış; öznel kombinasyon prosedürlerini de ele almıştır. Timmermann (2006) ise tahmin kombinasyonlarının teorik temelleri ve kapsamı üzerine, o zamana kadarki en kapsamlı çalışmadır; nokta, aralık ve olasılık tahminlerinin kombinasyonlarını tartışmıştır. Timmermann (2006)’ya göre literatürdeki tahmin kombinasyonlarının dört argümanı vardır: (1) İki modelden biri diğeri tarafından domine edilse dahi, her ikisinin kombinasyonu, bunların tek başına kullanımından daha iyi sonuçlar verebilir; (2) Kombinasyon modelleri, yapısal kırılmalardan çeşitli şekillerde etkilenebilen tekil tahmin modellerine kıyasla daha iyi performans gösterebilir; (3) Farklı modellerden elde edilen tahmin sonuçlarının birleştirilmesi, tahminlerin yanlılığına ve bireysel tahminlerin altında yatan, veri kümelerindeki ölçüm hatalarına karşı daha sağlam bir yaklaşım geliştirmenin bir yolu olarak görülebilir; (4) Farklı kayıp fonksiyonlarının optimizasyonuna dayanan tekil tahminler birleştirildikten sonra daha iyi tahmin sonuçları elde edilebilir. Tahmin kombinasyonlarıyla doğruluk performansı daha yüksek modellerin elde edilmesi için gerekli teorik altyapı ve ampirik sonuçlar uzun zamandır çalışılmaktadır. Tüm bu argümanlar ve kaynaklarla beraber tahminleri birleştirmenin nihai tahmin hatasını azalttığı konusunda genel bir görüş birliği bulunmaktadır; ancak bir amaç veya veri kümesi özelinde hangi kombinasyon yönteminin en iyi sonucu vereceğine yönelik net bir fikir birliği yoktur (Blanc ve Setzer, 2016). Wei & Yang (2012), kombinasyon modellerini (1) tahmin modellerinin varyans-kovaryans tahminine dayanan optimizasyon yaklaşımı (Bates ve Granger, 1969), (2) Bayesgil yaklaşım (Min & Zellner, 1993), (3) regresyon yaklaşımı (Granger & Ramanathan, 1984) ve (4) yapısal kırılmaları dikkate alan yaklaşım (Timmermann, 2006) olarak gruplamıştır.

ÇKKV Yaklaşımı Altında Zaman Serileriyle Tahminleme Yöntemleri

Literatürde ÇKKV yöntemlerini baz alarak tahmin kombinasyonu bileşenlerine ağırlık atayan bir metodolojiye rastlanmamıştır; ancak son yıllarda tek bir tahmin modelinin seçimine karar vermek için ÇKKV yöntemlerini kullanan kısıtlı sayıda çalışma yapılmıştır. Kumar & Kaur (2021), SOM (Self Organizing Map) ve SOMFTS (SOM Fuzzy Time Series) tabanlı tahminleme modellerini, dört farklı performans kriteri altında, yedi geleneksel tahminleme modeliyle karşılaştırmıştır. COVID-19 vakalarının gelecek 10 günlük tahminlemesi için geliştirdikleri

modelin, ölüm oranları için olmasa da, doğrulanmış ve iyileşen COVID-19 vakaları için bazı performans kriterleri altında diğer geleneksel tahminleme modellerine göre daha iyi sonuç verdiğini bulmuşlardır. Badulescu vd. (2021), nicel ve nitel modelleri birlikte değerlendirmek ve aralarından en iyi tahmin modelini seçmek için ANP ve TOPSIS temelli bir teknik kullanmıştır. Bu çalışmada model seçimi yapılırken, farklı performans ölçütlerinin birbirleriyle olan korelasyonlarının etkisini de göz önünde bulundurularak değerlendirme yapılmıştır. Kriter olarak atanan performans ölçütleri, iki uzmanın öznel karşılaştırmalarına dayalı olarak ANP ile ağırlıklandırıldığından, nesnel temelden uzak kalmaktadır. Çalışmada performans ölçütlerinin arasındaki ilişkinin göz önünde bulundurulduğu ve bulundurulmadığı durumlarda sonucun nasıl değişeceği de irdelenmiş, sıralanan yedi modelden sadece ikisinin yerinin birbiriyle değiştiği gözlemlenmiştir.

METODOLOJİ

Tahmin Modelleri

Bu çalışmada kullanılan tahmin teknikleri şunlardır: STL (Seasonal and Trend decomposition using Loess), ETS (Error, Trend, Seasonality), NNAR (Neural Network AutoRegressive), THETAM (Theta Model), Basit Hareketli Ortalama (SMA), Rastgele Orman (RandomForest), Destek Vektör Makinesi (Support Vector Machine), XgbOOST. Sayfa kısıtından dolayı açıklamalarına yer verilemeyen bu tekniklerle ait ayrıntılı bilgiye Hyndman & Athanasopoulos (2021), Shmueli & Lichtendahl (2018) ve Géron, (2019)'dan ulaşılabilir.

Performans Metrikleri

Bu çalışmada kullanılan tahminleme modellerinin karşılaştırılmasında yararlanılan performans metriklerinin tanımları ve formülleri Tablo 1'de verilmektedir (Theodosiou, 2011; Diebold, 2007; Hyndman & Athanasopoulos, 2018).

Tablo 1. Performans metrikleri

<i>n</i> : tahmin edilen gözlem sayısı; <i>a_i</i> : <i>i</i> zamanındaki gerçek değer; <i>f_i</i> : her <i>i</i> zamanındaki tahmin değeri.	
Tahmin Hataları (Forecast Errors)	
<i>Ortalama Hata (Mean Error)</i> : Tahmin değerleriyle gözlenmiş gerçek değerler arasındaki farkların ortalamasıdır.	$ME = \frac{1}{n} \sum_{i=1}^n (f_i - a_i)$
Ölçeğe Bağlı Hatalar (Scale-Dependent Errors)	
<i>Ortalama Mutlak Hata (Mean Absolute Error)</i> : Tahmin değerleriyle gözlenmiş gerçek değerler arasındaki mutlak farkların ortalamasıdır.	$MAE = \frac{1}{n} \sum_{i=1}^n f_i - a_i $
<i>Kök Ortalama Kare Hatası (Root Mean Squared Error)</i> : Tahmin değerleriyle gözlenmiş gerçek değerler arasındaki farkların karelerinin ortalamasının kareköküdür.	$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (f_i - a_i)^2}$
Ölçekli Hata (Scaled Errors)	
<i>Ortalama Mutlak Ölçekli Hata (Mean Absolute Scaled Error)</i> : Tahmin modelinin ortalama mutlak hatasını (MAE), basit bir naive modelin ortalama mutlak hatasıyla karşılaştırarak ölçekler.	$MASE = \frac{\frac{1}{n} \sum_{i=1}^n f_i - a_i }{\frac{1}{n-1} \sum_{i=2}^n a_i - a_{i-1} }$
Yüzde Hatalar (Percentage Errors)	
<i>Ortalama Hata Yüzdesi (Mean Percentage Error)</i> : Tahmin değerleriyle gözlenmiş gerçek değerler arasındaki farkların gerçek değerlere oranla yüzde hatalarının ortalamasıdır.	$MPE = \frac{1}{n} \sum_{i=1}^n \left(\frac{f_i - a_i}{a_i} \right) \times 100$
<i>Ortalama Mutlak Yüzde Hata (Mean Absolute Percentage Error)</i> : Tahmin değerleriyle gözlenmiş gerçek değerler arasındaki mutlak yüzde hatalarının ortalamasıdır.	$MAPE = \frac{1}{n} \sum_{i=1}^n \left \frac{f_i - a_i}{a_i} \right \times 100$
Korelasyon Metrikleri	
<i>ACF1 Ölçümü</i> : Zaman serisi tahmin modelleri bağlamında, artıkların ACF1'i, model kurulduktan sonra kalan herhangi bir otokorelasyon olup olmadığını değerlendirmek için kullanılır.	$ACF1 = \frac{\text{Cov}(f_i - a_i, f_{i-1} - a_{i-1})}{\text{Var}(f_i - a_i)}$

Kombinasyon Yöntemleri

Bu çalışmada önerilen kombinasyon modelinin sonuçlarının karşılaştırılması için literatürdeki öne çıkan basit ve sofistike kombinasyon yöntemlerine yer verilmiştir. Tekil modelleri birleştirmenin en etkili yollarından biri olan Basit Aritmetik Ortalama, karşılaştırma yapılacak modeller arasına özellikle eklenmiştir. İlgili yöntemlerin formülleri Tablo 2’de verilmektedir (Weiss vd., 2019).

Tablo 2. Önerilen kombinasyon yönteminin karşılaştırıldığı kombinasyon tahmin modelleri

R: birleştirilen tekil tahmin modellerinin sayısı; j: tekil tahmin model türü; f_{ji} : j. tekil tahmin modelinin i zamanında ürettiği tahminin değeri; λ : düzeltme faktörü; n: tahmin edilen gözlem sayısı; a_i : i zamanındaki gerçek değer; $\hat{\omega}_j$: j. tekil tahmin modeline atanan ağırlık; ε_i : i zamanındaki hata değeri.			
Kombinasyon Yöntemi	Formülasyon	Kombinasyon Yöntemi	Formülasyon
Basit Aritmetik Ortalama (BAO)	$f_i^c = \frac{1}{R} \sum_{j=1}^R f_{ji}$	Ayıklanmış Ortalama (AO)	$f_i^c = \frac{1}{R(1-2\lambda)} \sum_{j=\lambda R+1}^{(1-\lambda)R} f_{ji}$
Medyan (MED)	R tek sayı iken: $f_i^c = f_{(\frac{R}{2}+0,5,i)}$ R çift sayı iken: $f_i^c = \frac{1}{2} (f_{(\frac{R}{2},i)} + f_{(\frac{R}{2}+1,i)})$	Düzeltilmiş Ortalama (DO)	$f_i^c = \frac{1}{R} \left[\lambda R f_{(\lambda R+1,i)} + \sum_{j=\lambda R+1}^{R-\lambda R} \lambda R f_{(R-\lambda R,i)} \right]$
En Küçük Kareler Reg. (EKK)	$Z = \min \sum_{i=1}^n (a_i - f_i^c)^2$ $f_i^c = \hat{\alpha} + \sum_{i=1}^R \hat{\omega}_j f_{ji} + \varepsilon_i$	En Küçük Mutlak Sapma Reg. (EKMS)	$Z = \min \sum_{i=1}^n a_i - f_i^c $ $f_i^c = \hat{\alpha} + \sum_{i=1}^R \hat{\omega}_j f_{ji} + \varepsilon_i$

CRITIC-TOPSIS ile Çok Kriterli Karar Verme

Önerilen yöntem için tekil tahmin modellerine ağırlık ataması, performans kriterlerinin CRITIC ile ağırlıklandırıldığı bir karar matrisinin TOPSIS ile değerlendirilmesiyle gerçekleştirmiştir.

CRITIC Yöntemi

İlk olarak Diakoulaki vd. (1995) tarafından önerilen CRITIC, varyasyon değerleri üzerinden kriterler arasındaki çatışmayı ve zıtlık yoğunluğunu dikkate alır. Yöntem, bir kriterin standart sapması ve diğer kriterler arasındaki korelasyon ile ölçülen bilgi miktarına dayalı olarak ağırlıklar atar. CRITIC, birbiriyle bağımlı olabilen kriterler için de uygulanabilen nesnel bir ağırlıklandırma yöntemi olmasıyla diğer ağırlıklandırma yöntemlerinden ayrışır (Alinezhad ve Khalili, 2019). Yöntemin uygulama adımları aşağıdaki gibi sıralanabilir (Alinezhad ve Khalili, 2019; Diakoulaki vd., 1995):

Birinci Adım (Normalleştirilmiş karar matrisinin oluşturulması). r_{ij} , orjinal karar matrisinin i. alternatifinin j. performans değeri; x_{ij} , normalleştirilmiş karar matrisinde i. alternatifin j. performans değeri; $r_i^+ = \max(r_{i1}, r_{i2}, \dots, r_{im})$ ve $r_i^- = \min(r_{i1}, r_{i2}, \dots, r_{im})$ olmak üzere,

$$x_{ij} = \frac{r_{ij} - r_i^-}{r_i^+ - r_i^-}; i = 1, \dots, m \text{ ve } j = 1, \dots, n.$$

İkinci Adım (Her performans kriterinin standart sapmasının ölçümü). \bar{x}_j , j. kriterin değerler ortalaması ve n, alternatif

sayısı olmak üzere, $\sigma_j = \sqrt{\frac{1}{n-1} \sum_{j=1}^n (x_{ij} - \bar{x}_j)^2}; i = 1, \dots, m$

Üçüncü Adım (Kriterler arası korelasyon katsayısının hesaplanması). \bar{x}_k , k. kriterin ortalaması olmak üzere, $\rho_{jk} =$

$$\sum_{i=1}^m (x_{ij} - \bar{x}_j)(x_{ik} - \bar{x}_k) / \sqrt{\sum_{i=1}^m (x_{ij} - \bar{x}_j)^2 \sum_{i=1}^m (x_{ik} - \bar{x}_k)^2}$$

Dördüncü Adım (C indeksinin saptanması). $C_j = \sigma_j \sum_{k=1}^n (1 - \rho_{jk}); j = 1, \dots, n$

Beşinci Adım (Kriter ağırlıklarının atanması). $w_j = \frac{c_j}{\sum_{j=1}^n c_j}; j = 1 \dots, n$

TOPSIS Yöntemi

Pozitif ideal çözüme en yakın, negatif ideal çözüme ise en uzak mesafedeki alternatifin seçilmesine dayanan TOPSIS yönteminin uygulanma adımları aşağıdaki gibi sıralanabilir (Gaur vd., 2023; Hwang & Yoon, 1981):

Birinci Adım (Normalleştirilmiş karar matrisinin oluşturulması). TOPSIS için normalleştirilmiş karar matrisinin d_{ij} elemanı şu şekilde bulunabilir: $d_{ij} = \frac{r_{ij}}{\sum_{i=1}^m r_{ij}^2}; i = 1, \dots, m \quad j = 1 \dots, n.$

İkinci Adım (Ağırlıklandırılmış, normalize edilmiş karar matrisinin oluşturulması). $v_{ij} = w_j \cdot d_{ij}$

Üçüncü Adım (Pozitif İdeal ve Negatif İdeal çözümlerin belirlenmesi). v_j^* , j . kriterin sütunundaki “en iyi değeri” temsil eden pozitif ideal çözüm; v_j^- , j . kriterin sütunundaki “en kötü değeri” temsil eden negatif ideal çözüm olmak üzere,

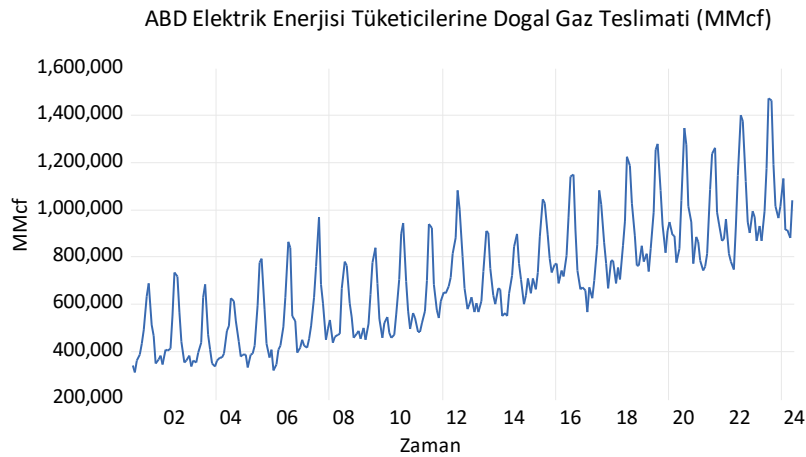
$$S_i^* = \sqrt{\sum_j (v_{ij} - v_j^*)^2} \quad \text{ve} \quad S_i^- = \sqrt{\sum_j (v_{ij} - v_j^-)^2}.$$

Dördüncü Adım (İdeal çözüme göreli yakınlığın hesaplanması). $C_i^* = \frac{S_i^-}{(S_i^* + S_i^-)}$. Negatif ideal çözüme uzaklığın toplam uzaklık içindeki payını temsil eden C_i^* katsayısından yola çıkılarak, en büyük C_i^* değerine sahip tahmin modelinin kombinasyon içerisinde en çok ağırlığa sahip olması gerektiğinin çıkarımı yapılır.

MODELİN UYGULANMASI

Verinin Hazırlanması

Önerilen yeni yaklaşım ABD Elektrik Enerjisi Tüketicilerine Doğal Gaz Teslimatı (MMcf) veri kümesi üzerinde test edilmiştir (*U.S. Energy Information Administration (EIA)*). Ocak 2001 ile Mayıs 2024 arasındaki dönemi kapsayan aylık veri kümesi toplam 281 gözlemlenir bir zaman serisidir ve davranışı Şekil 1’de resmedilmiştir.



Şekil 1. ABD elektrik enerjisi tüketicilerine doğal gaz teslimatı (MMCF) veri kümesinin grafiği

Trend ve mevsimsel etkiler içeren veri kümesinin Ocak 2001 – Aralık 2017 arasındaki dönemi eğitim, Ocak 2018 – Haziran 2022 arasındaki dönemi doğrulama, Temmuz 2022 – Mayıs 2024 arasındaki dönemi test amacıyla kullanılarak, daha önce bahsedilen geleneksel ve makine öğrenmesi tahmin algoritmaları uygulanmıştır. Tüm modellerin test ve doğrulama kümesi üstündeki performans ölçümleri not edilerek, önerilen yaklaşım, literatürdeki başka kombinasyon modelleriyle karşılaştırılmıştır.

Modellerin istatistiksel temele uygun olarak kurulabilmesi için öncesinde stokastik veya deterministik trendin varlığı, durağanlık testleri yapılarak sorgulanmıştır. İlgili sonuçlar Tablo 3’te verilmektedir.

Tablo 3. Durağanlık test sonuçları

Durağanlık testi	Denklem açıklaması	Kritik değer yüzdeleri (<i>t</i> -istatistiği)	Sonuç
ADF testi	Trend ve sabit terimli denklemde her iki terim de 0,000 <i>p</i> -değeri ile anlamlı	%1 seviyesinde (-3,99) %5 seviyesinde (-3,43) %10 seviyesinde (-3,13)	<i>t</i> -istatistiği: -4,33 <i>p</i> -değeri: 0,0033
KPSS testi	Trend ve sabit terimli denklemde her iki terim de 0,000 <i>p</i> -değeri ile anlamlı	%1 seviyesinde (0,22) %5 seviyesinde (0,15) %10 seviyesinde (0,12)	LM-istatistiği: 0,10

ADF testine göre, trend ve sabit terimli denklemle başladığında, terimler istatistiksel olarak anlamlıdır ve bu denklemin birim kökü yoktur (stokastik trendin varlığı reddedilebilir). Tek değişkenli bir zaman serisinin birim kökü olup olmadığını test etmek için kullanılan KPSS testi de aynı sonucu vermiştir. Grafik analiz verinin bir trend içerdiğine işaret etmektedir. ADF ve KPSS testleri bu trendin stokastik olmadığını ortaya koyduğundan, verinin deterministik trend içerdiği çıkarımı yapılarak, detrendize etme işlemi uygulanmıştır.

Özgün Yaklaşımın Uygulanması

Detrendize edilmiş veri kümesinin ilk 204 aylık kısmı eğitim, sonraki 54 aylık kısmı doğrulama, son 23 aylık kısmı ise test amacıyla kullanılmıştır. Yukarıda bahsedilmiş olan tahmin yöntemleri ve her yöntemin doğrulama kümesindeki performansından yola çıkılarak oluşturulan karar matrisi Tablo 4'te; CRITIC yönteminin performans metriklerine uygulanması neticesinde belirlenen kriter ağırlıkları ise Tablo 5'te verilmiştir. Kriter ağırlıklarının belirlenmesinin ardından, karar matrisi TOPSIS yönteminin uygulanmasına hazır hale gelmiştir. TOPSIS neticesinde her modelin pozitif ideal çözüm değerlerinden sapma oranları (S_i^*), negatif ideal çözüm değerlerinden sapma oranları (S_i^-), ideal çözüme göreli yakınlık değerleri (C_i^+) ve C_i^+ değerlerinin normalize edilmesinden elde edilmiş tahmin modellerinin kombinasyon içerisindeki ağırlıkları Tablo 6'da gösterilmiştir.

Tablo 4. Doğrulama kümesi üzerinde oluşturulmuş karar matrisi

Model	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
STL	30564,5	68454,7	55121,3	-35,5175	155,608	0,704979	0,516842
ETS	23342,6	69064,1	56298,1	-45,2766	152,852	0,809448	0,505613
NNAR	65322,5	137444,7	99896,3	-31,8329	186,327	2,494934	0,541119
THETAM	-157886,1	983127,3	523706,8	601,5116	1140,552	0,537545	0,001429
SMA(<i>n</i> =3)	50676,3	121633,3	93221,7	-120,263	214,812	1,947799	0,49149
RF	17520,6	118274,3	93115,1	-46,8808	157,530	1,013367	0,283048
SVM	22898,5	125190,0	100549,3	-95,3070	200,480	1,748163	0,462011
xgbOOST	52744,0	165314,1	114099,7	27,2878	123,799	2,129218	0,562711

Tablo 5. CRITIC yöntemine göre performans kriterlerinin ağırlıkları

Performans Kriteri	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Ağırlık	%17,7	%12,0	%11,7	%11,7	%12,2	%16,5	%18,2

Tablo 6. CRITIC ile ağırlıklandırılmış kriter ağırlıkları altında TOPSIS sonuçları (normalize edilmiş tekil tahmin modeli ağırlıkları ile birlikte)

Model	S_i^*	S_i^-	C_i^+	Model Ağırlıkları
STL	0,07391	0,28027	0,791	%13,9
ETS	0,07202	0,27612	0,793	%13,9
NNAR	0,11455	0,28675	0,715	%12,6
THETAM	0,26569	0,10711	0,287	%5,1
SMA(<i>n</i> =3)	0,09412	0,28565	0,752	%13,2
RF	0,04466	0,26930	0,858	%15,1
SVM	0,08098	0,26755	0,768	%13,5
xgbOOST	0,10557	0,27461	0,722	%12,7
		TOPLAM	5,686	%100,0

Bulguların Yorumlanması

Doğrulama kümesine uygulanan hibrit ÇKKV modeli sonrası elde edilen ağırlıklar kullanılarak oluşturulmuş yeni kombinasyon yöntemi, tekil modellerin ve diğer kombinasyon modellerinin test kümesindeki performansı ile Tablo 7’de karşılaştırılmıştır. Önerilen kombinasyon yöntemine oranla daha kötü performans gösteren modellere ait hücreler kırmızı renkli dolguyla boyanmıştır.

SONUÇ VE ÖNERİLER

Bulgular, belirlenen farklı istatistiksel yönelimleri yakalama potansiyeline sahip bazı geleneksel ve makine öğrenmesi tahminleme yöntemlerinin farklı performans metrikleri altında nesnel bir yöntemle bir arada değerlendirildiği ve sonucunda ideal çözüme yakınlığa göre tahminleyici ağırlıklarının atandığı CRITIC-TOPSIS tabanlı yeni modelin performansının, karşılaştırma yapılan tekil modellere ve kombinasyon modellerine nazaran oldukça iyi olduğuna işaret etmektedir. İlgili literatürde, etkili ve yenmesi zor olarak değerlendirilen BAO yöntemiyle karşılaştırıldığında, yeni modelin, ortalama hata ve ortalama mutlak ölçekli hata dışındaki tüm performans metriklerine göre, daha iyi olduğu veya BAO yöntemininkine yakın sonuçlar verdiği görülmüştür. Bu bağlamda, zaman serisi tahmin kombinasyonları literatürüne ÇKKV yöntemlerinin yeni bir soluk getirmesi ve ÇKKV çatısı altında başka modeller de üretilebilmesi mümkündür.

Tablo 7. Tekil ve kombinasyon tahmin modellerinin test kümesi performanslarının önerilen yönteminkiyle karşılaştırması

	Performans Sonucu Tahmin Modeli	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Tekil Modeller	STL	74540,3	99709,6	78176,6	-214,246	263,446	1,07420	0,51081
	ETS	67309,5	98850,8	76377,7	-186,928	240,237	1,17603	0,46837
	NNAR	123495,7	223858,8	144827,6	-260,786	367,994	6,44352	0,59011
	THETAM	-148631,5	929510,9	515490,8	-448,630	1227,260	0,52000	-0,03472
	SMA(n=3)	87749,0	158351,0	110997,5	-285,622	442,194	2,34143	0,45119
	RF	13825,3	134342,0	110412,8	-255,887	475,126	1,27798	0,25941
	SVM	41110,8	152384,0	120226,9	-188,174	470,140	1,96327	0,40108
	xgbOOST	72068,1	167241,7	121791,2	-202,141	313,970	2,78121	0,59907
Önerilen Yöntem		55273,4	105197,3	73307,0	-45,331	107,153	1,00536	0,35763
Diğer Komb. Modelleri	BAO	42381,3	104752,0	79098,4	-265,049	320,789	0,58009	-0,25958
	EKK	-171243,2	837294,3	451636,0	-364,290	1092,970	0,51670	-0,01781
	MED	72313,7	121158,0	88745,2	-270,623	338,078	2,08147	0,47321
	AO	72313,7	121158,0	88745,2	-270,623	338,078	2,08147	0,47321
	DO	71650,7	120103,7	87552,1	-268,531	334,534	2,06608	0,48900
	EKMS	-81423,2	440910,6	244502,3	-332,992	703,941	0,50501	-0,0678

Referanslar

- Alinezhad, A., & Khalili, J. (2019). *New Methods and Applications in Multiple Attribute Decision Making (MADM)* (Vol. 277). Springer International Publishing. <https://doi.org/10.1007/978-3-030-15009-9>
- Armstrong, J. S. (Ed.). (2007). *Principles of forecasting: A handbook for researchers and practitioners* (7. print). Springer.
- Badulescu, Y., Hameri, A.-P., & Cheikhrouhou, N. (2021). Evaluating Demand Forecasting Models Using Multi-Criteria Decision-Making Approach. *Journal of Advances in Management Research*, 18(5), 661–683. <https://doi.org/10.1108/JAMR-05-2020-0080>
- Bates, J. M., & Granger, C. W. J. (1969). The Combination of Forecasts. *Operational Research Society*, 20(4), 451–468.
- Blanc, S. M., & Setzer, T. (2016). When to choose the simple average in forecast combination. *Journal of Business Research*, 69(10), 3951–3962. <https://doi.org/10.1016/j.jbusres.2016.05.013>
- Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, 5(4), 559–583. [https://doi.org/10.1016/0169-2070\(89\)90012-5](https://doi.org/10.1016/0169-2070(89)90012-5)

- De Menezes, L. M., W. Bunn, D., & Taylor, J. W. (2000). Review of guidelines for the use of combined forecasts. *European Journal of Operational Research*, 120(1), 190–204. [https://doi.org/10.1016/S0377-2217\(98\)00380-4](https://doi.org/10.1016/S0377-2217(98)00380-4)
- Diakoulaki, D., Mavrotas, G., & Papayannakis, L. (1995). Determining objective weights in multiple criteria problems: The critic method. *Computers & Operations Research*, 22(7), 763–770. [https://doi.org/10.1016/0305-0548\(94\)00059-H](https://doi.org/10.1016/0305-0548(94)00059-H)
- Diebold, F. X. (2007). *Elements of forecasting* (4th ed). Thomson/South-Western.
- Gaur, S., Dosapati, S., & Tawalare, A. (2023). Stakeholder assessment in construction projects using a CRITIC-TOPSIS approach. *Built Environment Project and Asset Management*, 13(2), 217–237. <https://doi.org/10.1108/BEPAM-10-2021-0122>
- Géron, A. (2019). *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems* (Second edition). O'Reilly.
- Granger, C. W. J., & Ramanathan, R. (1984). Improved methods of combining forecasts. *Journal of Forecasting*, 3(2), 197–204. <https://doi.org/10.1002/for.3980030207>
- Hendry, D. F., & Clements, M. P. (2002). Pooling of forecasts. *The Econometrics Journal*, 7(1), 1–31. <https://doi.org/10.1111/j.1368-423X.2004.00119.x>
- Homepage—U.S. Energy Information Administration (EIA). (n.d.). Retrieved 12 September 2024, from <https://www.eia.gov/index.php>
- Hwang, C.-L., & Yoon, K. (1981). *Multiple Attribute Decision Making* (Vol. 186). Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-642-48318-9>
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: Principles and Practice*.
- Kumar, A., & Kaur, K. (2021). A hybrid SOM-Fuzzy time series (SOMFTS) technique for future forecasting of COVID-19 cases and MCDM based evaluation of COVID-19 forecasting models. *2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)*, 612–617. <https://doi.org/10.1109/ICCCIS51004.2021.9397216>
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and Machine Learning forecasting methods: Concerns and ways forward. *PLOS ONE*, 13(3), e0194889. <https://doi.org/10.1371/journal.pone.0194889>
- Mariñas-Collado, I., Sipols, A. E., Santos-Martín, M. T., & Frutos-Bernal, E. (2022). Clustering and Forecasting Urban Bus Passenger Demand with a Combination of Time Series Models. *Mathematics*, 10(15), 2670. <https://doi.org/10.3390/math10152670>
- Masini, R. P., Medeiros, M. C., & Mendes, E. F. (2023). Machine learning advances for time series forecasting. *Journal of Economic Surveys*, 37(1), 76–111. <https://doi.org/10.1111/joes.12429>
- Mehdiyev, N., Enke, D., Fettke, P., & Loos, P. (2016). Evaluating Forecasting Methods by Considering Different Accuracy Measures. *Procedia Computer Science*, 95, 264–271. <https://doi.org/10.1016/j.procs.2016.09.332>
- Min, C., & Zellner, A. (1993). Bayesian and non-Bayesian methods for combining models and forecasts with applications to forecasting international growth rates. *Journal of Econometrics*, 56(1–2), 89–118. [https://doi.org/10.1016/0304-4076\(93\)90102-B](https://doi.org/10.1016/0304-4076(93)90102-B)
- Shmueli, G., & Lichtendahl, K. C. (2018). *Practical time series forecasting with R: A hands-on guide* (Second edition). Axelrod Schnall Publishers.
- Theodosiou, M. (2011). Forecasting monthly and quarterly time series using STL decomposition. *International Journal of Forecasting*, 27(4), 1178–1195. <https://doi.org/10.1016/j.ijforecast.2010.11.002>
- Timmermann, A. (2006). Chapter 4 Forecast Combinations. In *Handbook of Economic Forecasting* (Vol. 1, pp. 135–196). Elsevier. [https://doi.org/10.1016/S1574-0706\(05\)01004-9](https://doi.org/10.1016/S1574-0706(05)01004-9)
- Wei, X., & Yang, Y. (2012). Robust forecast combinations. *Journal of Econometrics*, 166(2), 224–236. <https://doi.org/10.1016/j.jeconom.2011.09.035>
- Weiss, C., E., Raviv, E., & Roetzer, G. (2019). Forecast Combinations in R using the ForecastComb Package. *The R Journal*, 10(2), 262. <https://doi.org/10.32614/RJ-2018-052>

Xu, B., & Ouenniche, J. (2012). Performance evaluation of competing forecasting models: A multidimensional framework based on MCDA. *Expert Systems with Applications*, 39(9), 8312–8324. <https://doi.org/10.1016/j.eswa.2012.01.167>

Yang, Y. (2004). Combining Forecasting Procedures: Some Theoretical Results. *Econometric Theory*, 20(01). <https://doi.org/10.1017/S0266466604201086>

Bilgilendirme

Bu çalışma, İstanbul Teknik Üniversitesi'nde yürütülmekte olan, danışmanlığını Erkan Işıklı'nın yaptığı Tuğba Yasemin Karagöz'ün "A Novel Approach for Forecasting Combinations Based on Multi-Criteria Decision Making (MCDM)" başlıklı yüksek lisans tez çalışmasının bir uzantısıdır.

Çıkar çatışması

Yazarlar çıkar çatışması olmadığını beyan etmişlerdir.

Yazar Katkıları

Bu çalışmanın 1. yazarı model geliştirme, literatür taraması, veri uygulama, sayısal analiz, sonuçların elde edilmesi ve makalenin yazılmasından sorumlu olmuş; 2. yazarı araştırma metodolojisinin geliştirilmesinden sorumlu olmuş, tüm çalışmayı kavramsallaştırmış ve gözden geçirmiş, makalenin yazılması ve düzenlenmesine katkıda bulunmuştur.

Nesne Tanıma ve Sınıflandırmada Kullanılan Derin Öğrenme Yöntemleri

Erdem UYSAL¹, Hamza EROL²

¹Mersin Üniversitesi Mersin Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, Mersin, Türkiye

²Mersin Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Mersin, Türkiye

Özet

Bilgisayarlı görü çerçevesinde nesne tanıma ve sınıflandırma büyük veri analitiğinde geniş kullanım alanına sahiptir. Derin öğrenme algoritmaları görüntü işlemede problemlerin kolay çözülebilmesini sağlamaktadır. Bu çalışmada derin öğrenme yöntemleri kullanılarak nesne tanıma ve sınıflandırma için kullanılan modeller incelenmiş ve nesne tanıma ve sınıflandırmada kullanılan derin öğrenme yöntemlerinin nesne tanıma ve sınıflandırma performansları araştırılmıştır.

Anahtar Kelimeler: Derin Öğrenme; Nesne Tanıma; Nesne Sınıflandırma

1. Giriş

Derin öğrenme 2012 yılında nesne sınıflandırma için yapılan büyük ölçekli görsel tanıma (ImageNet) yarışmasında elde ettiği başarı ile ortaya çıkmıştır [1]. Derin öğrenmenin temelleri geçmişe dayansa da özellikle son yıllarda popüler olmasının sebeplerinden ilki eğitim için yeteri kadar verinin olması, ikinci olarak bu veriyi işleyecek donanımsal alt yapının bulunmasıdır.

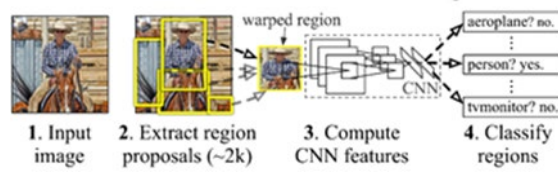
Çok katmanlı sinir ağı modelinin ortaya çıkmasıyla birlikte katman sayısının artırılarak daha iyi sonuçlar vermesi için Convolutional Neural Network (CNN) geliştirilmiştir [2]. Burada yer alan sinir ağı modelinde gizli katmanlar yer almakta ve elde edilen sonuçlar oldukça başarılı olmaktadır[2]. Çok katmanlı ileri beslemeli bir yapay sinir ağı olan CNN özellikle görüntü analizleri için kullanılmaktadır. CNN hayvan görü sistemine dayanan bir yaklaşımla ortaya atılmıştır. Filtrelemeye dayalı bir yapıdadır. Kullanılacak olan filtre ile görüntünün özelliğini belirtecek özellikler belirgin hale getirir [2]. Özellikle sınıflandırıcı işlemlerinde başarılı sonuçlar üretmektedir. Derin öğrenme yöntemleri birçok sektörde başarılı olarak kullanılmaktadır [1].

2. Yöntem

Parametrik modeller eğitim aşamasından sonra ağırlıklar ve bias (yanlılık) değerleri oluşturur. Ardından sınıflama aşamasında bu değerleri kullanarak kolay ve hızlı bir şekilde matris çarpımıyla sınıflama işlemi yapar. K-En yakın Komşuluk (KNN) gibi makine öğrenmesi yöntemleri her sınıflama yapılacağı zaman veri seti üzerinde gezmesi ve uzaklık hesaplaması gerektiğinden yavaş kalmaktadır [3]. Bu durum görüntü verisi gibi çok boyutlu veri olduğunda bellek yetersizliği yaşanmaktadır. Bu yüzden parametrik olmayan modeller çok boyutlu verilerde uygun değildir. Parametrik öğrenmenin 4 modülü vardır [3]. Bunlar: (i) Veri, (ii) Skor fonksiyonu, (iii) Kayıp Fonksiyonu, (iv) Ağırlıklar ve bias'ler. Bazı kayıp fonksiyonları: (i) Hinge Loss (Multi class SVM Loss), (ii) Cross Entropy Loss, (iii) Softmax Classifier. Bazı optimizasyon yöntemleri: (i) Stochastic Gradient Descent, (ii) Gradient Descent, (iii) Regularization (Genelleştirme) L1 Regularization, L2 Regularization, Elastic Net. Derin öğrenme tabanlı nesne tanıma ve tespit modelleri ise şunlardır: (i) R-CNN, (ii) Fast R-CNN, (iii) Faster R-CNN, (iv) Mask R-CNN, (v) YOLO, (vi) Retinanet, (vii) SSD(Single Shot Detector). Nesne tanıma için derin sinirsel ağların gücünden yararlanılmaktadır. Sadece nesne sınıflandırması değil nesne tanıma ve nesnenin görüntüdeki konumunun tespit edilmesi çalışılmaktadır. Yapay sinir ağlarının bir versiyonu olan CNN bu işlem için yaygın olarak kullanılmaktadır. Bu çalışmada nesne tanıma ve sınıflandırmada kullanılan derin öğrenme yöntemlerinin nesnelere tanıma ve sınıflandırma performansları araştırılmıştır. Nesne tanımda tek aşamalı model olan YOLO ve Dense Distinct Queries (DDQ), iki aşamalı model olan Faster R-CNN kullanılmıştır. Nesne sınıflandırma için kullanılan modeller hesaplamalar bölümünde verilmiştir.

2.1. Nesne Tanıma

Nesne tanıma görüntüdeki anlamlı nesnelere yerinin tespiti ve çerçeve kutu çizimi işlemidir. Bir nesneyi tanımak, onu bir sınıflandırma içerisinde değerlendirmek anlamına gelir [4]. Görüntünün sınıflandırılması için de görüntü üzerinden renk, şekil gibi yüksek düzeyde bilgileri çıkartacak görüntü işleminin gerçekleştirilmesi gerekir [4]. Nesne tanımda kullanılan mimariler: (i) Tek Aşamalı Model Mimarisi ve (ii) Çift Aşamalı Model Mimarisi (R-CNN) olmak üzere incelenir. Bunlar: (1) Tek Aşamalı Modeller olarak farklı YOLO versiyonları, SSD, DETR, DDQ ve (2) İki Aşamalı Modeller olarak: (i) R-CNN, (ii) Fast R-CNN, (iii) Faster R-CNN ve (iv) Mask R-CNN [5].



Şekil 1. İki Aşamalı Model R-CNN modeli.

Nesne tanımda İki Aşamalı Modeller:

R-CNN, iki aşamalı R-CNN modeli şekil 1’ de verilmiş ve 4 bölümden oluşmaktadır [5]. Bunlar: (I) Kategori bağımsız bölge önerileri üretmek, (II) Her bölge sunumundan sabit uzunluk özellik vektörlerini çıkarmak, (III) Bir görüntüdeki nesnelere sınıflandırmak için sınıf spesifik lineer SVM seti oluşturmak ve (IV) Çerçeve tespiti için regresör belirlemek.

Fast R-CNN, bir yıl sonra Ros Grishcik tarafından daha hızlı versiyonu olarak sunulmuştur [5]. R-CNN’ in her bölge sunumu için hesaplama paylaşımı yapmadan ConvNet ileri yayılım işlemi yapması yüzünden R-CNN SVM’ lerin sınıflamasında daha uzun zaman almaktadır. Fast R-CNN özellikleri (feature) tüm giriş görüntüsünden tek seferde çıkarılır ve region of interest pooling katmanına aktararak sabit sayıda feature’ ları çıkarır ve bu feature’ ları devamındaki sınıflama ve bounding box regresyonu için fully connected katmanlara input olarak verir. Feature’ lar tüm görüntüden tek seferde çıkarılır ve bir seferde sınıflama ve lokalizasyon için CNN’ e gönderilir. R-CNN’e kıyaslanırsa, region proposals CNN’e aktarılırken büyük miktarda zaman CNN hesaplaması sırasında kazanılmakta ve feature’ ları depolamak için ise geniş disk alanı kazanılmaktadır. R-CNN eğitimi multistage’ dir. Pre-training aşaması, finetuning aşaması, SVM’ lerin sınıflama aşaması ve bounding box regresyon aşamasıdır. Fast R-CNN her labeled RoI(Region of Interest) üzerinde multi-task loss kullanarak tek aşamalı end-to-end eğitim işlemidir. Bir diğer geliştirme ise RoI pooling katmanı kullanmasıdır. Bu sayede farklı boyutlardaki bölge önerilerinden sabit boyutlu özellik haritası çıkarılır. Bu işlem warping bölgeler olmadan ve bölge önerilerinin özelliklerinin spatial bilgilerini koruması açısından önemlidir. Hızlı tespit için yazar truncated SVD kullanır [5].

Faster R-CNN, Fast R-CNN çalışmasının ardından ortaya çıkmıştır. Bu çalışma daha da geliştirilerek bölge tabanlı CNN temelini iyileştirir. Fast R-CNN seçici arama (selective search) yaparak RoI önerir. Bu işlem yavaş ve tespit ağı kadar zaman aldığı için Faster R-CNN yeni bir RPN(Region Proposal Network) ile değiştirir. RPN geniş aralık en boy ve boyutlu bölge sunumları verimli bir şekilde tahmin edebilen tamamen bağlı konvolüsyon ağıdır. RPN bölge sunumlarının oluşturulmasını hızlandırır. Çünkü fully-image convolution feature’ larını paylaşır ve konvolüsyon setini tespit ağı ile ortaklaşa kullanır. Farklı boyutlarda nesne tespiti için multi-scale anchor’ lar kullanılmıştır. Anchor’ lar farklı boyutlarda bölge önerileri üretmeyi kolaylaştırmaktadır [5].

Mask R-CNN, temel olarak instance segmentation için kullanılmaktadır. Feature pyramid network içermektedir. Bu sayede önemli özellikler çıkarılabilir. Ayrıca başarıyı daha da arttırmak için RoI pooling’ i RoIAlign ile değiştirir [5].

Nesne tanımda Tek Aşamalı Modeller: (i) YOLO ve versiyonları, (ii) SSD

YOLO (You only look once),’yu diğer algoritmalarından ayıran gerçek zamanlı nesne tespiti yapmasıdır[6]. R-CNN 2000 bounding box tahmin ederken YOLO her görüntü için 100 bounding box tahmin etmektedir. Bu sayede gerçek zamanlı çalışabilmektedir. İkinci olarak YOLO problemi regresyon problemi olarak ele alır. Görüntü verisini SxS gride bölerek her hücresi nesnenin orta noktasının düştüğü nesneyi tespit etmektedir. Yeterlik skoru iki parçanın çarpımı ile oluşur. Nesnenin olma olasılığı IoU(Intersection over Union), o nesnenin ne kadar doğru olduğunu belirtir. Her grid hücresi B çerçeveleyici kutu ve herbiri için emin olma skorunu tahmin etmektedir. Ayrıca C sınıf için C boyutlu bağımlı sınıf olasılığı verir. Feature extraction ağı 24 convolution katmanının ardından average pooling katmanı ve fully connected katmanı ile devam etmektedir.

YOLO v2, Yolo nun hızını ve başarısını arttırmak için yeni tasarım içerir. Bu tasarımlar şunlardır: (i)Batch Normalizasyon [8], (ii) Yüksek Çözünürlüklü Sınıflayıcı [8], (iii) Anchor Boxlar ile Konvolüsyon [8], (iv) Anchor boxların boyutunu ve aspect ratio’ larını tespit etmek için Dimension Clusters [8], (v) Fine-grained Features [8], (vi) Multi-scale Training [8]. Ayrıca YOLOv2 yeni bir sınıflama backbone’ u önerir. Adı da Darknet-19. 19 convolution katmanı ve 5 max pooling katmanı ile görüntüyü işlemek için daha az işlem adımı ama daha fazla doğruluk elde eder.

YOLOv3, Yolo v2 nin geliştirilmiş versiyonudur. Çok etiketli sınıflama ve independent logistic classifier kullanılmaktadır [8].

SSD (Single Shot Detector), farklı boyutlarda ve en boy oranlarında varsayılan çerçeveleyici kutu kullanır. Farklı özellik haritalarında bu kutuların oranları en yüksek ve en düşük katman arasında belirli boşluklarla boyutları hesaplanır. Her spesifik özellik haritası nesnenin boyutlarına duyarlı olarak öğrenir. Her varsayılan kutu için tüm nesne kategorilerinde offset ve emin olma değerleri tahmin edilir [5].

DETR (End-to-End Object Detection), Facebook araştırma ekibi tarafından 2020 yılında çıkmıştır. Genellikle doğal dil işleme mimarilerinde kullanılan Transformers mimarisi tabanlı bir nesne tanıma modelidir. Nesne tanıma için region proposal veya son işlem adımları gibi ayrı bileşen ihtiyaçlarını elimine edecek birleşik yaklaşımı benimsemektedir. Paralel işlem yeteneğinden ve bünyesinde anchor boxes, NMS gibi kısıtlayıcı teknikler kullanmamasından dolayı hızlı bir yapısı vardır [11].

Dense Distinct Queries (DDQ), var olan modellerdeki sparse ve dense query' lerin limitlerini adresleyerek nesne tanımayı iyileştiren yeni bir yaklaşımdır. Geleneksel sparse query' ler recall değerleri ile sorun yaşarken, dense query' ler genellikle fazla sayıda olmasından dolayı optimizasyon sorunlarıyla karşılaşmaktadır.

Yaklaşım, ilk olarak dense query' ler ürettikten sonra tekil olanları sınıf yansız non-maximum suppression (NMS) tekniği ile seçerek her iki metodun güçlü yönlerini birleştirmektedir. Bu seçim işlemi sadece benzersiz olanların kullanımını garanti ettiği için fazlalığı düşürerek model başarısını iyileştirir. Ayrıca end-to-end yaklaşımını benimsediğinden geleneksel region proposal veya anchor boxes gibi teknikleri kullanmamaktadır.

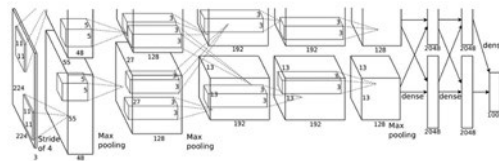
DDQ metodu farklı tespit mimarilerine entegre edilmiştir. Örneğin FCN (fully convolutional network), R-CNN ve DETR ile kullanıldığında performans artışı sağlamaktadır [12].

2.2. Nesne Sınıflama

Nesne sınıflama görüntüdeki anlamlı nesnenin veya nesnelerin bulunduğu sınıfa atama işlemidir. ImageNet yarışmasını 2012-2015 yıllarında kazanan modellerin açıklaması bu bölümde verilmiştir. Bu modellerin her biri derin öğrenmenin temel taşları olarak kabul edilmektedir.

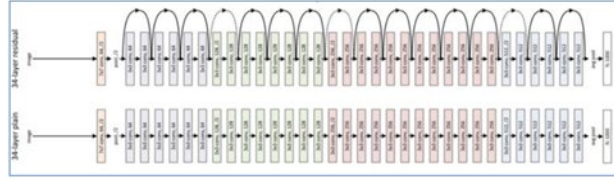
Başarılı sınıflama modelleri aşağıdaki gibidir.

Alex Net, derin öğrenmenin ilk olarak 1998 yılında Yann LeCun tarafından yayınlanmış olan makalede uygulandığı söylene de dünya çapında duyulması 2012 yılında olmuştur. O yıl gerçekleştirilen ImageNet yarışmasını derin öğrenme mimarisi ile tasarlanan AlexNet modeli kazanmıştır. Makalenin adı "ImageNet Classification with Deep Convolutional Networks"(Krizhevsky, Sutskever et al. 2012)' dir. Bu mimari ile bilgisayarlı nesne tanımlama hata oranı %26,2' den %15,4' e düşürülmüştür. Mimari 5 konvolüzyon katmanı, havuzlama katmanı ve 3 tam bağlantılı katmandan oluşmaktadır. Mimari 1000 nesneyi sınıflandıracak şekilde tasarlanmıştır. Filtreler 11x11 boyutunda ve adım kayma sayısı 4 olarak belirlenmiştir. Mimari görüntüsü şekil 2' de verilmiştir.



Şekil 2. Alex Net Mimarisi

ResNet (Residual Neural Network), CNN mimarisidir. Yüz, hatta binlerce konvolüzyon katmanına destek veren bir mimaridir. Daha önceki CNN mimarileri çok katman oluşturmaya izin vermemekteydi. Bu yüzden limitli bir performansları vardı. Ancak daha fazla katman eklenince araştırmacılar "vanishing gradient" problemi ile karşılaşmaktaydılar. ResNet mimarisi bir seri konvolüzyon bloklarına ara input ekleyerek yeni bir tasarım önermiştir. Bu model "skip connection" olarak da adlandırılır. Bu teknik geriye yayılım aşamasında gradient akışını düzeltir. Bu sayede ağıımız 50, 100 hatta 150 katman olabilmektedir. Skip connection ağa hesaplama yükü getirmez. ImageNet 2015 yarışmasının kazananı olmuştur. Microsoft ResNet' in ilk 34 katmanlı ağ mimarisi şekil 3' de gösterilmiştir [1].



Şekil 3. ResNet ilk 34 katmanının ağ mimarisi [1]

Bu mimari residual bloklardan oluşmaktadır. Residual blokta, x girişinin konvolüsyon ReLu-konvolüsyon serisinden sonra bir $F(x)$ sonucu vermektedir. Bu sonuç daha sonra orijinal x girişine eklenir ve $H(x)=F(x)+x$ olarak ifade edilir.

EfficientNet-L2, etiketli verinin bol olması varsayımına dayanan semi-supervised öğrenme yaklaşımını benimsemiş, gürültülü öğrenci eğitimi olarak adlandırılmıştır. Çalışma makine öğrenmesi yöntemi olan self-training metodundan ve saflaştırma için veri sıkıştırma tekniğinden yararlanmaktadır.

İlk olarak, EfficientNet modeli ImageNet’deki etiketli görüntüler ile eğitilir. Ardından model 300M görüntü için sahte etiket üretmesi için öğretmen olarak kullanılır. Ardından EfficientNet’ in daha geniş bir modeli ilk aşamada üretilen etiketlenmiş ve sahte görüntülerin kombinasyonu ile öğrenci olarak eğitilir. Bunun ardından, aynı işlem öğrenciye dropout, stochastic depth ve veri artırımı gibi farklı tiplerde gürültü eklenerek iteratif bir şekilde devam eder. Böylece bu model 2020 yılında ImageNet verisetinde Top-1 başarısı 85,5% olmuştur [10].

SqueezeNet, 18 katmana sahiptir. 2016 yılında sunulmuştur. Amacı daha az parametre ve bilgisayar belleğine sığabilecek boyutlarda olmasıdır. 2017’de Baidu, Xilinx, Imagination Technologies ve Synopsys gibi firmalar SqueezeNet’in akıllı telefon, FPGA ve custom işlemciler gibi düşük güçlü ortamlarda çalışabileceğini göstermişlerdir. AlexNet ile olan benzerliği AlexNet kadar başarılı fakat 50 kat daha az parametre içermesi özelliği ile öne sürülmüştür. AlexNet 240 MB parametre içerirken SqueezeNet 5MB parametre içermektedir. SqueezeNet’i görüntülerin semantic segmentation’ı ve style transfer gibi diğer alanlarda da uygulama imkanı bulunmaktadır[7].

GoogLeNet yapısındaki Inception modüllerinden dolayı karmaşık bir mimaridir. GoogLeNet 22 katmanlı ve %5.7 hata oranı ile ImageNet 2014 yarışmasının kazananı olmuştur. Bu mimari genel olarak ardışık bir yapıda konvolüsyon ve havzalama katmanlarını üst üste istiflemekten uzaklaşan ilk CNN mimarilerinden birisidir. Bu yeni model bellek ve güç kullanımı üzerinde önemli bir yere sahiptir. Katmanların hepsini yığlamak ve filtre eklemek hesaplama ve bellek maliyeti getirir, ezberleme olasılığını artırır. GoogLeNet bu durumun üstesinden gelmek için paralel olarak birbirine bağlı modüller kullanmaktadır. Şekil 4’de GoogLeNet mimarisi verilmiştir.



Şekil 4. GoogLeNet ağ mimarisi [1]

3. Hesaplamalar

Nesne tanıma için kullanılan derin öğrenme mimarisi Faster-RCNN (farklı backbone ağlarla eğitilmiş), YOLO ve DDQ’dir. Nesne tanıma modelleri Pascal VOC ve MS COCO veriseti kullanılarak eğitilmiştir. Eğitim parametreleri aşağıda verilmiş ve her modelin mAP oranları elde edilmiştir. Nesne sınıflandırma için ise VGG16, Inception v3, Mobilenet ve Efficient Net transfer öğrenme yöntemi ile eğitilmiş model parametreleri, başarı ve kayıp oranları grafikleri ile verilmiştir. Kullanılan veriseti ise MS-COCO’dur. Uygulama platformu olarak PyTorch 1.7.0, TensorFlow 1.15.5 ve darknet kullanılmıştır.

3.1. Nesne Tanıma Uygulama Adımları

Pytorch kullanılarak Faster R-CNN algoritması ve YOLO ise darknet kullanılarak elde edilmiştir [9]. DDQ modeli MS COCO 2017 veriseti ile diğer nesne tanıma modelleri Pascal VOC veriseti ile eğitilmiştir.

Pascal VOC verisetinde 20, MS COCO verisetinde ise 80 kategoride nesne bulunmaktadır.

Faster RCNN modeli için sonuçlar 14 epoch sonunda oluşmuştur (10 epoch 0.001 learning rate ve ardından 4 epoch 0.0001 learning rate). İki farklı backbone kullanılmıştır. Bunlar Resnet101 ve VGG16’dir.

ÇİZELGE I. FASTER RCNN EĞİTİM PARAMETRELERİ

Momentum	0.9
Weight decay	0.0005
Dropout	0
Augmentation	Evet

ÇİZELGE II. NESNE TANIMA SONUÇLARI (FASTER RCNN)

Backbone	Başarı(mAP)
Resnet101	74.92%
VGG16	70.62%

<pre> Average Precisions ----- cat : 85.7% car : 83.9% horse : 83.6% bicycle : 80.4% dog : 80.3% bus : 80.0% person : 79.5% motorbike : 77.2% train : 76.7% cow : 74.8% sheep : 71.8% tvmonitor : 70.3% aeroplane : 69.9% diningtable: 68.4% bird : 67.6% sofa : 66.8% bottle : 54.8% chair : 50.5% boat : 49.9% pottedplant: 40.2% ----- Mean Average Precision = 70.62% </pre> <p>(a)</p>	<pre> Average Precisions ----- cat : 89.1% car : 87.5% dog : 87.4% horse : 86.8% train : 84.5% bicycle : 82.3% person : 81.4% cow : 79.6% bus : 79.5% motorbike : 78.3% bird : 78.2% aeroplane : 75.7% sheep : 75.5% tvmonitor : 73.8% sofa : 72.8% diningtable: 66.2% boat : 61.4% bottle : 56.3% chair : 55.3% pottedplant: 46.6% ----- Mean Average Precision = 74.92% </pre> <p>(b)</p>
--	--

Şekil 5. (a) Faster R-CNN VGG16 Backbone Sonucu.

(b) Faster R-CNN ResNet101 Backbone sonucu.

YOLO versiyonları darknet framework' ü kullanılarak eğitilmiştir. YOLO v3 eğitim parametreleri çizelge III' de verilmiştir:

ÇİZELGE III YOLO V3 EĞİTİM PARAMETRELERİ

Learning rate	0.001
Momentum	0.9
Weight decay	0.0005
Width	416
Height	416
Channels	3
Saturation	1,5
Exposure	1,5
Hue	1.0

Sonuçlar ise çizelge IV' deki gibidir:

ÇİZELGE IV YOLO V3 EĞİTİM SONUÇLARI

detections_count = 31363, unique_truth_count = 12032
for conf_thresh = 0.25, precision = 0.78, recall = 0.72, F1-score = 0.75

for conf_thresh = 0.25, TP = 8688, FP = 2521, FN = 3344, average IoU = 60.09 %
IoU threshold = 50 %, used 11 Recall-points mean average precision (mAP@0.50) = 0.729993, or 73.00 %
Total Detection Time: 75 Seconds

YOLO v2 eğitim parametreleri çizelge V' de verilmiştir.

ÇİZELGE V YOLO v2 EĞİTİM PARAMETRELERİ

Learning rate	0.0001
Momentum	0.9
Weight decay	0.0005
Width	416
Height	416
Channels	3
Saturation	1,5
Exposure	1,5
Hue	1.0

Sonuçlar ise çizelge VI' da verilmiştir.

ÇİZELGE VI YOLO v2 EĞİTİM SONUÇLARI

detections_count = 102694, unique_truth_count = 12032
for conf_thresh = 0.25, precision = 0.62, recall = 0.72, F1-score = 0.67
for conf_thresh = 0.25, TP = 8674, FP = 5306, FN = 3358, average IoU = 48.11 %
IoU threshold = 50 %, used 11 Recall-points mean average precision (mAP@0.50) = 0.713296, or 71.33 %
Total Detection Time: 49 Seconds

Average Precisions	Average Precisions
cat : 84.23%	cat : 84.00%
car : 82.97%	car : 79.61%
horse : 80.74%	horse : 79.23%
bicycle : 83.35%	bicycle : 81.04%
dog : 83.04%	dog : 83.33%
bus : 78.42%	bus : 79.03%
person : 77.31%	person : 73.71%
motorbike : 79.58%	motorbike : 81.56%
train : 78.76%	train : 82.64%
cow : 75.86%	cow : 74.64%
sheep : 72.90%	sheep : 73.52%
tvmonitor : 68.42%	tvmonitor : 71.23%
aeroplane : 80.38%	aeroplane : 74.00%
diningtable : 70.25%	diningtable : 69.30%
bird : 74.11%	bird : 72.18%
sofa : 72.54%	sofa : 68.30%
bottle : 53.37%	bottle : 46.83%
chair : 54.25%	chair : 51.77%
boat : 62.91%	boat : 57.50%
pottedplant : 46.60%	pottedplant : 43.17%

(a)

(b)

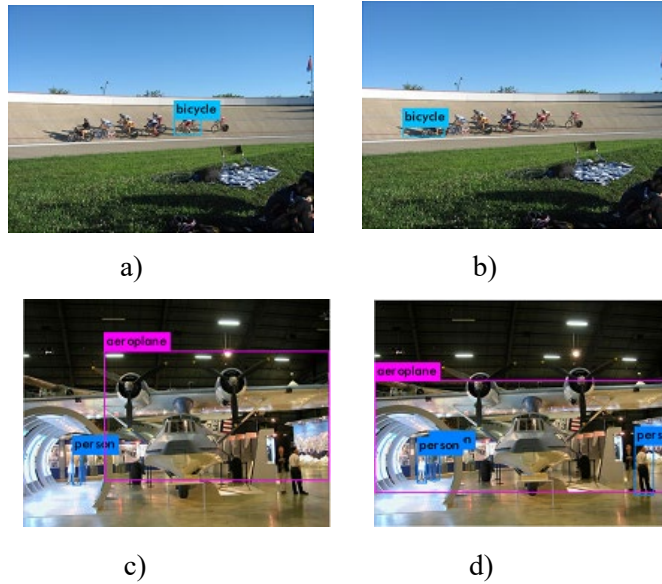
Şekil 6. (a) YOLO v3 Sınıf mAP Oranları
(b) YOLO v2 Sınıf mAP Oranları

DDQ modeli ise MS-COCO veriseti üzerinde DETR mimarisi ile birlikte kullanılmıştır. Özet sonuç tablosu Çizelge VII’ de verilmiştir.

ÇİZELGE VII NESNE TANIMA MODEL BAŞARILARI

Model		mAP@0.50
DDQ (DETR)		68.90%
YOLO v2		71.33%
YOLO v3		73.00%
Faster R-CNN	Resnet 101 Backbone	74.92%
	VGG16 Backbone	70.62%

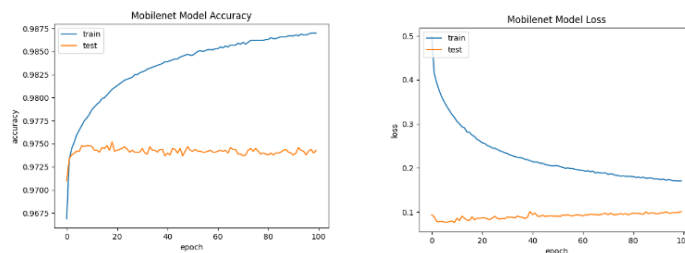
Bazı nesne tanıma görsel sonuçları şekil VII’ de verilmiştir.



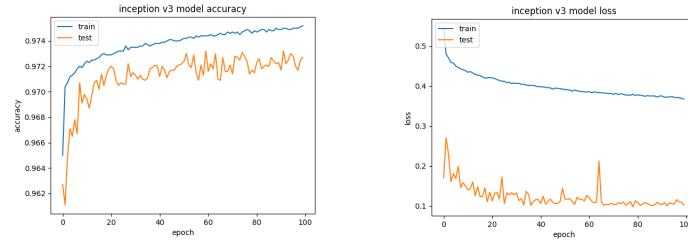
Şekil 7. (a,b,c,d) Örnek Nesne Tanıma Sonuçları

3.2. Nesne Sınıflandırma Uygulama Adımları

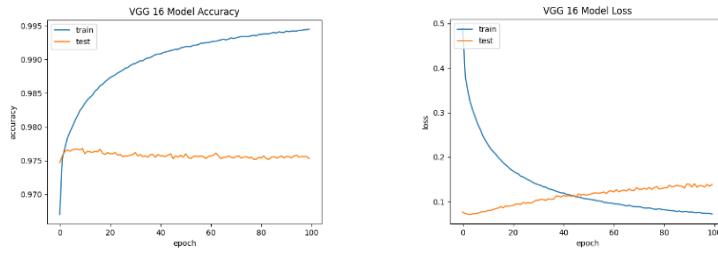
Bu uygulama adımında transfer öğrenme yöntemiyle imajenet veri setinde eğitilmiş 4 modelin (VGG 16, Mobilenet ve Inception v3, Efficient Net) MS COCO veri setine uyarlanması ile oluşmaktadır. MS COCO veriseti 80 kategoride nesne ve 118.287 adet eğitim görüntüsü içermektedir. Nesne tanıma, nesne sınıflama ve görüntü segmentasyonunda kullanılmaktadır. Tensorflow 1.15.5 versiyonu kullanılarak CUDA GPU üzerinde ilk 80.000 görüntü eğitim için, kalan 38.287 adet görüntü ise doğrulama seti olarak kullanılmıştır. Imajenet verisetinde eğitilmiş modellerden son tam bağlı katman çıkarılarak 2 tam bağlı katman eklenmiştir. İlk eklenen tam bağlı katman 512 düğüm içermekte ardından Batch Normalization, Dropout katmanı ve 80 çıkıtlık tam bağlı katman ile modellerimiz elde edilmiştir. Şekil 8-9-10’ da modellerin başarı ve kayıp grafikleri verilmiştir.



Şekil 8. Mobilenet Modeli Eğitim Sonucu



Şekil 9. Inception v3 Modeli Eğitim Sonucu



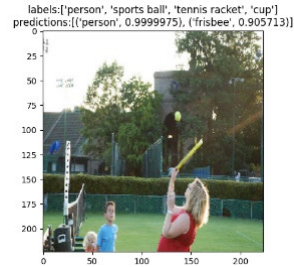
Şekil 10. VGG 16 Model Eğitim Sonucu

ÇİZELGE VIII. NESNE SINIFLANDIRMA SONUÇLARI

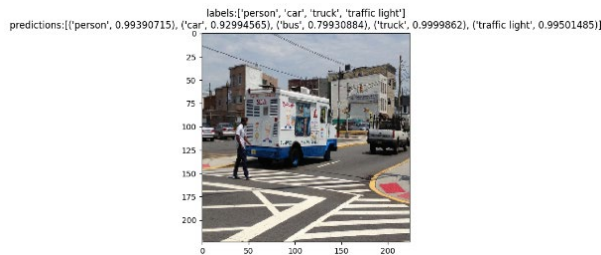
Model	Top-1 Başarısı	Top-5 Başarısı
VGG 16	0.5480	0.8676
MobileNet	0.5980	0.8648
Inception	0.5784	0.8692
Efficient Net B0	0.5156	0.8868



(a)



(b)



(c)

Şekil 13. (a, b, c) Örnek Nesne Sınıflandırma Sonuçları

4. SONUÇ VE ÖNERİLER

Bu araştırmada nesne tanıma ve sınıflandırmada mevcut derin öğrenme yöntemlerinin kısmen başarılı sonuçlar verdiği belirlenmiştir. Bununla birlikte nesne tanıma ve sınıflandırmada kullanılan derin öğrenme yöntemlerinin küçük nesnelere tanıma ve sınıflandırma performanslarının artırılması gerektiği sonucuna varılmıştır. İleriye dönük yapılacak çalışmalar ve araştırmalar için mevcut derin öğrenme yöntemlerinin küçük nesnelere tanıması ve sınıflandırması için modellerin yeniden düzenlenmesi ve geliştirilmesi gerektiği sonucuna varılmıştır.

Kaynaklar

- [1] Özkan, İ. N. İ. K., and Erkan Ülker. "Derin öğrenme ve görüntü analizinde kullanılan derin öğrenme modelleri." *Gaziosmanpaşa Bilimsel Araştırma Dergisi* 6.3 (2017): 85-104..
- [2] F. Doğan ve İ. Türkoğlu, "Derin Öğrenme Modelleri ve Uygulama Alanlarına İlişkin Bir Derleme", *DÜMF Mühendislik Dergisi*, c. 10, sy 2, ss. 409-445, Haz. 2019, doi: 10.24012/dumf.411130.
- [3] A. Rosebrock, "Deep Learning for Computer Vision with Python. 2 - Practitioner Bundle".
- [4] S. Kevser, "MAKİNE ÖĞRENMESİ VE DERİN ÖĞRENME: NESNE TANIMA UYGULAMASI".
- [5] L. Jiao vd., "A Survey of Deep Learning-Based Object Detection", *IEEE Access*, c. 7, ss. 128837-128868, 2019, doi: 10.1109/ACCESS.2019.2939201.
- [6] F. G. Tan, A. S. Yüksel, E. Aydemir, ve M. Ersoy, "DERİN ÖĞRENME TEKNİKLERİ İLE NESNE TESPİTİ VE TAKİBİ ÜZERİNE BİR İNCELEME", *European Journal of Science and Technology*, May. 2021, doi: 10.31590/ejosat.878552.
- [7] "SqueezeNet". *Wikipedia*, 29 Ocak 2024. *Wikipedia*, <https://en.wikipedia.org/w/index.php?title=SqueezeNet&oldid=1200466107>.
- [8] Terven, J., & Cordova-Esparza, D. (2023). *A comprehensive review of YOLO: From YOLOv1 to YOLOv8 and beyond*. *arXiv preprint arXiv:2304.00501*.
- [9] B. Trzynadlowski, "trzy/FasterRCNN". 09 Şubat 2024. Erişim: 14 Şubat 2024. [Çevrimiçi]. Erişim adresi: <https://github.com/trzy/FasterRCNN>
- [10] T. Turay and T. Vladimirova, "Toward Performing Image Classification and Object Detection With Convolutional Neural Networks in Autonomous Driving Systems: A Survey," in *IEEE Access*, vol. 10, pp. 14076-14119, 2022, doi: 10.1109/ACCESS.2022.3147495.
- [11] Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A., & Zagoruyko, S. (2020, August). End-to-end object detection with transformers. In *European conference on computer vision* (pp. 213-229). Cham: Springer International Publishing.
- [12] Zhang, S., Wang, X., Wang, J., Pang, J., Lyu, C., Zhang, W., ... & Chen, K. (2023). Dense distinct query for end-to-end object detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 7329-7338).

Classification of Brand Perception Using Random Forest: Brand Preference, Brand Loyalty and Brand Trust

Muhammed İkbal Yılmaz¹, Özge Yıldız² and Şahika Gökmen^{3,4}

^{1,2,3} Department of Econometrics, Faculty of Economics and Administrative Sciences,
Hacı Bayram Veli University, Ankara, Turkey.

⁴ Statistics Department, Uppsala University, Uppsala, Sweden.

Abstract

Brands enable strategy development, competitive advantage, and audience connection in modern society. This study addresses gaps in Brand Preference (BP), Brand Loyalty (BL), and Brand Trust (BT) among university students using Machine Learning (ML), particularly Random Forest (RF) Classification. Data from Ankara Hacı Bayram Veli University (Spring 2023/2024) is analyzed with decision tree classifiers. A meta-predictor reduces overfitting, while the 'Brand Perception Scale', with 22 independent variables (x_i), assesses BP, BL, and BT.

Keywords: Machine Learning Classification, Brand Trust, Brand Loyalty, Brand Preference, Random Forest, Decision Trees

INTRODUCTION

In Artificial Intelligence (AI), McCulloch et al. (1943) introduced the artificial neuron model, laying the foundation for Deep Learning (DL). Turing's (1950) query, "Can Machines Think?" hinted at machine emulation of human responses. On August 31, 1955, McCarthy and colleagues coined "AI" while defining machine learning capabilities. Samuel (1959) later defined ML as a computer's ability to improve from experience. Initially, brands served as simple identifiers (Davis, 1964) but have since evolved into complex assets essential for growth, rooted in 'brand image' factors like quality and reputation (Aaker, 1996; Keller, 2001). BT is linked to belief and behavior (Moorman et al., 1992), while early identification of BL factors strengthens brand ties (Schee, 2010). Key drivers of BL include quality (Zehir et al., 2011), and brands have become complex, influential assets (Bastos and Levy, 2012). BP reflects individual preferences shaped by multiple factors (Ebrahim, 2013; Yang et al., 2015). While research on BL clustering is limited, BL aligns with repurchasing loyalty (Gumus, 2016; Li and You, 2021). BT has become critical for customer loyalty and long-term relationships (Lee et al., 2014; Alhaddad, 2015), with technological innovation as a strategic asset (Morgan-Thomas and Veloutsou, 2013). Despite extensive ML-brand research (Chaudhary et al., 2016; Sayli et al., 2016; Pamuksuz and Yun, 2021; Dong, 2023), studies integrating BP, BL, and BT are rare. To address this, we applied a RF multi-class classification model to clarify relationships among these constructs.

DECISION TREES WITH MACHINE LEARNING

Decision Trees

This study employed proportional quota sampling to select a representative sample of 407 undergraduate students from the Econometrics and Economics departments at Ankara Hacı Bayram Veli University, each exposed to brands. The dataset was split into a 75% training and 25% test set, revealing significant class imbalance (Figure 1), which complicates classification and necessitates adjustments to enhance predictive accuracy. Chawla et al. (2002) demonstrated the efficacy of SMOTE (Synthetic Minority Over-sampling Technique) for improving minority class prediction in imbalanced datasets. Pears et al. (2014) observed that while synthetic data generation can balance classes, it may risk overfitting and potential information loss. SMOTE mitigates these risks by interpolating within the minority class data, ensuring synthetic samples align with the original data without introducing external information. Here, x_i represents a minority class sample, with J selected via the k-nearest neighbors method (Tesfahun and Bhaskari, 2013). In this approach, λ is a random variable within $[0,1]$, and the difference vector is $\Delta = J - x_i$. A new synthetic data point is generated as $y_{new} = x_i + \lambda \cdot \Delta$, creating a sample between x_i and J . If $\lambda = 1$, y_{new} matches J ; if $\lambda = 0$, y_{new} remains x_i . Quinlan (1986) made key contributions to decision tree algorithms, essential for RF classification, which begins with tree construction. In RF, data and variables are randomly sampled as given in equation (1) to create decision trees, each based on unique subsets of the training dataset:

$$D_n^{(j)} = \left\{ \left(x_1, y_1^{(j)} \right), \left(x_2, y_2^{(j)} \right), \dots, \left(x_n, y_n^{(j)} \right) \right\} \quad (1)$$

where $y_i^{(j)} = h(x_i; \theta_j, D_n^{(j)})$. Each tree $h(x_i; \theta_j, D_n^{(j)})$ predicts the class of x_i using tree-specific split criteria. The j -th tree's prediction is $y_i^{(j)}$. Final classification of x_i is determined by majority vote across M trees, with the highest voted class selected as:

$$\hat{y}(x_i) = \underset{c \in C}{\operatorname{argmax}} \left\{ \sum_{j=1}^M 1 \left[h \left(x_i; \theta_j, D_n^{(j)} = c \right) \right] \right\} \quad (2)$$

In equation (2), C denotes all possible classes (As given in Table 1), and $h(x_i; \theta_j, D_n^{(j)})$ represents tree j 's prediction of x_i based on parameters θ_j . This ensemble approach combines multiple trees for a robust model (Kursa, 2014).

Machine Learning

Data analysis was conducted in Python 3.12.5 (Windows 64-bit) on Jupyter Notebook, using a system with an 11th Generation Intel Core i5-11260H, 24 GB RAM, and NVIDIA GeForce RTX 2050ti GPU (Micro-Star International Co.). Key libraries included pandas for data handling, numpy for computations, imbalanced-learn.over_sampling.SMOTE for class imbalance, and sklearn.ensemble.RandomForestClassifier for classification modeling, with sklearn.metrics.classification_report and sklearn.metrics.confusion_matrix for evaluation. BP, BL, and BT models were based on averages from 7, 4, and 3 questions, respectively. Internal consistency (Cronbach's Alpha, α) values were 0.7417 for BP, 0.7821 for BL, and 0.7696 for BT, with an overall reliability of $\alpha = 0.8399$.

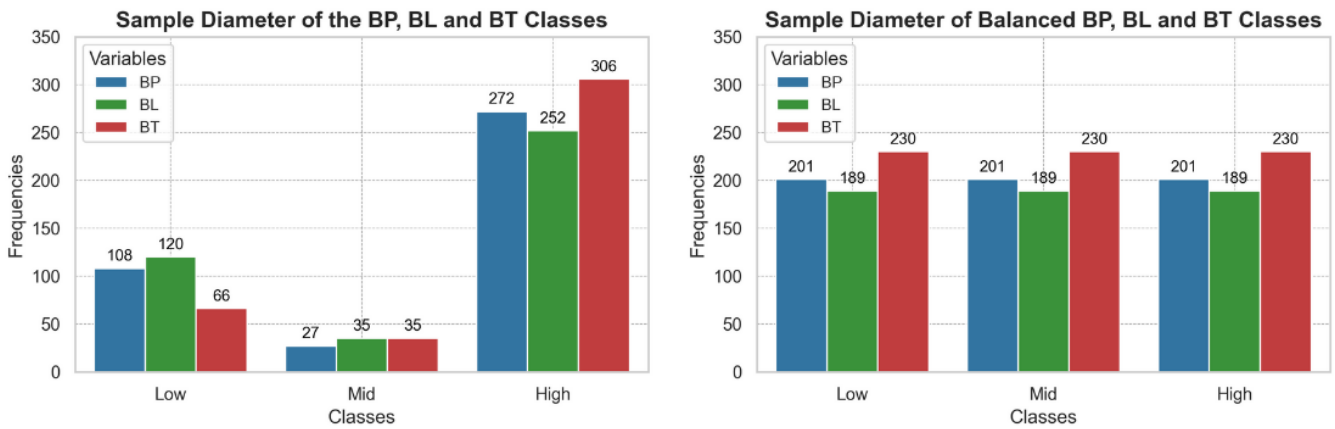


Figure 1. Original Dataset and Balanced Dataset

Table 1 details y_i and x_i in the dataset. Three classification models categorized observations into Low, Mid, and High classes using the RF algorithm with Breiman's bootstrap resampling on training set ($D_n(D_n^{(j)} \sim \text{Bootstrap}(D_n))$) and a voting mechanism (Breiman, 1996). Each model included 22 independent variables. RF, introduced by Breiman (1984, 2001), is a prominent ML algorithm advancing decision tree methods (Quinlan, 1986).

Table 1. Variables and Classes

Variable Type	Variables	Classes
Dependent (y_i)	BP	Low, Mid, High
	BL	Low, Mid, High
	BT	Low, Mid, High
Independent (x_i)	Age, Gender, Department, Year, BandR, Place, Income, Socialmedia, Ishopping, BrandMonitoring, Price, Quality, Timeofuse, ProductionDate, Seasonality, Material, Discount, Fashion, Design, Psychologicalimpulses, Country, Recognizability	

Table 2 presents classification performance pre- and post-SMOTE, showing BP, BL, and BT model differences. BP achieved high Precision (0.83) and Recall (0.78), while BL and BT had Recalls of 0.74 and 0.70. BP and BL excelled in the High class (Recalls 0.88 and 0.91), and both BP and BT showed balanced Precision and Recall in the Low class (0.82 and 0.89 for BP; 0.85 and 0.81 for BT).

Table 2. Performance Metrics for Imbalanced and Balanced Data

Imbalanced					Balanced				
Variables	Classes	Precision	Recall	F1-score	Variables	Classes	Precision	Recall	F1-score
\hat{y}_{BP}	Low	0.73	0.42	0.54	\hat{y}_{BP}	Low	0.82	0.89	0.85
	Mid	1.00	0.00	0.00		Mid	0.83	0.78	0.80
	High	0.77	0.94	0.85		High	0.91	0.88	0.89
Accuracy: 0.76					Accuracy: 0.85				
\hat{y}_{BL}	Low	0.80	0.48	0.60	\hat{y}_{BL}	Low	0.80	0.83	0.81
	Mid	1.00	0.00	0.00		Mid	0.86	0.74	0.79
	High	0.73	0.95	0.83		High	0.82	0.91	0.86
Accuracy: 0.75					Accuracy: 0.82				
\hat{y}_{BT}	Low	1.00	0.31	0.48	\hat{y}_{BT}	Low	0.85	0.81	0.83
	Mid	1.00	0.00	0.00		Mid	0.81	0.70	0.75
	High	0.78	1.00	0.88		High	0.75	0.87	0.80
Accuracy: 0.79					Accuracy: 0.80				

Figure 2’s confusion matrices display model predictions on the balanced dataset. BP improved with some Mid and High reassignments to Low. BL accuracy increased with refined Low and High predictions, while BT showed strong Mid class performance. Adding High and Mid confidence students to Low BT supported model balance.

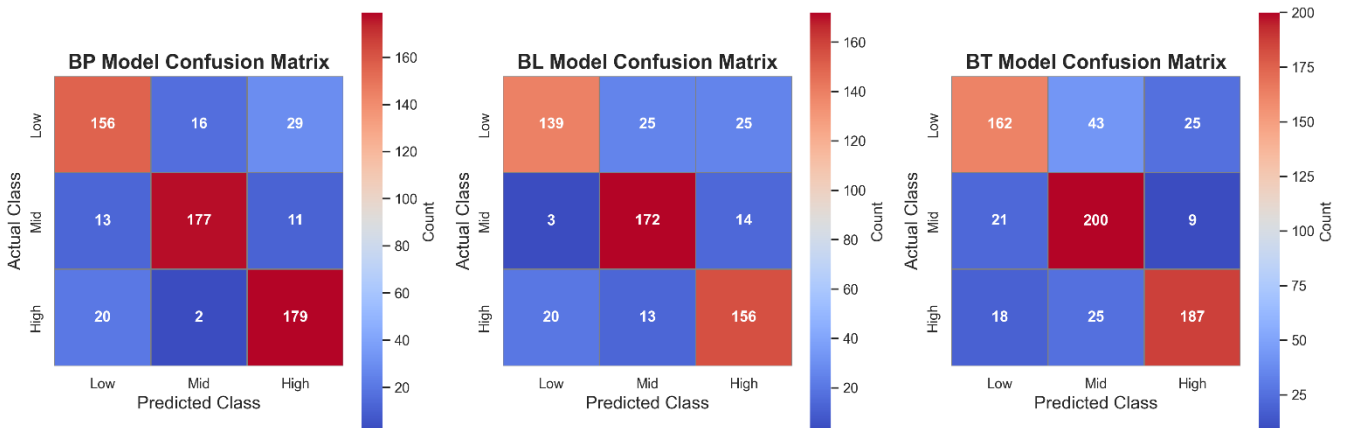


Figure 2. Confusion Matrices

Figures 3, 4, and 5 show decision trees where entropy at each node measures system uncertainty (Shannon, 1948; Tian et al., 2019). Social media (≤ 3.5) is primary BP driver, especially affecting Mid and High classes along with fashion. Low social media use but High quality (≤ 4.5) links to stronger BP, while Low brand monitoring (≤ 1.5) and seasonality (≤ 2.5) link to Mid BP. Lower fashion (≤ 3.5) and design (≤ 4.5) scores associate with Low BP. Findings show fashion and social media significantly influence BP: students engaged with fashion, social media, or higher income tend toward High BP, while less engaged students align with Mid or Low BP. Financial awareness also impacts BP, especially for Low social media users (Figure 3). The BL decision tree in Figure 4 identifies design perception as the main factor influencing BL. Students with lower design appreciation tend toward Low BL, while higher ratings align with Mid or High BL. Online shopping interactions boost loyalty, though this effect declines for older products. Production recency positively impacts BL, with discount perceptions further boosting loyalty, especially for students active on social media. Recognizability and income also influence BL, with higher incomes linked to greater loyalty. Key BL factors include design, shopping behavior, and production recency, while low design interest and online engagement relate to Low BL.

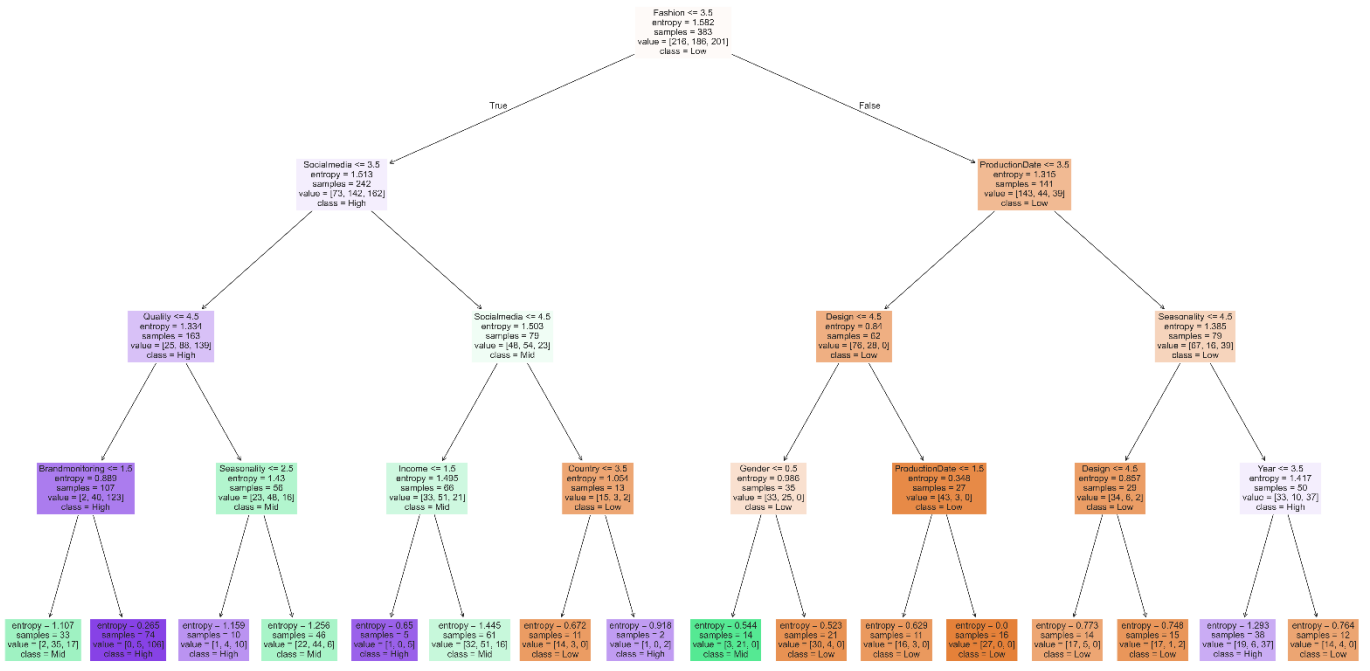


Figure 3. BP Model Decision Tree

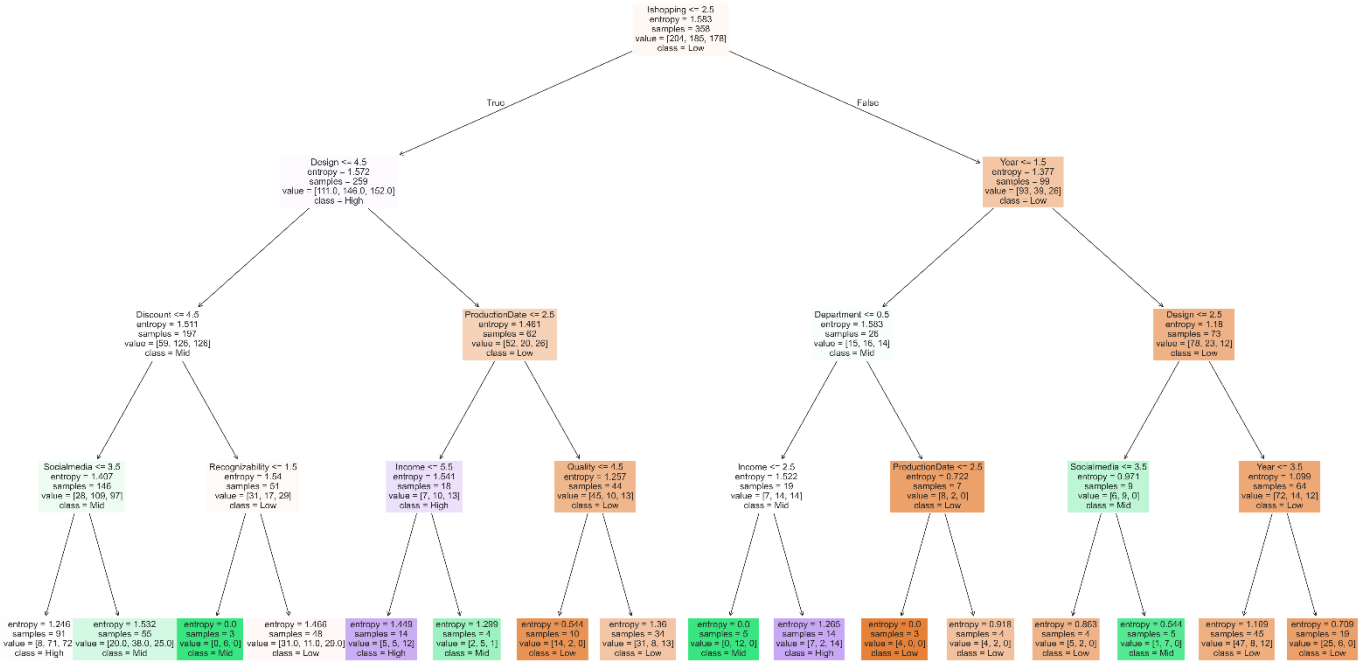


Figure 4. BL Model Decision Tree

The BT decision tree in Figure 5 shows online shopping as the top trust factor, with lower engagement linked to Low BT and higher engagement showing varied trust levels. Design perception also impacts BT: low appreciation aligns with Low BT, high appreciation with Mid or High BT. Recognizability strengthens trust, especially for high-income students, while Low income students tend to remain in Low BT. Ignoring production dates and design lowers trust among online shoppers; even quality-conscious students may report reduced trust when production dates are overlooked. Thus, trust is shaped by online shopping, design, recognizability, and economic factors.

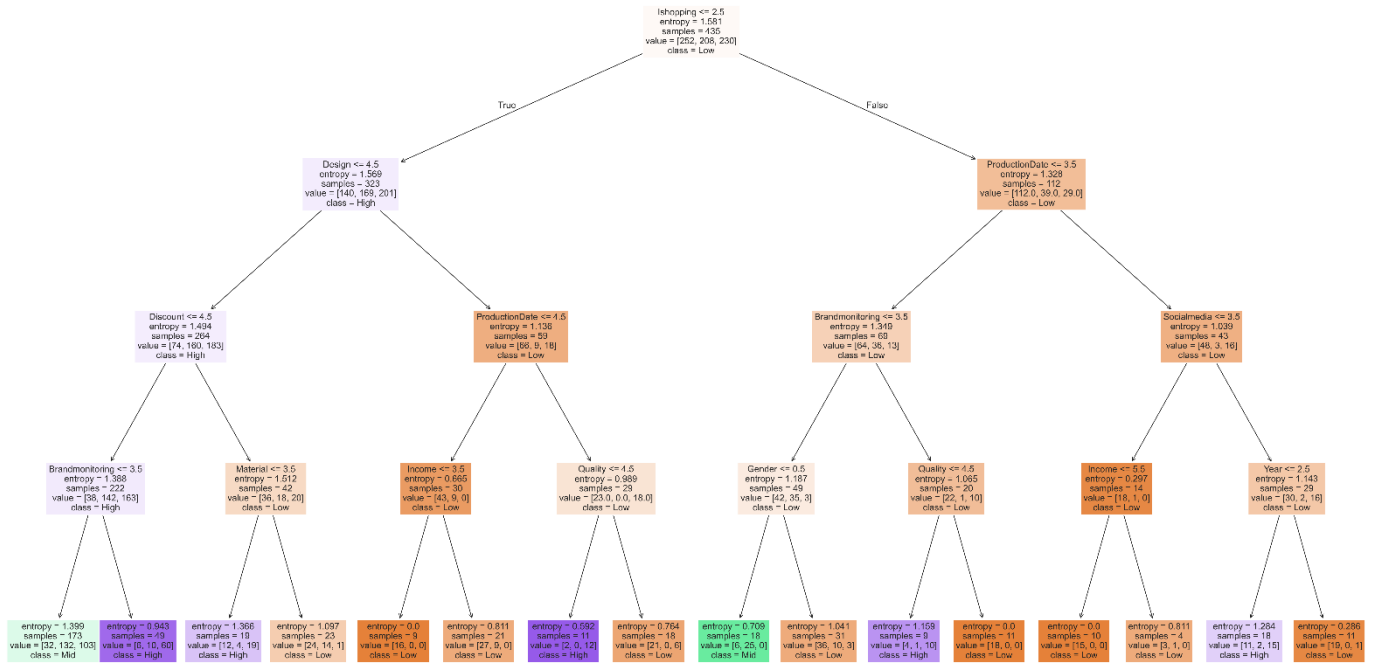


Figure 5. BT Model Decision Tree

Feature importance values for all models appear in Figure 6. The results align with decision tree nodes, demonstrating consistency across model visuals. Our findings provide several conclusions and recommendations. Figure 6’s Feature Importance Plots show that, as in Ebrahim’s thesis, BP is shaped by a range of factors across 22 variables (Ebrahim, 2013). In contrast, BL and BT models emphasize brand awareness, linking it strongly to loyalty and trust (Oliver, 2010; Chen, 2021). Feature importance analysis confirms social media’s role in shaping brand perception and enhancing marketing impact among students (Pamuksuz et al., 2021). Technological innovations and gender-specific campaigns align with consumer preferences, while materialistic values notably impact BL and BT. Additionally, design, fashion, and social media blend functionality with aesthetics, fostering confidence and proactive loyalty (Oliver, 2010).

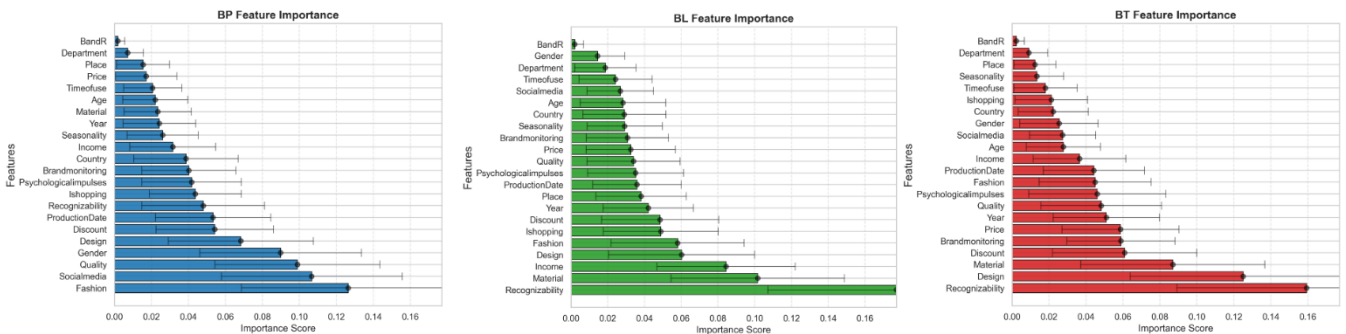


Figure 6. Feature Importance Plots in Balanced BP, BL, and BT Models

RESULTS AND DISCUSSION

Our findings align with existing literature on BP, identifying strategic factors influencing BL and BT among young consumers, such as online shopping, social media engagement, student demographics, product quality, and innovation (Zehir et al., 2011; Gumus, 2016; Bastos and Levy, 2012). Companies targeting students are advised to adopt personalized strategies using social media algorithms and AI tools, particularly AI-driven recommendation systems that enhance trust in online shopping (Chawla et al., 2002; Yang et al., 2015). These insights reinforce the quality-loyalty relationship highlighted by Zehir et al. (2011) and Gümüş (2016), with Bastos and Levy (2012) noting the growing complexity of brands, making accessible and diverse online options crucial. Yang et al. (2015) further stress the importance of integratrating personality traits into BP analysis through social media data and psychometric surveys. Methodologically, SMOTE effectively addresses class imbalance (Chawla et al., 2002), though binary classification and alternative resampling methods may be beneficial in cases of severe imbalance. RF algorithms are

effective in determining feature importance, and combining them with coefficient-based models, such as logistic regression, improves interpretability (Ebrahim, 2013). Probabilistic methods, including Bayesian Inference and Maximum Likelihood Estimation (MLE), also enhance performance in imbalanced datasets (Bastos and Levy, 2012). In the original dataset, accuracy values for BP, BL, and BT were 0.76, 0.75, and 0.79, respectively, increasing by 0.09, 0.07, and 0.01 post-balancing. Future studies should expand sample diversity by including students from foundation universities and explore techniques like ADASYN, TomekLinks, and other resampling methods for low sample sizes (Yang et al., 2015). Ensemble methods with class weights and multi-class classifiers are recommended for additional accuracy improvements. This study builds on our previous research (Yılmaz et al., 2024), demonstrating the effectiveness of the Random Forest algorithm in multi-class classification. Consistent with Ebrahim (2013), our findings confirm that BP is shaped by multiple factors, establishing a foundation for scalable marketing solutions and suggesting that similar future studies are likely to yield consistent results.

References

- Aaker, D. A. (1996). *Managing Brand Equity*. New York: The Free Press.
- Alhaddad, A. (2015). Perceived quality, brand image, and brand trust as determinants of brand loyalty. *Journal of Research in Business and Management*, 3(4), 1–8.
- Bastos, W., & Levy, S. J. (2012). A history of the concept of branding: Practice and theory. *Journal of Historical Research in Marketing*, 4(3), 347–368.
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24, 123–140.
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>.
- Breiman, L., Friedman, J., Olshen, R. A., & Stone, C. J. (1984). *Classification and regression trees* (1st ed.). Chapman and Hall/CRC. <https://doi.org/10.1201/9781315139470>.
- Chaudhary, A., Kolhe, S., & Kamal, R. (2016). An improved random forest classifier for multi-class classification. *Information Processing in Agriculture*, 3(4), 215–222.
- Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321–357.
- Davis, S. M. (1964). *Brand asset management: Profitable growth through your brands* (2nd ed.). San Francisco: Jossey-Bass.
- Dong, Y. (2023). Application of user preference mining algorithms based on data mining and social behavior in brand building.
- Ebrahim, R. S. (2013). A study of brand preference: An experiential view (Doctoral dissertation, Brunel University Brunel Business School).
- Gümüş, I. (2016). Brand gender, brand personality, and brand loyalty relationship. *Communication in Mathematical Modeling and Applications*, 1(2), 8–41.
- Keller, K. L. (2001). Building customer-based brand equity. *Marketing Management*, 10(2), 14–19.
- Kursa, M. B. (2014). Robustness of random forest-based gene selection methods. *BMC Bioinformatics*, 15, 1–8.
- Lee, J. L., James, J. D., & Kim, Y. K. (2014). A reconceptualization of brand image. *International Journal of Business Administration*, 5(4), 1.
- McCulloch, W. S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, 5, 115–133.
- McCarthy, J., Minsky, M. L., Rochester, N., & Shannon, C. E. (2006). A proposal for the Dartmouth summer research project on artificial intelligence, August 31, 1955. *AI Magazine*, 27(4), 12.
- Moorman, C., Zaltman, G., & Deshpande, R. (1992). Relationships between providers and users of market research: The dynamics of trust within and between organizations. *Journal of Marketing Research*, 29(3), 314–328.

- Morgan-Thomas, A., & Veloutsou, C. (2013). Beyond technology acceptance: Brand relationships and online brand experience. *Journal of Business Research*, 66(1), 21-27.
- Oliver, R. L. (2010). Consumer brand loyalty. *Wiley international encyclopedia of marketing*. Chichester, West Sussex, UK: Wiley-Blackwell.
- Pamuksuz, U., Yun, J. T., & Humphreys, A. (2021). A brand-new look at you: Predicting brand personality in social media networks with machine learning. *Journal of Interactive Marketing*, 56, 1-15.
- Pears, R., Finlay, J., & Connor, A. M. (2014). Synthetic minority over-sampling technique (SMOTE) for predicting software build outcomes. *arXiv preprint*, arXiv:1407.2330.
- Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1, 81-106.
- Samuel, A. L. (1959). Some studies in machine learning using the game of checkers. *IBM Journal of Research and Development*, 3(3), 210-229.
- Saylı, A., Ozturk, I., & Ustunel, M. (2016). Brand loyalty analysis system using K-means algorithm. *Journal of Engineering Technology and Applied Sciences*, 1(3), 107-126.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379-423.
- Tesfahun, A., & Bhaskari, D. L. (2013). Intrusion detection using random forests classifier with SMOTE and feature reduction. *2013 International Conference on Cloud & Ubiquitous Computing & Emerging Technologies* (pp. 127-132). IEEE.
- Tian, J., Liu, L., Zhang, F., Ai, Y., Wang, R., & Fei, C. (2019). Multi-domain entropy-random forest method for the fusion diagnosis of inter-shaft bearing faults with acoustic emission signals. *Entropy*, 22(1), 57.
- Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433-460. <https://doi.org/10.1093/mind/LIX.236.433>.
- Schee, B. A. (2010). Students as consumers: Programming for brand loyalty. *Services Marketing Quarterly*, 32(1), 32-43.
- Yılmaz, M. İ., Yıldız, Ö., & Gökmen, Ş., (2024) Marka Tercihine Etki Eden Faktörler: Ankara Hacı Bayram Veli Üniversitesi Ekonometri ve İktisat Bölümü Lisans Programı Öğrencileri Üzerine Bir Uygulama. 5. *International Applied Statistics Congress (UYIK-2024)*, 864-874.
- Yang, C., Pan, S., Mahmud, J., Yang, H., & Srinivasan, P. (2015). Using personal traits for brand preference prediction. *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*.

Acknowledgements

We thank TÜBİTAK BİDEB, the Turkish Statistical Association, IDSSC'24, Ankara Hacı Bayram Veli University, and Sezin Yılmaz for their support in data acquisition and collection. Our gratitude extends to all contributors to this study.

Conflict of Interest

The authors declare that there is no conflict of interest.

Yapay Sinir Ağları ile İstanbul Elektrik Talep Tahmini

Hayriye Yasak ÖZKAL^{1*}, Buse KAYLAN², Meltem SİPAHİ³, Muammer Altan ÇAKIR⁴

^{1,2,3} Veri Mühendisliği ve İş Analitiği Anabilim Dalı, İstanbul Teknik Üniversitesi, İstanbul, Türkiye

⁴Fizik Mühendisliği Bölümü, İstanbul Teknik Üniversitesi, İstanbul, Türkiye

Özet

Elektrik tüketimine yönelik arz, talep ilişkisi incelendiğinde, elektrik yükü, fazla tahmin edilirse gereksiz enerji alınmasına ve fazla rezerv oluşturulmasına yol açarken; yük tahminleri gerçekleşen tüketimden düşükse, enerji arzının yetersiz kalmasına, kesintilerin artmasına, dolayısıyla sistem kararsızlığına neden olabilmektedir. Elektrik talep tahminlerinde optimizasyon sağlanması, kaynakların verimli kullanılması, kesintilerin azaltılması vb. hususlar enerji üreticileri, dağıtım şirketleri ve tüketiciler için önemli bir sorun haline gelmektedir. Bu çalışma ile sıcaklık, nem gibi mevsim değişkenlerinin yanı sıra resmî tatiller ile hafta tatillerinin İstanbul ili bazında elektrik tüketimine etkisi, EPİAŞ şeffaflık raporları ile birlikte değerlendirilmiş, bahse konu parametrelerin tüketim tahminlerindeki etkisi analiz edilerek tahmin doğruluğunu artırmak amaçlanmıştır. Veri seti zaman serisi özellikleri açısından incelenmiş olup NARX modelleri ve Prophet / LSTM hibrid modelleri uygulandıktan sonra en düşük hata metriklerini ve en yüksek tahmin doğruluğunu sağlayan nihai model seçilmiştir. İlgili model sonucunda İstanbul ili bazında elektrik tüketiminin nem, sıcaklık, gün içi saat ve tatil günleri ile ilişkili olduğu görülmüştür.

Anahtar Kelimeler – Elektrik talep tahmin, Makine öğrenmesi algoritmaları, Prophet/LSTM modeli, NARX model, Çekişmeli Üretici Ağ (Generative Adversarial Networks (GAN))

GİRİŞ

Ülkemizde yıllara göre artan nüfus, sanayi ve teknoloji temelli ticari faaliyetler ile teknolojiye olan talebin artması gibi nedenlerle elektrik enerjisine olan talep de artmaktadır. Elektrik enerjisinin doğrudan depolanamaması elektrik üretimi, iletimi ve tüketimi ile ilgili planlama ve operasyonel stratejileri doğrudan etkilemektedir.

Tüketilen elektrik miktarının üretilenden fazla olması veya üretilen elektriğin tüketimden fazla olması, enerji dengesizliğine yol açmaktadır. Buna göre yük tahminleri, fazla tahmin edilirse gereksiz enerji alınmasına ve fazla rezerv oluşturulmasına yol açarken; yük tahminleri gerçekleşen tüketimden düşükse, enerji arzının yetersiz kalmasına ve kesintiler nedeniyle sistem kararsızlığına yani enerji dengesizliğine neden olabilmektedir. Bu nedenle enerji piyasalarının verimli işlemesi, kaynak israfının önlenmesi ve maliyetlerin azaltılması açısından tüketici talebinin doğru tahmin edilmesi amacıyla arz ve talep dengesinin sağlandığı doğru yük tahminleri, elektrik piyasalarında alınan pek çok kararın temelini oluşturmaktadır. (Akman ve diğerleri, 2018)

Elektrik ticareti, uzun dönemli ikili anlaşmaların yanı sıra, arz-talep yönetiminde sistem etkinliğinin sağlanması için Gün Öncesi Piyasası (GÖP) ve Gün İçi Piyasası (GİP) üzerinden yapılmaktadır. GÖP, bir gün öncesinde üretim ve tüketim planlaması yaparak arz-talep dengesini sağlamaya çalışmaktadır. Katılımcılar, bir gün sonrası için saatlik alış ve satış tekliflerini sunmakta, EPİAŞ ise bu teklifleri eşleştirmektedir ancak GÖP ile sağlanamayan arz-talep dengesi, GİP ve Dengeleme Güç Piyasası aracılığıyla sağlanmaktadır. GİP, aynı gün içinde enerji dengesizliklerini düzeltmek amacıyla esnek ve anlık işlemler yapılmasını sağlamaktadır. Böylece hem üreticiler hem de tüketiciler için arz-talep dengesi anlık olarak korunarak enerji ticaretinin daha verimli hale getirilmeye çalışılmaktadır.

Bahse konu piyasalarda elektrik tüketim tahmini hissedilen sıcaklık, elektrik dağıtım ve iletim kaybı gibi belirsiz faktörlerin yanı sıra hafta tatili veya resmî tatillerin de tüketime olan etkisi dikkate alınmaktadır. Bu çalışma ile elektrik tüketimini etkileyen değişkenlerin tahmin doğruluğu üzerindeki dönemsel etkisi incelenecek olup bahse konu

değişkenlerin tüketim trendine etkisi analiz edildikten sonra geliştirilen derin öğrenme modeli ile tahmin doğruluğunu artırmak amaçlanmaktadır.

ARAÇ VE YÖNTEMLER

Malzeme

Kullanılacak modelin tespitine yönelik literatür incelendiğinde; geçmiş veriler ve uygun tahmin modelleri kullanılarak, uzun vadeli (yıl), orta vadeli (ay), kısa vadeli (gün/saat) ve gerçek zamanlı (dakika/saniye) tahminler yapılabileceği; piyasa koşulları, mevsimsel değişkenler, sosyal ve demografik olayların araştırma kapsamında ele alınarak yük tahminleri yapılmaya çalışıldığı, bu amaçla çoklu lineer regresyon analiz (ÇLR), yapay sinir ağları (YSA), ARIMA, en küçük kareler destek vektör makineleri (EKK-DVM) metotlarına dayalı çeşitli modeller geliştirildiği, çalışma sonuçlarına göre ise YSA, EKKK-DVM'nin kısa vadeli tahminlerde gerçekleşen değerler ile uyumlu sonuç verdiği anlaşılmıştır.(AKMAN,YILMAZ, SÖNMEZ;2018). Buradan hareketle İstanbul ili için gerçekleşen tüketim verileri kullanılarak SimpleRNN (Tekrarlayan Sinir Ağları) ve LSTM (Uzun-Kısa Vadeli Bellek) yapay sinir ağlarına dayalı NARX (Doğrusal Olmayan Otoregresif Dışsal Girdili Model) modelleri ve Prophet/LSTM hibrid modelleri geliştirilmiş ve modellerin kısa dönem (1 saat) performansları test edilmiştir.

Metodlar

Verilerin Toplanması

Elektrik tüketimiyle ilgili olabilecek olası faktörler ve veri kaynakları belirlenmiş ve ham verilerin temininde Apache Airflow, Python gibi açık kaynaklı büyük veri araçları kullanılmıştır. AWS S3 bulut depolama hizmetinde depolanan ham veriler Nifi ETL aracı ile işlenerek Elasticsearch motoruna aktarılmıştır. Çalışmada yer alan veriler 2021 Ocak-2024 Mart aylarını kapsamakta olup saat bazlı sıcaklık ve nem verileri açık kaynaklı Open-Meteo web servisinden, gerçekleşen saat bazlı elektrik tüketim miktarı ise megawatt saat (MWh) cinsinden olup EPİAŞ üzerinden temin edilmiştir.

EPİAŞ üzerinden elde edilip işlenen saatlik tüketim veri seti Türkiye geneline ait olup il bazlı saatlik veri bulunmamaktadır. Kategorik tüketim bilgileri (genel aydınlatma, tarımsal sulama, mesken, sanayi vb.) yüzdesel olarak il bazlı EPİAŞ şeffaflık sayfasında yer almaktadır. Bu nedenle, İstanbul için sentetik saatlik elektrik verileri, Generative Adversarial Networks (GAN'lar) kullanılarak üretilmiştir. Bu çalışmada kullanılan GAN modeli, saatlik elektrik tüketim verilerinin sentetik olarak üretilmesi üzerine kurulmuştur. Modelin geliştirilme süreci iki aşamadan oluşmaktadır: öncelikle, Türkiye'nin saatlik elektrik tüketim verisinin sentetik versiyonu üretilmiş, ardından bu veri İstanbul'un yıllık tüketim oranlarına göre ölçeklendirilerek İstanbul'un saatlik elektrik tüketim verisi elde edilmiştir.

GAN modeli, bir Generator ve bir Discriminator ağından oluşmaktadır. Generator, rastgele bir girdi vektöründen gerçekçi görünen veri noktaları üretmek için tasarlanmıştır. Bu çalışmada kullanılan Generator, artan birim sayısına sahip üç yoğun katman (Dense Layer) ile yapılandırılmıştır. Bu katmanlar, sırasıyla 256, 512 ve 1024 birime sahiptir ve her bir katman LeakyReLU aktivasyon fonksiyonu ve Batch Normalization ile desteklenmiştir. LeakyReLU, özellikle negatif giriş değerlerine karşı duyarlı bir şekilde tepki vererek öğrenme sürecini hızlandırırken, Batch Normalization, katmanlar arasındaki veri dağılımlarının normalize edilmesini sağlar. Çıkış katmanı, tek bir birime sahip olup Linear aktivasyon fonksiyonu ile sonlandırılmıştır.

Discriminator, verinin gerçek mi yoksa Generator tarafından mı üretildiğini ayırt etmek amacıyla tasarlanmıştır. Discriminator, büyükten küçüğe doğru birim sayısına sahip üç yoğun katman ile yapılandırılmıştır: 1024, 512 ve 256 birim. Bu katmanlar LeakyReLU aktivasyon fonksiyonuna sahip olup aşırı öğrenmeyi önlemek amacıyla Dropout (0.3) uygulanmıştır. Discriminator'ın çıkış katmanında ise Sigmoid aktivasyon fonksiyonu kullanılmıştır; bu, verinin gerçek olma olasılığını belirlemeye yönelik bir çıktı sağlar.

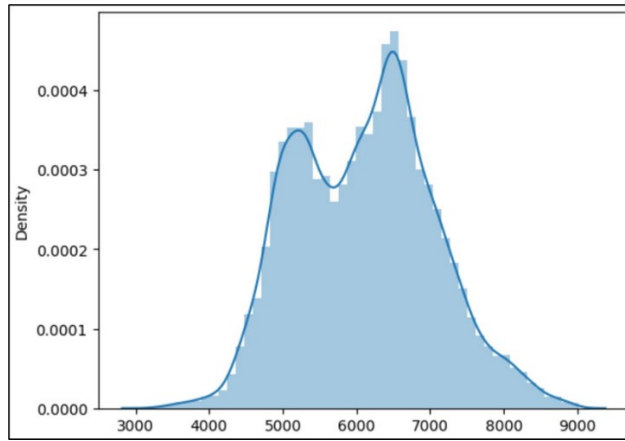
Modelin eğitimi, Generator ve Discriminator ağlarının sırasıyla eğitildiği ve her ikisinin de iteratif olarak birbirini yenmeye çalıştığı Adversarial Training süreci ile gerçekleştirilmiştir. Bu eğitim süreci 10,000 epoch boyunca devam

etmiş ve modelin yeterince uzun süre eğitilmesi amaçlanmıştır. Ancak, modelin sentetik olarak ürettiği verinin gerçekçi olmadığı tespit edilmiştir.

Bu çalışmanın ikinci aşamasında, üretilen sentetik veri, EPDK Elektrik Piyasası Gelişim Raporlarından elde edilen İstanbul'un Türkiye içindeki yıllık tüketim yüzdelerine göre ölçeklendirilmiştir. Ancak, modelin çıktılarını analiz edildiğinde, üretilen sentetik verinin gerçek elektrik tüketim verilerini yeterince iyi taklit edemediği gözlemlenmiştir. Bu sebeple, bu sentetik veriler araştırmanın açıklayıcı istatistiksel analizinde ve nihai sonuçlarında kullanılamamıştır. Bağımsız değişken olan İstanbul Tüketim verisi, Türkiye geneli saatlik tüketim verisinin EPDK Elektrik Piyasası Gelişim Raporlarında yer alan İstanbul iline ait yıl bazlı faturalanan tüketim yüzdesi ile çarpımından elde edilmiştir.

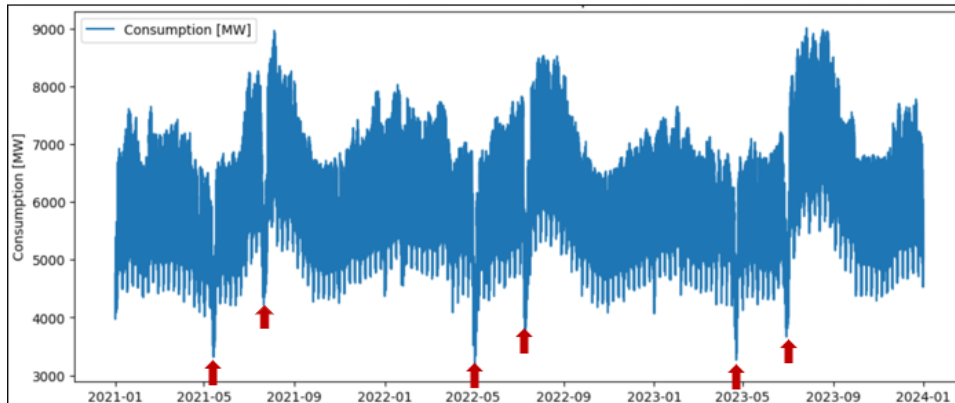
İstatistiksel Analiz

İstanbul'un saatlik elektrik tüketim verilerinin analizi, enerji talebinin zaman içerisindeki değişkenliğini anlamak ve çeşitli faktörlerin bu talep üzerindeki etkilerini belirlemek amacıyla gerçekleştirilmiştir. Bu veriler incelendiğinde, Şekil-1'de görüldüğü üzere, tüketim değerlerinin genellikle 4.000 ila 8.000 MW arasında değiştiği, ancak bu aralığın dışında da bazı değerlerin bulunduğu gözlemlenmiştir. Verinin bimodal bir dağılıma sahip olduğu, tüketimin iki farklı zirve noktasında yoğunlaştığı görülmektedir; bu zirveler 5.000 MW ve 6.000 MW seviyelerinde yer almaktadır. Bu durum, farklı dönemlerdeki tüketim alışkanlıklarına işaret etmektedir. Özellikle mevsimler, hafta sonları ve tatiller gibi çeşitli faktörlerin elektrik tüketimi üzerinde önemli etkileri olduğu anlaşılmaktadır.



Şekil 1 - İstanbul Elektrik Tüketim Verisi

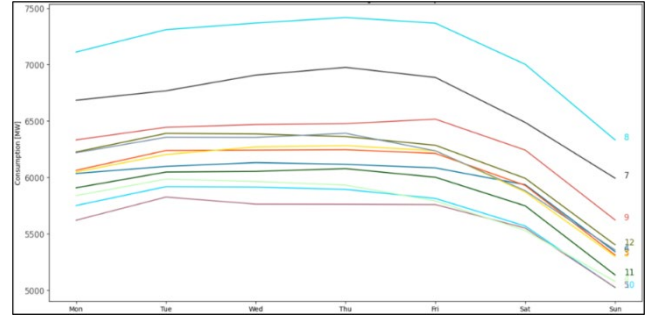
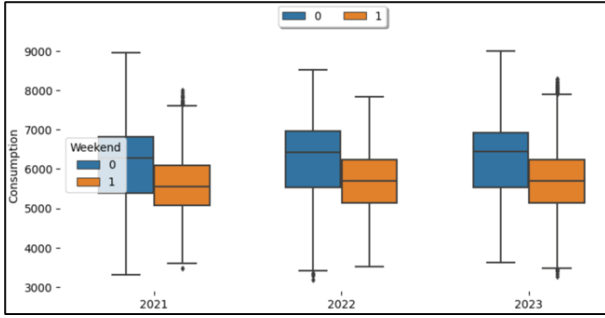
Şekil-2'de gösterildiği gibi, yılın ortalarında ve yılın sonuna doğru elektrik tüketiminde belirgin zirveler gözlemlenmiştir. Bu dönemler, yaz ve kış aylarına denk geldiği için mevsimsel etkilerin belirgin olduğu düşünülebilir. Özellikle kış aylarında ısınma ihtiyacının artması nedeniyle tüketimdeki bu zirveler daha da belirgin hale gelmektedir. Aynı şekilde, yaz aylarında soğutma ihtiyacının artmasıyla tüketim tekrar yükselmektedir.



Şekil 2 - İstanbul Yıllık Elektrik Tüketim Dağılımı

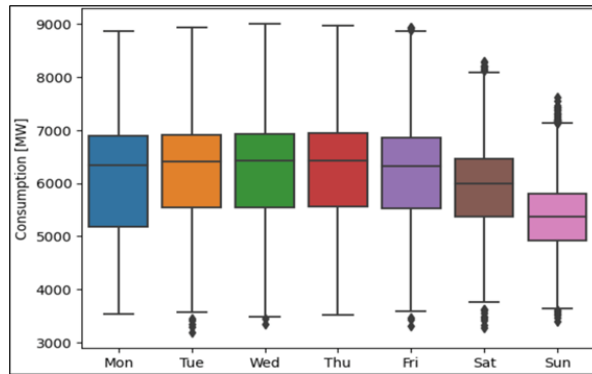
Dini bayramlar gibi özel dönemlerde ise elektrik tüketiminde önemli düşüşler yaşanmaktadır. Bu düşüşler, özellikle toplumsal etkinlikler ve tatil günleri sırasında ekonomik aktivitelerin yavaşlamasından kaynaklanıyor olabilir.

Hafta içi ve hafta sonu elektrik tüketimi arasında belirgin bir ayırım gözlenmektedir. Şekil-4 ve Şekil-5'te gösterildiği üzere; hafta içi günlerinde, sanayi ve ticari faaliyetlerin yoğunluğu nedeniyle elektrik tüketimi daha yüksek seviyelerde seyretmektedir. Hafta sonlarında ise bu faaliyetlerin azalması, genel tüketimin düşmesine neden olmaktadır.



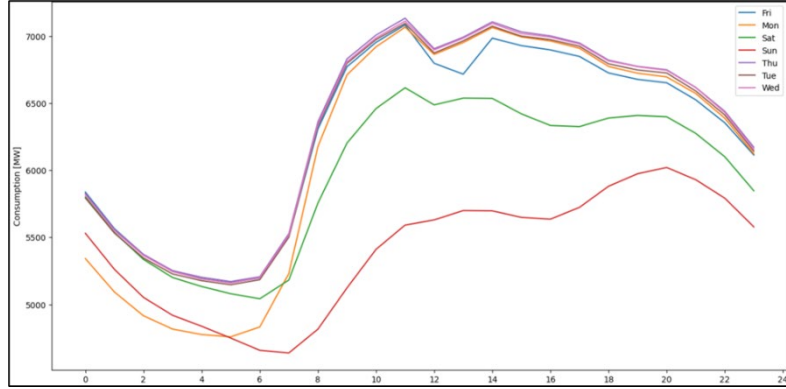
Şekil 4-İstanbul Hafta içi ve Hafta sonu Elektrik Tüketim Dağılımı Şekil 5 - İstanbul Elektrik Tüketim Verisi Günlük Dağılımı

Şekil-6'da görüldüğü üzere hafta içi günlerinde (özellikle salı ve çarşamba günleri) daha düşük tüketim değerlerinin yoğunlaştığı gözlemlenmiştir. Bu durum, evden çalışma gibi faktörlerle ilişkilendirilebilir. Öte yandan, pazar günü ise daha yüksek tüketim değerleri gözlemlenmektedir. Bu, insanların genellikle evde dinlenme eğiliminde oldukları ve bu durumun da tüketim seviyelerini yukarı çektiği anlamına gelebilir.



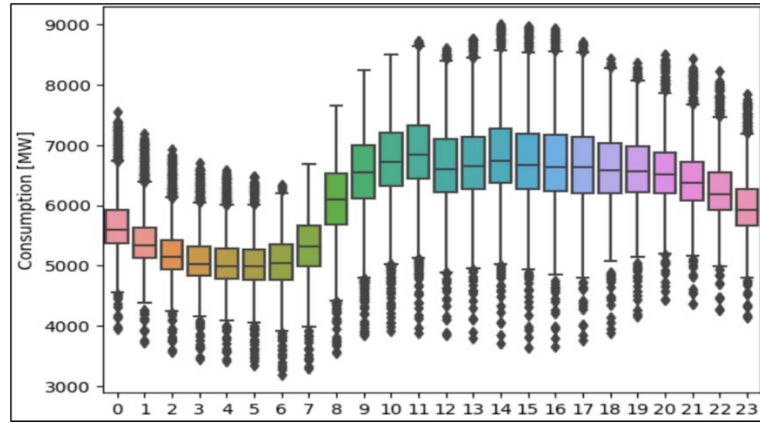
Şekil 6 - İstanbul Günlük Elektrik Tüketim Dağılımı

Şekil-7'de İstanbul'un saatlik elektrik tüketim trendini incelediğimizde ise, sabah saat 5-6 civarında tüketimde ani bir artış gözlemlenmekte ve bu artış saat 11'e kadar devam etmektedir. Bu saatten sonra tüketimde öğleden sonra saat 5-6'ya kadar kademeli bir azalma yaşanmaktadır. Bu düşüş, iş yerlerinin kapanma saatlerine paralel olarak gerçekleşmektedir. Ayrıca, öğle saatlerine yakın bir dönemde de tüketimde bir düşüş görülmektedir. Gece ve akşam saatlerinde tüketimde yukarı yönlü sapmalar gözlemlenirken, gündüz saatlerinde ise aşağı yönlü sapmalar dikkat çekmektedir. Gündüz saatlerinde insanların ofislerde bulunması tüketim seviyelerini düşürebilirken, gece saatlerinde tarımsal sulama ve vardiyalı çalışma gibi aktiviteler tüketimi artırmaktadır.



Şekil 7 - İstanbul Saatlik Elektrik Tüketim Dağılımı

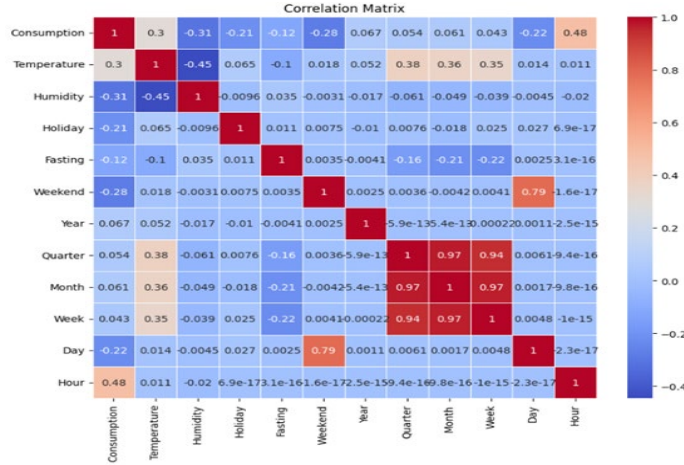
Mevsimsel elektrik tüketim trendleri incelendiğinde (Şekil – 8); yaz aylarında daha fazla sapma (outlier) gözlemlenmiştir, bu da yaz dönemlerinin daha dengesiz olduğunu göstermektedir. Bu dönemde tarımsal sulama, turizm ve soğutma ekipmanlarının kullanımındaki artış gibi faktörler tüketim dalgalanmalarına neden olabilmektedir.



Şekil 8 - İstanbul Aylık Elektrik Tüketim Dağılımı

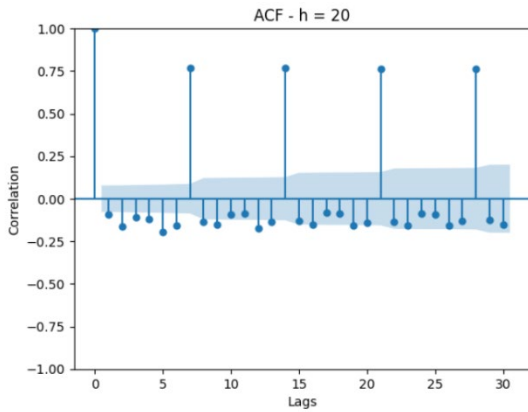
Eylül ayında belirgin bir düşüş yaşandıktan sonra yılın sonuna doğru tüketimde bir toparlanma gözlemlenmektedir. Veride tekrar eden düzenli aralıklarla ortaya çıkan döngüsel örüntüler, mevsimsel etkiler ve çalışma takvimleri gibi faktörlerle açıklanabilir. Mayıs ve temmuz aylarında belirgin tüketim artışları ise, bu dönemlerde meydana gelen beklenmedik olaylar veya anomalilerle ilişkilendirilebilir, bu durum yazın başlangıcı ve bitişi ile ilgili olabilir.

Veri setinin, ardışık ve sabit aralıklı gözlemlerden oluşmasından dolayı, set zaman serisi özellikleri açısından da incelenmiştir. Bu kapsamda, bağımlı ve bağımsız değişken adaylarının durağanlıkları Augmented Dickey Fuller, Phillips-Perron ve KPSS testleriyle ölçülmüş, başarısız olan bağımsız değişkenler elenmiştir. Ayrıca bağımlı değişken ile tüm bağımsız değişkenlerin korelasyon değerleri gözlemlenmiş ve karar mekanizmasındaki eşik değer ve korelasyon matrisi (Şekil- 9) yardımıyla elenmesi gereken bağımsız değişkenlere karar verilmiştir.

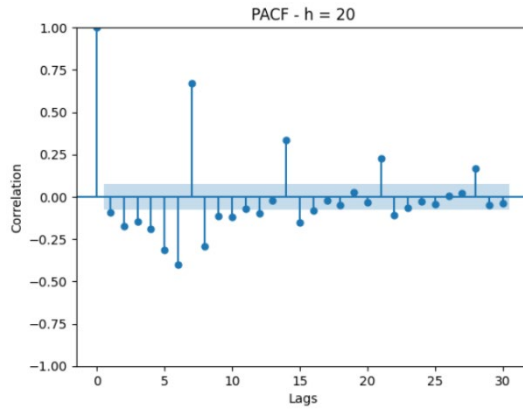


Şekil 9 – Korelasyon Matrisi

Bağımlı değişkenin otokorelasyon (ACF) ve kısmi otokorelasyon fonksiyonları (PACF) incelenmiş, otokorelasyon grafiğinde (Şekil -10) her 5 döngüde bir düzenli sıçramalar görülmüş, PACF grafiği (Şekil -11) incelendiğinde, ilk birkaç gecikme için anlamlı otokorelasyonların varlığı gözlemlenmiştir.



Şekil 10 – Otokorelasyon Grafiği



Şekil 11 – Kısmi Otokorelasyon Grafiği

Bu açıklayıcı istatistiksel analizler, İstanbul'un elektrik tüketim verilerinin zaman içerisindeki dinamiklerini ve çeşitli faktörlerin bu dinamikler üzerindeki etkisini anlamada önemli bir çerçeve sunmaktadır. Bu analizler, ilerleyen bölümlerde yapılan modelleme çalışmalarının temelini oluşturmuştur.

Modelleme

Otokorelasyon grafikleri sonucunda tüketim verisinin mevsimsel özellikler sergilediği ve geçmiş değerlerin mevcut tüketim değerinde anlamlı etkiler barındırdığı görülmüş, dolayısıyla veri setine uygulanacak modellerde AR(Otoregresif) tekniklerin kullanılmasına karar verilmiştir.

Elektrik tüketim verisinde mevsimsel döngülerin görülmesi, hafta içi ve hafta sonu tüketim farklılıklarının ortaya çıkmasından dolayı elektrik tüketim tahminlemede SimpleRNN (Recurrent Neural Networks), LSTM (Long Short-Term Memory) yapay sinir ağlarından yararlanan NARX (Nonlinear AutoRegressive with eXogenous inputs) modelleri ve Prophet zaman serisi modelini içeren hibrid yapılar tercih edilmiştir.

Veri Seti

Veri keşfinden ve zaman serisi analizlerinden yapılan çıkarımlar ile elektrik tüketim verileri ve ilişkili olduğu düşünülen faktörler üzerinden gidilerek nihai veri seti hazırlanmıştır. Bu kapsamda veri seti 1 Ocak 2021 saat

00:00'dan 31 Aralık 2023 saat 23:00'e kadar olan zaman aralığı modelleme, 1 Ocak 2024 saat 00:00'dan 31 Mart 2024 saat 23:00'e kadar olan zaman aralığı validasyon olacak şekilde iki kısma ayrılmış ve her iki veri seti de standart MinMaxScaler kullanılarak normal dağılıma dönüştürülmüştür. Modelleme veri seti normalleştirildikten sonra; kronolojik sıralamaya dikkat edilerek 80%'i modellemede kullanılacak eğitim verisi, 20%'si model performansını ölçmek amacıyla test verisi olarak ayrılmıştır.

Performans Metrikleri

Eğitim verisi kullanılarak geliştirilen modellerin, test verisi üzerinden MAE, RMSE hata metrikleri kullanılarak ve kayıp fonksiyonu, gerçekleşen ve tahminlenen tüketim verilerinin grafiklerin yardımıyla performansları karşılaştırılmıştır.

TARTIŞMA VE SONUÇ

Bilimsel çalışma sonucunda elde edilen bilgiler ışığında sonuçların değerlendirilmesi ve literatür bilgileri ile karşılaştırılması bu bölümde yapılmalıdır. İkinci ve sonraki paragraflarda paragraf başı 0,5 cm girintili olmalıdır. Metin Gövdesi Metin Gövdesi Metin Gövdesi Metin Gövdesi Metin Gövdesi (Yazar Soyadı, 2021; Yazarın Soyadı ve Yazarın Soyadı, 2021). Metin gövdesi metin gövdesi metin gövdesi metin gövdesi metin gövdesi metin gövdesi (Yazarın soyadı, 2021; Yazarın soyadı vd., 2021).

Referanslar

Gündüz, S, Uğurlu, U, Öksüz, İ, 2021. Gün İçi Piyasası Elektrik Fiyat Tahmini için Eksik Verilerin Tamamlanması. Avrupa Bilim ve Teknoloji Dergisi, (25), 334-340.

Kaysal K, Akarslan E, Hocoğlu FO, 2022. Türkiye Kısa Dönem Elektrik Yük Talep Tahmininde Makine Öğrenmesi Yöntemlerinin Karşılaştırılması. BŞEÜ Fen Bilimleri Dergisi 9(2), 693-702.

Biçen Y, 2016. Türkiye Elektrik Enerjisi Piyasası Gelişim Süreci: Gün Öncesi ve Dengeleme Güç Piyasası Özellikleri, Karaelmas Fen ve Müh. Derg. 6(2):432-438

Akman T, Yılmaz C, Sönmez Y, 2018. Elektrik Yükü Tahmin Yöntemlerinin Analizi, Gazi Mühendislik Bilimleri Dergisi 2018, 4(3): 168-175

Karaçor Z, Güvenek B, 2010. Enerji Piyasası Reformlarının Elektrik Enerjisi Piyasasına Etkisi: EÜAŞ ve Ayrıcalıklı Şirketler Üzerine Bir Analiz, Yönetim ve Ekonomi 17/1 (2010) 147-166

Demirtaş I, Tarı EI, 2021. Türkiye'de Elektrik Piyasasının Rekabete Açılması ve Rekabetçi Elektrik Piyasasında Elektrik Tüketiminin Belirleyicileri: ARDL Sınır Testi, Journal of Emerging Economies And Policy 6(1) 148-164

Kızıl Voyvoda Ö, Voyvoda E, 2019. Türkiye'de Enerji Sektörünün Yeniden Yapılandırılması Sürecinde Hukuk Düzenlemeleri, Çalışma ve Toplum, 2019/1 127-154

<https://www.sciencedirect.com/science/article/abs/pii/S1574013720304561> Energy price prediction using data-driven models: A decade review Hongfang Lu ^a, Xin Ma ^b, [Minda Ma](#) ^c, Senlin Zhu [\[MS\(1\)\]](#) ^{d e}

EPDK Elektrik Piyasası 2023 Yılı Piyasa Gelişim Raporu, 2024, Ankara

EPDK Elektrik Piyasası 2022 Yılı Piyasa Gelişim Raporu, 2023, Ankara

EPDK Elektrik Piyasası 2021 Yılı Piyasa Gelişim Raporu, 2022, Ankara

EPDK Elektrik Piyasası 2020 Yılı Piyasa Gelişim Raporu, 2021, Ankara

TEDAŞ, 2023 yılı Türkiye Elektrik Dağıtım Sektör Raporu, 2023, Ankara

TEDAŞ, 2022 yılı Türkiye Elektrik Dağıtım Sektör Raporu, 2022, Ankara.



Integrating RFM Analysis with Power BI for Enhanced Customer Segmentation and CLV Insights

Duygu Demiray Akkaya¹, Ergun Eray Akkaya^{2*}, Muhammet Mehmet Kır³

¹Istanbul Topkapı University, Faculty of Engineering, Department of Computer Engineering, , Istanbul, Turkiye

²Istanbul Topkapı University, Faculty of Economics, Administrative and Social Sciences, Department of Management Information Systems, Istanbul, Turkiye

³Beykoz University, Faculty Engineering and Architecture, Department of Computer Engineering, Istanbul, Turkiye

Abstract

This study presents the implementation of advanced customer segmentation and Customer Lifetime Value (CLV) prediction by integrating Recency, Frequency, Monetary (RFM) analysis with predictive models including Beta Geometric/Negative Binomial Distribution (BG/NBD) and Gamma-Gamma. The transactional data from an e-commerce platform is analyzed and preprocessed for outlier removal. The proposed methodology identifies customer segments based on their purchasing behavior and forecasts future transactions. RFM analysis segments customers, while BG/NBD model predicts future transactions, and Gamma-Gamma estimates the monetary value of those transactions. Combining these models and generating CLV predictions for each customer enables the business to create insights into their future value.

The analysis results are visualized using Power BI. This comprehensive analysis not only offers decision-makers a user-friendly interface but also provides real-time insights into CLV predictions. Furthermore, an interactive and dynamic structure is observed in the analysis results. Finally, this study optimizes marketing and customer retention strategies in the e-commerce domain with actionable insights.

Key words: Customer segmentation, RFM analysis, CLV prediction, BG/NBD model, Power BI.

INTRODUCTION

In today's competitive business environment, achieving long-term success requires understanding customer behavior and effectively managing customer relationships. Businesses may increase customer loyalty, enhance financial performance, and optimize marketing efforts with the use of Customer Relationship Management (CRM) analytics. Through CRM, companies can gather, store, and analyze customer data, which enables data-driven decision-making in marketing campaigns and consumer interactions (Kumar and Reinartz, 2018). CRM analytics has been widely used in numerous industries to segment customers and predict customer value, enabling for the strategic use of marketing resources (Venkatesan and Farris, 2012).

A popular method in CRM is Recency, Frequency, and Monetary (RFM) analysis, a technique that segments customers based on how recently they made a purchase, how often they buy, and how much they spend. In retail and online sectors, where knowing consumer behavior is crucial for customizing promotions, this technique has worked well (Lemon and Verhoef, 2016). However, while RFM analysis provides valuable insights into past customer behavior, it has limitation to predict future actions. Predictive models such as the Beta Geometric/Negative Binomial Distribution (BG/NBD) and Gamma-Gamma models fill this gap by forecast future transactions and spending and this provides a more comprehensive view of Customer Lifetime Value (CLV) (Fader and Hardie, 2019).

The BG/NBD model is very beneficial in non-contractual contexts, where customers may make purchases irregularly. This model offers a probabilistic framework for identifying high-value customers by estimating both the frequency of future transactions as well as the probability of customer dropout (Fader and Hardie, 2019). Assuming that spending behavior is independent of purchase frequency, the Gamma-Gamma model complements BG/NBD by estimating average transaction value and has demonstrated efficacy in retail and e-commerce applications (Kumar and Reinartz, 2018). When combined, these models provide a powerful approach for predicting CLV, enabling businesses to segment customers based on expected future value rather than just historical behavior.

This research aims to utilize these predictive models within a CRM framework, integrating RFM analysis with BG/NBD and Gamma-Gamma models to achieve more dynamic customer segmentation. Transaction data from an online retail store is analyzed to calculate RFM scores, forecast CLV, and group customers into strategic segments. Targeted marketing campaigns are informed by this segmentation, which enhances customer engagement and loyalty. Furthermore, Power BI is used to visualize these insights which optimizes marketing and retention strategies and gives decision-makers real-time access to customer value dynamics.

Although previous research has demonstrated their effectiveness separately, few studies have combined RFM analysis and CLV models into a comprehensive CRM analytics framework. This study contributes to the literature by providing a thorough application of RFM, BG/NBD, and Gamma-Gamma models in a commercial setting and also highlights the potential of CRM analytics to enhance customer satisfaction and improve business performance in the e-commerce industry.

MATERIAL AND METHODS

Material

Predictive statistical models and CRM data analytics approaches are the main sources of information used in this study. Two key prediction models are used in the analysis: the Gamma-Gamma model and the Beta Geometric/Negative Binomial Distribution (BG/NBD) model. Both purchase patterns and dropout probability are captured by the BG/NBD model, which is used to predict the frequency of client transactions. On the other hand, the Gamma-Gamma model is used to predict the average transaction value, assuming consistency in individual spending behavior over time.

These models provide more accurate customer segmentation based on predicted future value and are essential tools for calculating Customer Lifetime Value (CLV). The outcomes from these models are visualized through Power BI, which supports data-driven decision-making by providing a user-friendly interface for viewing and interacting with customer insights.

Methods

The Collection of the Data

The UCI Machine Learning Repository provided the data for this study, which included transactional records from a UK-based online retailer from 2009 to 2011. This data set includes both domestic and international sales, especially souvenirs and promotional products purchased by corporate customers (wholesalers). Key fields within the dataset are InvoiceNo, StockCode, Description, Quantity, InvoiceDate, UnitPrice, CustomerID, and Country.

Several preprocessing procedures were carried out to ensure data integrity:

Removal of Canceled Transactions: Since they reflect canceled transactions, entries with an invoice number beginning with "C" were not included.

Handling of Missing Values: Rows with missing values were removed, as the large sample size permitted minimal data loss without affecting the overall accuracy of the analysis.

Outlier Suppression: To ensure a more consistent and reliable dataset for further research, severe outliers were found and suppressed using the Interquartile Range (IQR) approach.

These preprocessing steps prepared the dataset for Recency, Frequency, and Monetary (RFM) analysis, serving as the basis for customer segmentation.

Statistical Analysis

RFM Analysis Process

Using the preprocessed data, RFM analysis was performed as follows (calculation of RFM metrics):

Recency: Calculated as the difference between the analysis date (set two days after the last transaction) and the most recent purchase date.

Frequency: Determined by counting distinct transactions per customer.

Monetary: Sum of total spending for each customer.

RFM scoring process can be given by the following steps:

Step 1 Determine the Analysis Date: The analysis date is identified as a fixed reference point to calculate the recency of each customer. It is typically set to a specific date shortly after the most recent transaction in the dataset.

Step 2 Groupby Process to Calculate RFM Values: Once the analysis date is set, Recency, Frequency, and Monetary (RFM) values can be calculated for each customer. Functions Employed in Groupby: max(): Used to determine each customer's most recent purchase date in order to compute recency. nunique(): This function counts the number of unique transactions for each client in order to determine the frequency. a sum(): This function adds up all of the purchases to determine the total amount paid (monetary value).

Step 3 RFM Scoring (qcut): Scoring is used following the computation of Recency, Frequency, and Monetary values. Customers are divided into quantile-based score groups using the qcut function; these groups usually range from 1 (lowest) to 5 (highest) for each parameter. Inversely, recent clients receive greater points for recency. Monetary and frequency scores are explicitly assigned; greater scores correspond to higher values.

Step 4 Final RFM Score: To create a total RFM score for each customer, add the individual Recency, Frequency, and Monetary scores (for example, RFM score = 535 for high Recency, low Frequency, and moderate Monetary value). Customers are categorized into meaningful segments (e.g., Champions, At Risk, Hibernating) based on these scores.

Predictive Models in CLV Calculation

In this study, two predictive models—the Beta Geometric/Negative Binomial Distribution (BG/NBD) model and the Gamma-Gamma model—were applied to predict Customer Lifetime Value (CLV). These models enable for a more precise calculation of customer value by predicting both the frequency and monetary value of future transactions. The expected number of orders for each customer is predicted using the BG/NBD model. This model consists of two main processes:

Transaction Process: As long as a client is "alive" (i.e., hasn't churned), it is expected that they will continue to make purchases at a gamma-distributed average transaction rate. This transaction rate distribution throughout the client population is represented by the model's parameters, r and α .

Dropout Process: Each customer has a probability of "dropping out" (i.e., stopping purchases) after each transaction and this probability is modeled using a beta distribution with parameters a and b which represents the likelihood that a customer will no longer be active after a certain transaction.

Combination of the transaction and dropout processes allow the BG/NBD model to predict future transaction counts based on a customer's purchase history. BG/NBD model formula is given as follows (Fader, and Hardie, 2005; Schmittlein et al. , 2007):

$$E(Y(t) | X = x, t_x, T, r, \alpha, a, b) = \frac{a + b + x - 1}{a - 1} \times \frac{\left[1 - \left(\frac{\alpha + T}{\alpha + T + t} \right)^{r+x} {}_2F_1\left(r + x, b + x; a + b + x - 1; \frac{t}{\alpha + T + t}\right) \right]}{1 + \delta_{(x>0)} \frac{a}{b + x - 1} \left(\frac{\alpha + T}{\alpha + t_x} \right)^{r+x}}$$

x :A customer's number of repeat purchases. The recency value is denoted by t_x . The amount of time in weeks that passed between the first and last purchases. Every customer has a distinct recency value.

T : It is the customer age value each week. The customer's age is computed by deducting the date of their first purchase from the analysis date that was established. When all customers are taken into account, the probabilistic state of purchasing behavior will be shaped by the gamma distribution parameters (r, α) . The drop rate likelihood is modeled by the beta distribution parameters (a, b) (Fader and Hardie, 2007).

The Gamma-Gamma model was applied to predict the expected average revenue per transaction. According to this model, the average transaction value varies across customers but remains relatively stable for an individual customer over time. The Gamma-Gamma model provides a probabilistic estimate of this transaction value, with parameters p , q , and γ influencing the distribution of consumer spending.

By combining the BG/NBD model (which predicts the frequency of purchases) with the Gamma-Gamma model (which predicts the average transaction value), the expected Customer Lifetime Value for each customer is calculated. Businesses can use this combination to identify high-value customers based on their projected future transactions and revenue, which enables more informed decisions in resource allocation and targeted marketing. Gamma-Gamma submodel is given as follows (Allenby et al., 1999):

$$E(M|p, q, \gamma, m_x, x) = \frac{(\gamma + m_x x)p}{px + q - 1} = \left(\frac{q - 1}{px + q - 1}\right) \frac{\gamma p}{q - 1} + \left(\frac{px}{px + q - 1}\right) m_x$$

x : Number of repeat sales of a customer

m_x : Total Monetary / transaction numbers

(p, q, γ) : Parameters shaping the distribution of the average transaction value

Calculation of CLV

By multiplying the expected number of transactions (from the BG/NBD model) by the expected average revenue per transaction (from the Gamma-Gamma model), the combined model generates an estimate of CLV. With frequency settings determined by daily or weekly transactions, the customer lifetime value (CLV) for a given time period was calculated using the `customer_lifetime_value()` function, which integrates both models. Customers were subsequently divided into strategic groups based on the computed CLV values in order to optimize resources and conduct targeted marketing.

Data-driven customer segmentation and enhanced CRM strategies are made possible through this predictive approach to CLV, providing firms with better insights into customer value.

Database and Power BI Integration

Microsoft SQL Server was chosen as the primary database for storing RFM scores and CLV results due to its strong compatibility with Power BI, as well as its ability to ensure data security and scalability. SQL Server's secure data storage capabilities are crucial for CRM analytics, where protecting sensitive customer data is vital. By hosting customer segmentation and CLV prediction data in SQL Server, we created a secure, centralized repository that allows efficient real-time data flow to Power BI.

As shown in Figure 1, the database structure includes tables for raw transaction data (`clean_data`), RFM analysis results (`rfm_data`), and CLV predictions (`cltv_data`).

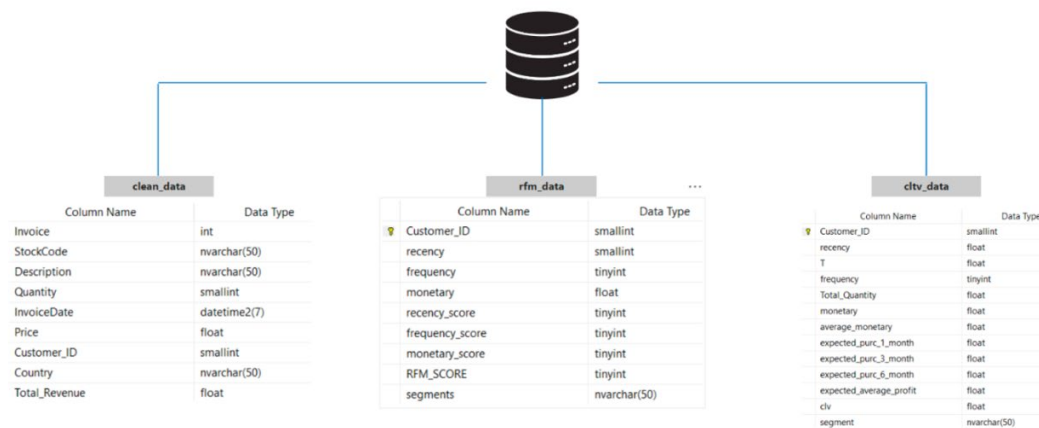


Figure 1. Database tables and columns

Power BI was used to visualize customer segments and provide stakeholders with actionable insights. Key dashboard features include:

Customer Segmentation Maps: Visuals representing customer groups based on RFM scores, highlighting high-value segments.

CLV Projections: Real-time insights into each segment’s potential revenue.

Interactive Panels: Allow users to filter customer data by attributes such as frequency and monetary value, supporting in-depth analysis.

These interactive dashboards are accessible to non-technical users, and Power BI’s secure web publishing options ensure controlled access for internal teams, preserving data confidentiality.

Software Design

There are four integrated layers in the software design:

Database Layer: Preprocessed data, RFM analysis, and CLV predictions are stored in SQL Server and are arranged for effective retrieval.

Power BI Layer: Provides secure access for authorized users by visualizing CRM analytics with real-time SQL Server updates.

Web Product: Using pre-trained models, a Streamlit interface offers real-time CLV predictions, supporting fast decision-making.

User Interaction and Data Export: Power BI supports offline analysis, allowing users to export data in Excel or CSV formats. Interactive features, such as detailed drill-downs and custom filters, provide users with the flexibility to explore customer data based on their unique needs.

Figure 2 illustrates the stages of the software design in the CRM analytics process:

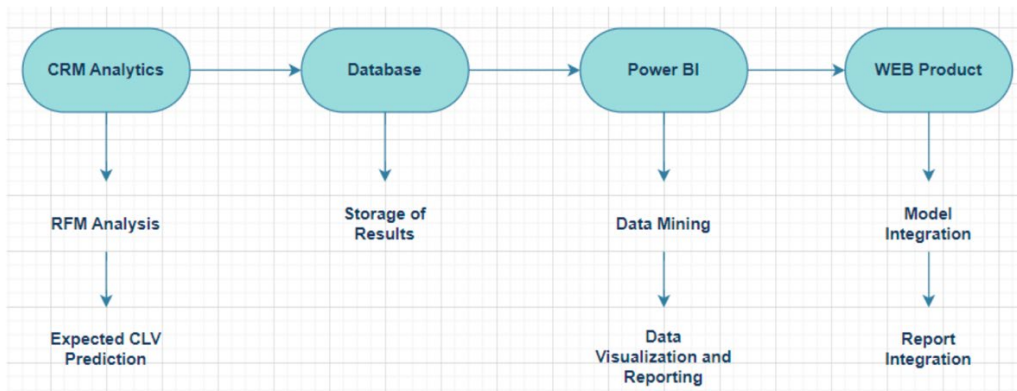


Figure 2. Software Design Document

RESULTS

This section presents the key findings of the study, including customer segmentation results from the RFM analysis, CLV predictions using BG/NBD and Gamma-Gamma models, the impact of outlier suppression on data accuracy, and the interactive visualization of customer segments in Power BI. These outcomes demonstrate how well the techniques used to identify high-value customer groups and optimize CRM strategies.

RFM Analysis Results: Table 1 provides a list of the most valuable customers based on their Customer Lifetime Value (CLV) scores, highlighting segments that represent high revenue potential.

Table 1. Most valuable customers according to CLV

Recency	T (Tenure)	Frequency	Expected Avg. Monetary	CLV
53.29	53.43	89	3533.39	63037.49
51.71	53.29	78	3115.77	48446.52
52.29	53.43	102	1794.85	36335.92
53.29	53.43	205	699.54	28389.97
10.00	10.57	7	5457.79	28068.62
51.57	53.00	94	1367.49	25619.75
0.00	3.43	2	6172.35	19350.42
52.71	53.29	31	2741.73	17751.59

Customers with high CLV ratings typically have a combination of recent, frequent transactions and greater average spending, according to this data. A lower recency score, which indicates recent contacts, may enhance the likelihood of a purchase if a customer remains engaged and maintains a consistent transaction amount. Recency, Tenure, Frequency, and Expected Average Monetary values are among the characteristics that contribute to high CLV and should be examined collectively. These insights are made possible by the use of the BG/NBD and Gamma-Gamma models, which give the company the ability to identify these valuable customers for focused engagement initiatives and resource distribution.

Customer Segmentation in Power BI:

The Power BI dashboard overview, which provides a thorough picture of the customer segmentation results and key CRM metrics, is shown in Figure 3. Users can evaluate CLV predictions, dive down into particular segments based on RFM scores, and interactively explore customer groups using this dashboard. The figure illustrates how the dashboard enables data-driven insights, allowing stakeholders to track engagement trends, identify high-value customer segments, and adjust marketing strategies accordingly.

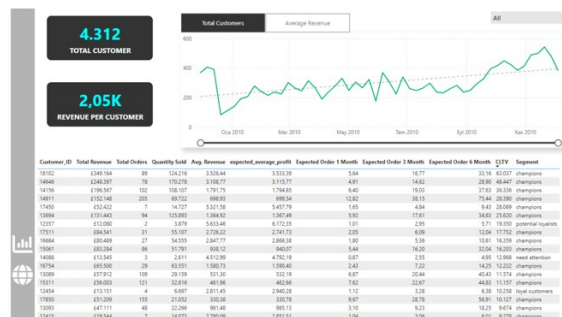


Figure 3. Power BI ‘Customer Detail’ Page

Impact of Outlier Suppression:

Outlier suppression was applied to the dataset in order to guarantee data stability and enhance model accuracy. By eliminating extreme numbers that may affect the RFM and CLV calculations, this preprocessing procedure produced more consistent and reliable customer insights. Due to extreme spending patterns among a tiny fraction of

customers, the data metrics prior to outlier suppression showed significant variability, especially in monetary values (Tables 2). The averages for recency, frequency, and monetary values stabilized following suppression, providing a clearer and more typical representation picture of consumer behavior. By minimizing the impact of outliers and allowing the models to concentrate on real trends within the larger customer base, this modification enables more accurate customer segmentation and CLV prediction.

Table 2. Comparison Before and After Outlier Suppression

Measure	Mean(Before)	Standard Deviation(Before)	Mean(After)	Standard Deviation(After)
Quantity	13.59	96.84	11.91	28.73
Price	3.29	34.76	3.01	3.25

Web-Based Customer Insights Platform:

Streamlit was used to create an interactive online interface that gave users access to real-time customer insights and CLV forecasts. Managers and stakeholders were able to access customer details based on RFM and CLV scores thanks to this interface's capabilities for dynamic filtering and data exploration. Actionable insights are made available by this user-friendly platform, facilitating quicker and more efficient CRM decisions.

DISCUSSION AND CONCLUSION

RFM analysis, BG/NBD, and Gamma-Gamma models were used in this study to segment customers according to their potential for future revenue and forecast Customer Lifetime Value (CLV). The outcomes show how well these models work when combined with interactive Power BI visualizations to create more focused CRM strategy. Businesses may increase customer loyalty and revenue by identifying high-value consumers and allocating resources as efficiently as possible while customizing engagement strategies.

Future improvements might concentrate on using AI-driven segmentation models, such K-Means, to enable real-time, dynamic consumer grouping that adjusts to shifting customer behavior. Maintaining accurate and current CLV predictions would require integrating real-time data sources, such as Apache Kafka. Increasing the scope of the dataset to incorporate information from various e-commerce platforms and sectors would improve the model's applicability to various market segments. Comparing the Gamma-Gamma and BG/NBD models with substitutes like Pareto/NBD may also offer further light on how to best optimize CLV forecasts. Lastly, creating a web platform that is automated for real-time marketing strategies will give companies strong, data-driven CRM decision-making capabilities.

References

Allenby, GM, Leone, RP, Jen, L, 1999. A dynamic model of purchase timing with application to direct marketing. *Journal of the American Statistical Association*, 94(446), 365-374.

Fader, PS, Hardie, BGS, 2005. A note on deriving the Pareto/NBD model and related models: Foundations and trends in marketing. *Marketing Science*, 24(2), 275-282.

Fader, PS, Hardie, BGS, 2007. How to project customer retention. *Journal of Interactive Marketing*, 21(1), 76-90.

Fader, PS, Hardie, BGS, 2019. *Probability Models for Customer-Base Analysis*. Philadelphia: Wharton Digital Press. ISBN: 9781613630810.

Kumar, V and Reinartz, W, 2018. *Customer Relationship Management: Concept, Strategy, and Tools* (3rd ed.). Springer International Publishing.

Lemon, KN, Verhoef, PC, 2016. Understanding customer experience throughout the customer journey. *Journal of Marketing*, 80(6), 69-96.

Schmittlein, DC, Morrison, DG, Colombo, R, 1987. Counting your customers: Who are they and what will they do next? *Management Science*, 33(1), 1-24.

Venkatesan, R, Farris, P, 2012. Measuring and Managing Customer Profitability. Foundations and Trends® in Marketing, 5(1), 1-87.

Avrupa İnsan Hakları Mahkemesi Yargı Kararlarının Metin Madenciliği Yöntemleri İle Tahmin Edilmesi

Nagihan ÜNAL¹, Oktay YILDIZ²

¹Gazi Üniversitesi Bilişim Enstitüsü, Ankara, Türkiye

²Gazi Üniversitesi Mühendislik Fakültesi, Ankara, Türkiye

Özet

Makine Öğrenmesi, Doğal Dil İşleme, Büyük Veri gibi alanlarda yaşanan gelişmeler her alanda olduğu gibi hukuk alanında da inovasyon ve verimlilik artışına katkıda bulunmuştur. Hukuki belgeler karmaşık ve özel bir dil içerdiğinden, bu belgelerin analiz edilmesi, yorumlanması ve sınıflandırılması kritik bir ihtiyaçtır. Doğal dil işleme (NLP) teknolojisi, metinlerin bilgisayarlar tarafından anlaşılmasını ve hukuki süreçlerin hızlanmasını sağlar. Bu çalışmada, Avrupa İnsan Hakları Mahkemesi'nin Avrupa İnsan Hakları Sözleşmesi'nin (AİHS) 6. Maddesi (adil yargılanma hakkı) kapsamında verdiği kararların ihlal olup olmadığını tahmin etmek için çeşitli makine öğrenimi algoritmaları kullanılmıştır. Dengesiz veri setinde nadir olan sınıfa ait veriler, WordNet adlı sözcüksel veritabanı kullanılarak veri artırma yöntemi ile çoğaltılmıştır. Lojistik Regresyon, Destek Vektör Makineleri (SVM), Gradyan Artırma (Gradient Boosting) algoritmalarının performansları değerlendirilmiştir. Sonuçlar, SVM sınıflandırıcısının 10,6 saniye eğitim süresiyle %98 F1-skor oranına ulaştığını göstermektedir.

Anahtar Kelimeler: Avrupa İnsan Hakları Mahkemesi, Makine Öğrenmesi, Doğal Dil İşleme, Metin Sınıflandırma, Metin Verisi Artırma.

GİRİŞ

Büyük veri ve yapay zekâ teknolojileri birçok alanda olduğu gibi hukuk alanında da büyük etkiye neden olmuştur. Hukukun çeşitli alanlarında bu teknolojilerin yaygın kullanımı, hukuk bilimine değerli katkılar sağlamaktadır. Bu teknolojiler hukuki süreçlerin daha etkili ve verimli bir şekilde yürütülmesine olanak tanımaktadır.

Her dil, kendine özgü gramer yapıları ve kurallar bütünü ile karmaşık bir yapıya sahiptir, bu nedenle dil verisi oldukça çeşitli ve karmaşık olabilir. Özellikle hukuki metinler üzerinde çalışırken, bu tür metinlerin ayrıntılı ve katı yapısı, yapay zekâ ile hukuk alanları arasındaki iş birliğini zorunlu hale getirir. Yapay zekâ modellerinin verimli bir şekilde çalışabilmesi için en temel gereksinim, kaliteli ve temiz veri kullanmaktır. Hukuki belgeler, insanların konuştuğu doğal dilde yazılır; bu dilin, bilgisayarların anlayabileceği bir forma dönüştürülmesi gereklidir. Bu aşamada, doğal dilin işlenmesi süreci devreye girer ki bu süreç, bilgisayar bilimlerinin bir alt disiplini olan doğal dil işleme (NLP) ile ilgilidir. Doğru şekilde işlenen metinler, yapay zekâ teknikleri ile analiz edilerek anlaşılabilir ve bu sayede hukuki sorunlar üzerinde değerli bilgiler elde edilebilir.

Doğal dil işleme ve yapay zekanın birleştirilmesi, hukuk profesyonellerinin günlük rutinini büyük ölçüde değiştirebilecek potansiyele sahip, hızla büyüyen bir araştırma alanıdır. Bu birleştirmenin en umut verici uygulamalarından biri hukuki karar tahminidir. Hukuki karar tahmini, geçmiş hukuki kararlar ve mahkeme içtihatları üzerinden yeni bir hukuki durumu değerlendirerek, belirli bir davada mahkemenin muhtemel bir kararını tahmin etmeye odaklanır. Hukuki karar tahmini, hukuki profesyonellerin vakaların sonuçlarına dair otomatik tahminler sağlayarak zaman ve kaynak tasarrufu yapmalarına olanak tanır ve daha karmaşık ve kritik görevlere öncelik verip odaklanmalarını sağlar. Özellikle Avrupa İnsan Hakları Mahkemesi (AİHM) gibi yüksek mahkemelerde, çok sayıda başvurunun olduğu karmaşık hukuki durumları çözmek zor olabilir. Hukuki karar tahmini, bu tür mahkemelerde iş yükünü hafifletebilir ve hakimlere, avukatlara ve diğer hukuki profesyonellere, benzer önceki davaların incelenmesi ve tahminlerin dikkate alınması yoluyla daha bilinçli kararlar almalarına yardımcı olabilir.

Son yıllarda, araştırmacılar hukuki karar temsili için modeller ve yöntemler geliştirmekte büyük ilerlemeler kaydetmiştir. Bu modeller, metin temsili, özellik çıkarma ve makine öğrenme algoritmalarını içeren çeşitli doğal dil işleme tekniklerini kullanarak hukuki davaların kararlarını analiz etmek ve tahmin etmek için tasarlanmıştır. Büyük hukuki veri setlerinin kullanılabilirliği ve makine öğrenmesindeki ilerlemeler, bu modellerin başarısına katkıda bulunmuştur. Literatür incelendiğinde doğal dil işleme kullanılarak hukuki karar tahmini yapılan birçok çalışma bulunmaktadır. Bu çalışmalardan bazıları şu şekildedir:

AİHM'nin yargı kararlarının doğal dil işleme yöntemleriyle tahmin edildiği ilk sistematik çalışmada [1] AİHS'nin bir maddesinin ihlal olup olmadığı şeklinde iki adet sınıfa sahip olan bir problem tanımlanmıştır. SVM modeli

kullanılmıştır. Sözleşme Madde 3, 6 ve 8'e göre 250, 80 ve 254 dava veri setini oluşturmaktadır. Veriler HUDOC [2] veri tabanından çekilmiştir. Metinsel veriler n-gramlar kullanılarak temsil edilmiştir. Mahkeme kararları %79 doğruluk oranıyla tahmin edilmiştir.

Farklı yargı alanlarındaki hukuki kategorilerin tahminin yapıldığı çalışmada [3] metin temsili için iki tür n-gram kullanılmış: char-gram ve n-gram. Veri seti, İspanyol yasal sisteminden alınan 96.163 karardan oluşmaktadır. Veri seti 42 farklı hukuk kategorisine sahiptir. Sınıflandırma için scikit-learn Python kütüphanesinin SVM, DT, RF ve Gradient Boosting sınıflandırıcıları kullanılmıştır. Doğruluk duyarlılık ve F1 metriklerine göre sonuçlar değerlendirilmiş ve tüm bu metriklerde %90'ın üzerinde oranlar ölçülmüştür.

Şulea ve arkadaşlarının yapmış olduğu çalışmada [4] Fransız Yüksek Mahkemesi davalarının sonuçlarını tahmin etmek için Destek Vektör Makineleri kullanılmıştır. 126.000'den fazla belgeden oluşan veri setinde, her vakadan unigramlar ve bigramlar çıkarılmış ve karşılık gelen etiket atanmıştır. Sonuçları tahmin etmek için SVM ensemble sınıflandırıcı eğitilmiş ve vaka sonuçları tahmin edilmiştir.

Medvedeva ve arkadaşlarının yapmış olduğu çalışmada [5] AİHM kararlarının tahmini için makine öğrenmesi algoritması olan SVM kullanılmıştır. AİHS'nin 9 maddesi için "ihlal" ya da "ihlal değil" şeklinde sınıflandırma yapmak için HUDOC veri tabanından davalara ilişkin metinler alınmıştır. N-gram kullanılarak metin temsil edilmiştir. Davaların yaklaşık %75 inde kararlar doğru tahmin edilmiştir.

Bir başka çalışmada [6] makine öğrenmesi yöntemleri kullanılarak mahkeme kararları TF-IDF metin sayısallaştırma yöntemi ve CURE, K-MEANS, AGNES, AFFINITY ve BIRCH gibi yapay zekâ yöntemleri ile kümelenebilir. Performans metriklerine göre en iyi sonucu BIRCH vermiştir.

Gerçekleştirilen bu çalışmada, AİHM dava metinleri veri olarak kullanılmış, ardından doğal dil işleme ve TF-IDF öznitelik çıkarımı süreçlerinden geçirilmiş ve yapay zekâ yöntemlerinin önemli bir alt dalı olan makine öğrenmesi yöntemleri ile analiz edilmiştir. Çalışmanın temel amacı, AİHM kararlarının makine öğrenimi teknikleri ile tahmin edilmesinin pratikte ne kadar başarılı olabileceğini değerlendirmesidir. Çalışmada AİHM'nin AİHS'nin [7] 6. Maddesine (adil yargılanma hakkı) ilişkin kararlarını içeren veri kümesi kullanılmıştır.

MATERYAL VE METODOLOJİ

Veri Kümesi

Çalışmada AİHM tarafından yayınlanan halka açık dava verileri kullanılmıştır. 1959 yılında kurulan AİHM, Avrupa İnsan Hakları Sözleşmesi'nde (AİHS) belirtilen hakların ihlali iddiasıyla yapılan bireysel ve devlet başvurularını inceler ve karara bağlar. Davaların görülmesi için kabul edilebilirlik aşamasını geçmeleri gerekmektedir. Mahkemenin kararları çevrimiçi olarak erişilebilir [2]. Bu çalışmada yayımlanmış karar metinleri kullanılmıştır. Tahmin görevi için esas olması gereken, Mahkemeye taraflarca sunulan argümanları içeren başvuru metinleridir. Ancak bu başvuru metinleri yayımlanmamaktadır. Bu nedenle ana varsayım, yayımlanmış karar metinlerindeki belirli bölümler (Olaylar) ile Mahkeme'ye sunulan başvurular arasında yeterli benzerlik olduğudur. Eğer bu iki metin arasında yeterli benzerlik varsa, bu çalışma başvuru metinleri analiz edilerek sonuçların tahmin edilmesinde kullanılabilir. AİHM kararları genellikle şu bölümlerden oluşur [8]:

- Giriş (Introduction): Dava bilgileri ve mahkeme kadrosu.
- Prosedür (Procedure): Dava süreci ve adımlar.
- Olaylar (The Facts): Davanın arka planı ve olayların ayrıntıları.
- Hukuk (Law): İddiaların hukuki tartışması.
- Karar (Judgement): Nihai karar.
- Karşıt/Ortak Görüşler (Dissenting/Concurring Opinions): Ek görüşler.

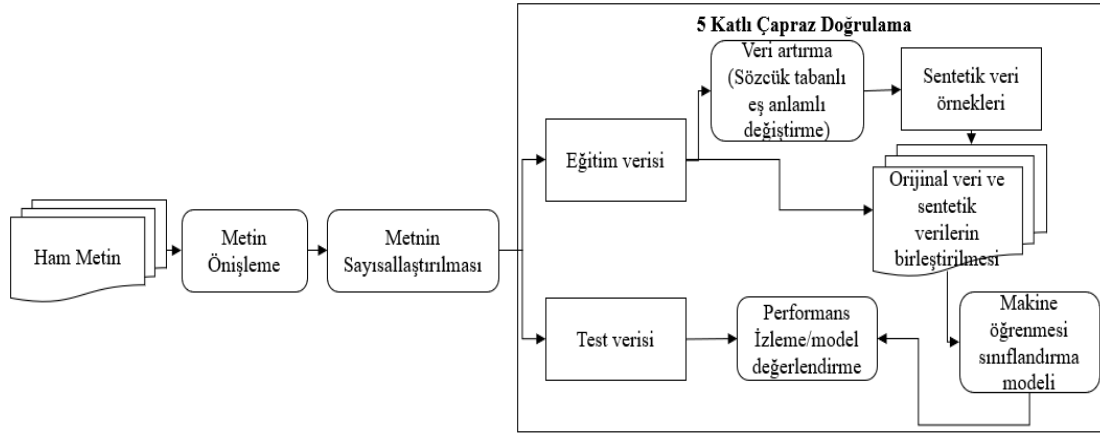
Bu çalışmada karar metinlerinde "Olaylar (The Facts)" bölümü kullanılacaktır. Bu bölüm, davanın tüm ilgili gerçeklerini ve olayların bağlamını kapsadığı için önemlidir. Olaylar bölümü, davaya neden olan olaylar, taraflar arasındaki ilişki ve anlaşmazlık hakkında detaylı bilgi verir. Ayrıca, davanın tarafları ve ilgili diğer kişilerin rolleri de bu bölümde belirtilir. Bu bilgiler, davayı çözmeye yardımcı olmak ve hukuki sonuçlara varmak için kullanılır.

HUDOC (Human Rights Documentation) veri tabanı, AİHM tarafından sağlanan ve AİHS ile ilgili dava kararlarını, yargı içtihatlarını ve diğer ilgili dokümanları içeren kapsamlı bir çevrimiçi kaynaktır. HUDOC, AİHM kararlarını içeren ücretsiz bir veri tabanıdır. Bu çalışmada, HUDOC'tan AİHS'nin 6. maddesi ile ilgili ihlal ve ihlal olmayan

dava verileri çekilmiştir. Çalışma, sadece tek bir kararın bulunduğu davalara odaklanmış ve her iki kararın da bulunduğu davalar veri setinden çıkarılmıştır. Bunun sebebi bu iki kararı (ihlal ve ihlal değil) etkileyen kalıpları daha net belirlemektir. Karar metinlerinde "Olaylar (The Facts)" bölümü düzenli ifadelerle çıkarılmış ve ihlal ve ihlal olmayan davalar sırasıyla 1 ve 0 olarak etiketlenmiştir. Sonuçta, 715 adet 'ihlal değil', 4154 adet 'ihlal' etiketli toplamda 4869 adet davadan oluşan veri seti elde edilmiştir.

Metodoloji

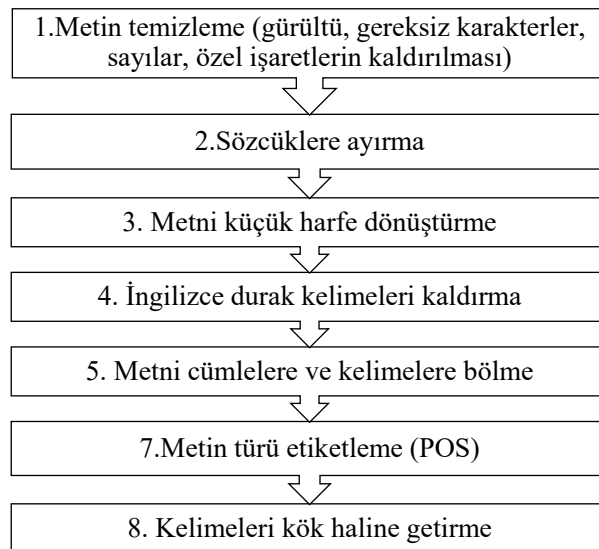
Çalışma, makine öğrenme yöntemlerini kullanarak HUDOC veri tabanından elde edilen dava metinlerinden oluşan veri kümesi ile ilgili dava sonuçlarının ihlal olup olmadığını tahmin etmeyi amaçlamaktadır. Yapılan deneysel çalışma, aşağıdaki adımlardan oluşmaktadır.



Şekil 14. Model süreç diyagramı

Veri önışleme

Metin verisi ön işleme, doğal dil işleme uygulamalarında metin verisinin temizlenmesi, düzenlenmesi ve hazırlanması sürecidir. Bu süreç, verinin daha etkili bir şekilde analiz edilmesini ve makine öğrenimi modellerine beslenmesini sağlar. Geliştirilen sistemde gerçekleştirilen önışleme adımları Şekil 2’de gösterilmiştir.



Şekil 15. Metin önışleme adımları

Tablo 1’de orijinal veri ve önışleme adımından geçirilmiş hali verilmiştir.

Tablo 3. Orijinal ve önişleme adımından geçirilmiş veri örneği

Orijinal Veri	Önişleme yapılmış veri
In reply to a question concerning his diploma, he explained that he had studied at the relevant university between 1994 and 1998 and had sat and completed the fifth-year examinations as an external student during the 1997-1998 academic year.	reply question concern diploma explain study relevant university sit complete fifth year examination external student academic year

Metnin sayısallaştırılması

Sınıflandırma algoritmasına girdi olarak sağlanan metin verilerinin sayısallaştırılmış halleridir. Çalışmada TF-IDF yöntemi uygulanmıştır. Bu teknik, önce bir kelimenin belirli bir belgede ne sıklıkta görüldüğünü ölçen Terim Frekansı'nı (Term Frequency) ve bu terimin korpusta ne kadar önem arz ettiğini ölçen Ters Belge Frekansı'nı (Inverse Document Frequency) hesaplayarak çalışır [9].

Terim Frekansı (TF- Term Frequency): Bir terimin belirli bir belgede ne kadar sık geçtiğini ölçer. Genellikle şu şekilde hesaplanır:

$$TF(t) = \frac{\text{Belgedeki terim } t' \text{ nin sayısı}}{\text{Belgedeki toplam terim sayısı}} \quad (1)$$

Ters Belge Frekansı (IDF- Inverse Document Frequency): Bir terimin tüm belge koleksiyonunda ne kadar yağın olduğunu ölçer. Şu şekilde hesaplanır:

$$IDF(t) = \frac{\text{Toplam belge sayısı}}{\text{Terim } t' \text{ nin bulunduğu belgesayısı}} \quad (2)$$

TF-IDF, belgeler arasındaki kelime önemini belirlemek için kullanılır. Bir kelime, belirli bir belgede sıkça görünüyorsa ve genel korpusta nadirse, TF-IDF ağırlığı yüksek olacaktır. Bu, bir kelimenin belirli bir belge için ne kadar anlamlı olduğunu gösterir. Şu şekilde hesaplanır:

$$TF - IDF(t) = TF(t) \times IDF(t) \quad (3)$$

K-Kat Çapraz Doğrulamada Veri Artırımının Uygulanması

Çalışmada k-kat çapraz doğrulama (k-fold cross-validation) sırasında her deney için eğitim verileri artırılmıştır. Bunun nedeni, modelin performansını ve genelleme yeteneğini artırmaktır. Bu özellikle dengesiz veri setlerinde önemlidir. Çalışmada nadir olan sınıfa (ihlal değil) ait veriler artırılmış, WordNet [10] adlı geniş bir sözcüksel veritabanı kullanılarak metin içindeki kelimeleri eş anlamlılarıyla değiştirildiği yöntem kullanılmıştır. Bu sayede, metnin anlamını bozmadan yeni ve çeşitlendirilmiş veri örnekleri oluşturulmuştur.

Hiperparametre Optimizasyonu

Modelin hiperparametreleri öğrenme süreci başlamadan önce kullanıcı tarafından ayarlanabilen parametrelerdir [11]. Sınıflandırıcıyı daha da geliştirmek için her algoritma için en uygun parametrelerin seçildiği bir hiperparametre optimizasyonu işlemi gerçekleştirilmiştir. Her parametre için farklı değerlerle birçok sayıda iterasyon çalıştırılarak, hiperparametre ayarlanması yapılmıştır.

Sınıflandırma

Sınıflandırma algoritması olarak Destek Vektör Makinesleri, Gradyan Artırma, Lojistik Regresyon algoritmaları

kullanılmıştır.

Performans Değerlendirme

Çalışmada sınıflandırma algoritmalarının performansını değerlendirmek için doğruluğun yanısıra kesinlik, duyarlılık ve F1 skoru metrikleri kullanılmıştır. Bu metrikler hesaplanırken karmaşıklık matrisi dikkate alınmıştır.

Karmaşıklık matrisine göre;

Doğru Pozitif (DP): Gerçekte ihlal olan ve model tarafından doğru bir şekilde ihlal olarak tahmin edilen dava sayısı.

Yanlış Pozitif (YP): Gerçekte ihlal olmayan ve model tarafından ihlal olarak tahmin edilen dava sayısı.

Doğru Negatif (DN): Gerçekte ihlal olmayan ve model tarafından doğru bir şekilde ihlal değil olarak tahmin edilen dava sayısı.

Yanlış Negatif (YN): Gerçekte ihlal olan ve model tarafından ihlal değil olarak tahmin edilen dava sayısını ifade etmektedir.

Yukarıda verilen tanımlamalara göre performans değerlendirme metriklerine ait hesaplama formülleri aşağıda verilmiştir:

$$\text{Doğruluk} = \frac{DP+DN}{DP+YP+DN+YN} \quad (4)$$

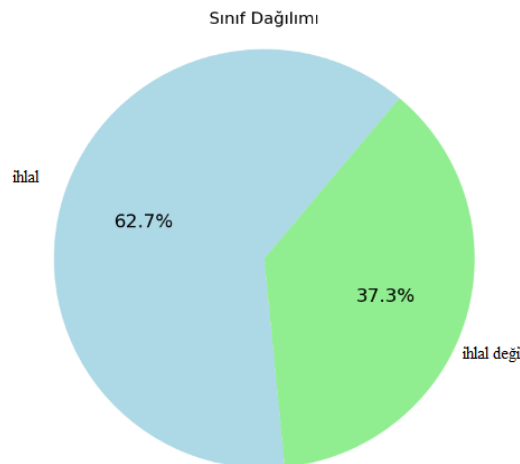
$$\text{Duyarlılık} = \frac{DP}{DP+YN} \quad (5)$$

$$\text{Kesinlik} = \frac{DP}{DP+YP} \quad (6)$$

$$F1 = 2 \times \frac{\text{Kesinlik} \times \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}} \quad (7)$$

DENEYSEL ÇALIŞMA

Çalışmada AIHS 6. Maddesine ilişkin verdiği kararlar kullanılmıştır. Tablo 1'deki veri oldukça dengesizdir. Performansı olumsuz etkileyeceği değerlendirildiğinden tüm veri seti kullanılmamıştır. Dengesiz olan veri setinden 1200 adet 'ihlal', 715 adet 'ihlal değil' etiketli veri kapsama alınmış (Şekil 3) ve nadir olan sınıf için veri artırımı yapılarak sonuçlar karşılaştırılmıştır. 5 katlı çapraz doğrulamada her katta eğitim veri setindeki 'ihlal değil' etiketli veriler 350 adet (%60) artırılmıştır. Modellerin veri artırımı ile performans ve sürelerinin karşılaştırması Tablo 2'de verilmiştir.



Şekil 16. Kapsama alınan verilere ilişkin sınıf dağılımı

Tablo 2’de görüldüğü üzere veri artırımı işlemi sonrası sınıflandırıcıların genelinde performansta artış olduğu görülmektedir. Destek Vektör Makineleri ve Gradyan Artırma algoritmaları en iyi sonuçları vermiştir. Ancak eğitim ve test süresi toplamına bakıldığında Destek Vektör Makineleri sınıflandırıcı Gradyan Artırma sınıflandırıcısına kıyasla yaklaşık %84,62 daha kısa sürede eğitilmekte ve test edilmektedir. Eğitim süresi ve performans artışı parametreleriyle birlikte değerlendirildiğinde Destek Vektör Makineleri sınıflandırıcısının en optimum algoritma olduğu değerlendirilmektedir.

Tablo 4. Modellerin veri artırımı ile performans ve eğitim/test süreleri

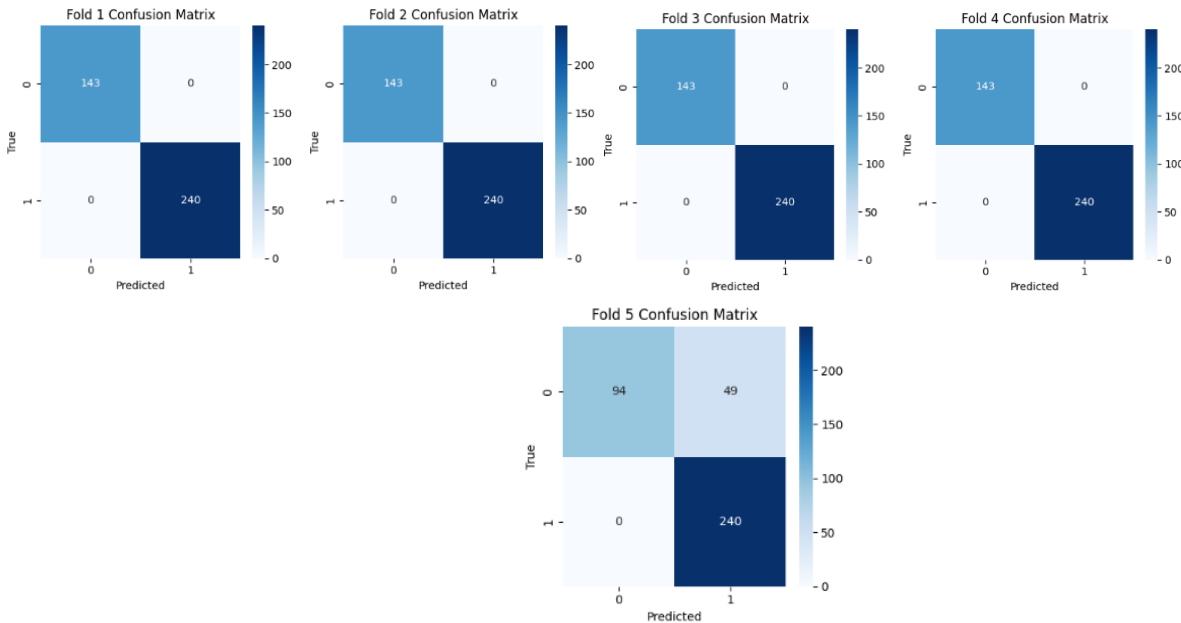
Model	Veri Artırımı	Ortalama Eğitim Süresi (sn.)	Ortalama Test Süresi (sn.)	F1-Skor (%)
Destek Vektör Makineleri	Yok	15.1	1.8	80
Destek Vektör Makineleri	Var	10.6	2.2	98
Grandyan Artırma	Yok	29.2	0.82	80
Grandyan Artırma	Var	78.4	0.32	99
Lojistik Regresyon	Yok	15.04	0.82	81
Lojistik Regresyon	Var	21.15	0.86	97

. Hiperparametre optimizasyonu sonrası bu algoritma ve kullanılan TF-IDF vektörizasyon yöntemi için bulunan en iyi hiperparametreler Tablo 3’te verilmiştir.

Tablo 5. Destek vektör makineleri için en iyi hiperparametreler

Parametre	Değer
kernel	rbf
gamma	0.099
C	100.0
Tfidf max features	None
Tfidf ngram range	(1, 3)

Destek Vektör Makineleri algoritması ile veri artırımı sonrası elde edilen en iyi hiperparametrelerle, 5 katlı çapraz doğrulamada her katta elde edilen karmaşıklık matrisleri Şekil 4’te gösterilmektedir.



Şekil 17. Veri artırma sonrası Destek Vektör Makineleri algoritması karmaşıklık matrisleri (0: ihlal değil, 1: ihlal)

Tablo 4’te Destek Vektör Makineleri için veri artırma sonrası 5 katlı çapraz doğrulamada her katta kesinlik, duyarlılık ve F1-skor değerleri Tablo 4’te verilmiştir.

Tablo 6. Veri artırımı sonrası destek vektör makineleri kesinlik, duyarlılık, F1-skor değerleri

Metrik	Kat 1	Kat 2	Kat 3	Kat 4	Kat 5
Kesinlik (ihlal)	1.0	1.0	1.0	1.0	0.83
Kesinlik (ihlal değil)	1.0	1.0	1.0	1.0	1.0
Duyarlılık (ihlal)	1.0	1.0	1.0	1.0	1.0
Duyarlılık (ihlal değil)	1.0	1.0	1.0	1.0	0.65
F1-Skoru (ihlal)	1.0	1.0	1.0	1.0	0.90
F1-Skoru (ihlal değil)	1.0	1.0	1.0	1.0	0.79

TARTIŞMA VE SONUÇ

Bu çalışmada Avrupa İnsan Hakları Sözleşmesi'nin 6. Maddesine ilişkin başvuruların yapmış olduğu başvuruların ihlal olup olmadığı tahmin edilmiştir. Destek Vektör Makineleri, Lojistik Regresyon ve Gradyan Artırma sınıflandırma algoritmaları kullanılmıştır. Veri seti dengesizliği probleminde çözüm olarak nadir olan sınıfta, WordNet adlı geniş bir sözcüksel veritabanı kullanılarak metin içindeki kelimelerin eş anlamlılarıyla değiştirildiği artırma yöntemi kullanılmış ve tahmin sonuçları karşılaştırılmıştır. Veri artırımı kullanılan algoritmaların (Lojistik Regresyon, Destek Vektör Makineleri, Gradyan Artırma) tümünde performans artışı sağlamış, eğitim süresi dikkate alındığında Destek Vektör Makineleri diğer algoritmalara kıyasla daha iyi performans göstermiştir. Gelecekteki çalışmalarda AİHS'nin diğer maddelerine ilişkin başvuruların da kapsama alınarak veri setinin artırılması, metin temsiline hukuk alanıyla ilgili metinler için özelleştirilmiş kelime vektörlerinin [12] kullanılması hedeflenmektedir.

Referanslar

- [1] N. Aletras, D. Tsarapatsanis, D. Preoşiu-Pietro ve V. Lampos, «Predicting judicial decisions of the European Court of Human Rights: a Natural Language Processing perspective,» *PeerJ Computer Science*, 2016.
- [2] [Çevrimiçi]. Available: <https://hudoc.echr.coe.int>. [Erişildi: 23 Haziran 2023].
- [3] J. González-González, F. d. Arriba-Pérez, S. García-Méndez, A. Busto-Castiñeira ve F. J. González-Castaño, «Automatic explanation of the classification of Spanish legal judgments in jurisdiction-dependent law categories with tree estimators,» *Journal of King Saud University*, cilt 35, no. 7, 2023.
- [4] O. M. Şulea, M. Zampieri, S. Malmasi, M. Vela, L. P. Dinu ve J. V. Genabith, «Exploring the Use of Text Classification in the Legal Domain,» 2017.
- [5] M. Medvedeva, M. Vols ve M. Wieling, «Using machine learning to predict decisions of the European Court of Human Rights,» *Artificial Intelligence and Law*, cilt 28, no. 2, pp. 237-266, 2020.
- [6] M. B. GÖRENTAŞ ve T. UÇKAN, «Makine Öğrenmesi Yöntemleri Kullanılarak Mahkeme Kararlarının Kümeleneşmesi,» *Journal of Computer Science*, cilt 8, no. 2, pp. 148-158, 2023.
- [7] «European Convention on Human Rights,» [Çevrimiçi]. Available: https://prd-echr.coe.int/documents/d/echr/convention_eng. [Erişildi: 8 Haziran 2023].
- [8] [Çevrimiçi]. Available: https://www.echr.coe.int/documents/d/echr/Rules_Court_ENG. [Erişildi: 11 Haziran 2023].
- [9] B. C. ÖĞE ve F. KAYAALP, «Farklı Sınıflandırma Algoritmaları ve Metin Temsil Yöntemlerinin Duygu Analizinde Performans Karşılaştırılması,» *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, cilt 9, no. 6, pp. 406-416, 2021.
- [10] «Princeton University "About WordNet.",» Princeton University, 2010. [Çevrimiçi]. Available: wordnet.princeton.edu. [Erişildi: 23 Ağustos 2023].

- [11] M. Emeç ve M. H. Özcanhan, «Makine Öğrenmesi Algoritmalarında Hiper Parametre Belirleme,» *Mühendislikte Öncü ve Çağdaş Çalışmalar*, İzmir, Duvar Yayınları, 2023.
- [12] I. Chalkidis, D. Karpas ve A. I. & Claims, «Deep learning in law: early adaptation and legal word embeddings trained on large corpora,» *Artificial Intelligence and Law*, cilt 27, no. 2, pp. 171-198, 2019.

Birliktelik Kuralı Analizi ile Ürün Kod Atama Uygulaması

Kadir Korhan BABADAĞ^{1*}, Semih ERGİŞİ²

^{1,2}Yöntem Araştırmaları Daire Başkanlığı / Türkiye İstatistik Kurumu, Ankara, Türkiye

Özet

Bu çalışmada denetimsiz makine öğrenmesi yöntemlerinden birisi olan birliktelik kuralı kullanılarak metin alanlara kod ataması yapılması amaçlanmaktadır. Bu kapsamda verinin temizlenmesi, eğitim verisinin oluşturulması, eğitim verisinden makine öğrenmesiyle kural tablosunun çıkarılması ve kurallar yardımıyla tahminlerin yapıp performans değerlendirilmesi gerçekleştirilmiştir.

Birliktelik kuralı, veri madenciliğinde öğeler arasındaki ilişkileri keşfetmek için kullanılan bir tekniktir ve genellikle Apriori algoritması gibi yöntemlerle gerçekleştirilir. Apriori algoritması, öğelerin sık tekrar eden kümelerini bulmak için destek eşiği kullanarak veriyi tarar ve bu süreç, verinin tüm olası kombinasyonlarını inceleyerek gerçekleştirilir. Büyük veri kümelerinde ilişkileri anlamak ve stratejik kararlar almak için güçlü araçlar olarak kabul edilir.

Çalışmadaki ana yaklaşım, ürün tanımı içinde geçen kelimeler ile ürün kodu arasındaki ilişkinin tespit edilmesidir. Bu süreçte yapılan işlemler 3 ana başlık üzerinde ifade edilebilir.

*İfadeler üzerinde yapılan dilbilimsel işlemler,
Modelleme için eğitim veri setinin hazırlanması,
Modelleme, tahmin ve performans ölçümü.*

Çalışmada karşılaşılan en büyük sorunlardan birisi açık kaynak programlarda Türk dili desteğinin yetersiz oluşudur. Analiz içerisindeki adımların tamamında python programlama dilindeki hazır kütüphanelerin yanı sıra ağırlıklı olarak kendi fonksiyonlarımız kullanılarak çözümler üretilmiştir.

Çalışmada 519.290.605 gözleme denk gelen (1.916.781 tekil ifade) veri seti kullanılmıştır. Bu veri setindeki ürün tanımı kullanılarak her kayıt için COICOP atama yapılması hedeflenmiştir. Birliktelik kuralıyla birlikte ifadeler ve kodlar arasındaki ilişkiler bulunmuş (42.142 kural) ve bunlar dikkate alınarak tüm veri seti skorlanmıştır. Kayıtlara %85 oranında kod ataması yapılmış ve bunların %99 doğrulukla yapıldığı görülmüştür.

Anahtar Kelimeler: Doğal dil işleme, Python, Düzenli ifadeler, Denetimsiz öğrenme, Birliktelik kuralı

GİRİŞ

Çalışmada denetimsiz makine öğrenmesi yöntemlerinden birisi olan birliktelik kuralı kullanılarak metin alanlara COICOP kod ataması yapılması hedeflenmektedir. Bu kapsamda verinin temizliği aşamasında yaşanan en büyük sıkıntı açık kaynak programlardaki Türkçe dil desteğinin yetersizliği olmuştur. Bu eksikliğin giderilmesi için kendi yazdığımız python fonksiyonları kullanılmış ve bu sayede veri karakteristiğine uygun çözümler elde edilmiştir. Veri setinde temel düzeltmeler yapıldıktan sonra veri tutarlılığını öne çıkaran mantıksal kontroller yapılmış ve uyumsuz verilerin eğitim veri setine dahil olmasının önüne geçilmiştir. Eğitim veri setinin elde edilmesinden sonra birliktelik kuralı analizi yapılmıştır. Mevcut birliktelik kuralı kütüphaneleri bizim veri yapımızla uyumlu olmadığı için kendi yazdığımız fonksiyonlardan oluşan python kodları kullanılmıştır. Kuralların elde edilmesiyle birlikte yine python kodları kullanarak tüm veri seti üzerine skorlama yapılmıştır.

Çalışmanın tamamlanmasıyla üç noktası öne çıkmıştır. Bunlardan ilki, Türkçe dilinin açık kaynak programlarda yeterli desteği olmaması nedeniyle kendi geliştirdiğimiz fonksiyonlarla bu soruna çözüm aranmıştır. Diğer taraftan, hazır kütüphanelerdeki birliktelik kuralı ile mevcut veri yapısı uyum ve beklentiler farklı olduğu için bu nokta da kendi birliktelik kuralı fonksiyonumuz yazılmıştır. Son öne çıkan nokta ise, birliktelik kuralı denetimsiz bir öğrenme tekniğiyle bu çalışmada birliktelik kuralının belirlenen bir hedef değişkeni tahmin etmek için kullanılmasıdır.

ARAÇ VE YÖNTEMLER

Veri Seti

Çalışmada ürün tanımı ve bu tanıma ilişkin verilmiş COICOP kodu değerleri kullanılmıştır. COICOP, "Classification of Individual Consumption by Purpose" (Bireysel Tüketim Harcamalarının Amacına Göre Sınıflandırılması) kısaltmasıdır. Bu, bireylerin ve hanelerin tüketim harcamalarını belirli kategorilere ayırmak için kullanılan uluslararası bir standarttır. COICOP, özellikle ekonomik ve sosyal analizlerde, tüketim harcamalarının yapısını incelemek için kullanılır ().

Temel özellikleri

Uluslararası Standart: COICOP, Birleşmiş Milletler tarafından geliştirilmiştir ve birçok ülkede ekonomik veri toplama ve analizlerinde standart bir sınıflama sistemi olarak kabul edilir.

Tüketim Kategorileri: COICOP, tüketim harcamalarını çeşitli kategorilere ayırır. Bu kategoriler genellikle gıda, giyim, konut, sağlık, eğitim, ulaşım gibi geniş grupları içerir.

Düzeyler: COICOP, farklı düzeylerde ayrıntılı sınıflama sunar. Genellikle bu düzeyler ana kategorilerden daha ayrıntılı alt kategorilere kadar genişleyebilir.

Tüketim Analizi: COICOP, hanehalkı bütçeleri ve tüketim harcamaları ile ilgili ekonomik analizlerde, enflasyon hesaplamalarında ve yaşam standartlarını değerlendirmede kullanılır.

Kullanım Alanları

Ekonomik Araştırmalar: COICOP, ekonomik araştırmalarda tüketim alışkanlıklarının analizi ve yaşam standartlarının ölçülmesi için kullanılır.

Enflasyon Hesaplamaları: Tüketici fiyat endeksleri (TÜFE) gibi enflasyon ölçümleri, COICOP kategorilerine dayanarak yapılır. Bu, tüketim maliyetlerindeki değişiklikleri takip etmek için önemlidir.

Sosyal Politika: Sosyal yardımlar ve diğer politika kararları, COICOP kategorilerinin analizi ile daha etkili bir şekilde planlanabilir.

Kategori Örnekleri

Gıda ve İçecekler: Gıda ürünleri, içecekler, restoran ve otel harcamaları.

Konaklama: Kiralar, ev bakım masrafları.

Sağlık: Hastane giderleri, ilaçlar.

Eğitim: Okul harçları, eğitim materyalleri.

Ulaşım: Araç alımı, yakıt, toplu taşıma.

COICOP'un belirli bir ülke veya bölgeye özgü uyarlanmış versiyonları da olabilir. Bu nedenle, her ülkenin COICOP uygulaması, yerel ekonomik koşulları ve veri toplama yöntemlerine göre değişebilir. Kurumumuzda bu sınıflama çok yoğun bir şekilde kullanılmaktadır. İfadeler ve kodların arasındaki uyum yapılan çalışmaların sonucunu ve kalitesini önemli biçimde etkilemektedir.

Çalışmada kullanılan veri seti örneği Tablo 1'de verilmiştir.

Tablo 7: Kullanılan veri seti örneği

PRODUCT_NAME	MADDE_KOD	TOPLAM_SAYI
TORKU TORKU TATKRAK BAHARATLI KRAKER 100 GR torku KG 120062440	01114020200	2
ÜLKER 9 KATTAT RULOKAT FINDIKLI 170GR ülker KG 766086473	01114040200	2,030
SNICKERS SNICKERS 53 ML snickers KG 159343961	01114040200	370
MARKASIZ KESTANELİ 4 6 KİŞİLİK KG 190000008	01114050101	158

Tablo 1’den görüleceği üzere veri setinde ifadelerin Türkçe tanımları, ifadeye karşılık gelen COICOP kodları ve bu kaydın veri setinde kaç kez tekrar ettiği bilgileri mevcuttur. Veri seti 1.916.781 tekil kayıttan oluşmaktadır. Veri setinin toplam büyüklüğü 519.290.605 kayda denk gelmektedir.

Metod

Çalışmada ilk olarak veri setinin temizlenmesi yapılmıştır. Temizleme sürecinde düzenli ifadeler(regular expression) kullanılmıştır (Lopez ve Romero, 2014). Sırasıyla aşağıdaki adımlar uygulanmıştır.

Büyük küçük harf dönüşümü ve noktalama işaretlerinin kaldırılması

“Şamp. 700 ml.” → “şamp 700 ml” , “İĞDIR KAYISI” → “İğdir kayısı”

Bağlaçların kaldırılması

“kız ve erkek çocuk giyim” → “kız erkek çocuk giyim”

Birlikte yazılan harf ve rakamların ayrı olarak düzenlenmesi

“70gr bitter çikolata” → “70 gr bitter çikolata”

Türkçe karakter kaynaklı düzeltme

“Kırmızı Gomlek” -> “kırmızı gömlek”

Türkçe karakter düzeltmelerde diğer kelimelerden faydalanma

“dis cephe” → “dış cephe”, “dis doktor” → “dış doktor”

Kısaltma kaynaklı düzeltme

“çik. Bisk.” → “çikolatalı bisküvi”, “bul mak” → “bulaşık makinesi”

Eklerde yapılan düzeltmeler

“ayakkabi, ayakkabısı, ayakkabıları, ayakabı” → “ayakkabı”

Birimlerden kaynaklı düzeltmeler

“sarımsak 250 gr, sarımsak 1 kg, sarımsak 500 gr” → “sarımsak gr_”

İfade kaynaklı düzeltme

“sarı gömlek, mavi gömlek, yeşil gömlek” → “renk_ gömlek”

“xl pantolon, m pantolon, s pantolon” → “beden_ pantolon”

İfade kaynaklı mantıksal düzeltme

“baklava üretim” → “baklava imalat”, “kayısı üretim” → “kayısı yetiştirme”

İfade içinde tekrarlayan kelimelerin kaldırılması

“abc çay 250 gr. Demlik poşet çay abc” → “abc çay 250 gr demlik poşet”

Şekil 1’de yapılan bu işlemlerden sonra veri setindeki değişim gösterilmiştir.

Alandan Gelen İfade	Dönüşüm Sonrası İfade
0 Kol Detaylı - Elbise Fusya Yırtmaç Detaylı Örme Elbise xl	0 kol detay elbise renk_ yırtmaç örme beden_
00 0013 BN AYK BEJ 40 Terlik	00 0013 <u>bn</u> ayakkabı renk_ beden_ terlik
9930190 - Basic Eşofman Alt Tekstil PANÇO	basic eşortman alt tekstil panço

Şekil 18: Doğal dil işleme adımından sonra veri seti

Doğal dil işleme adımları sonrasında veri de mantıksal tutarlılık kontrolü yapılmıştır (Jurafsky ve Martin, 2021). Mantıksal tutarlılıktan kastedilen eğitim veri setinde aynı ifadenin aynı kodu almasıdır. Bir ifadenin farklı kod alması veri tutarlılığını bozacaktır. Bu kapsamda aynı ifadenin farklı kod aldığı kayıtlar eğitim veri setine dahil edilmemiştir. Şekil 2’de veri tutarsızlık kapsamında veri setindeki örnek bir gözlem sunulmuştur.

Yeni	Madde_Kod	Toplam_Sayı
indomie doodle 75 gr köri çeşni p indomie kg 908020014	1115010110	788
indomie doodle 75 gr köri çeşni p indomie kg 908020014	1115010100	2862

Şekil 19: Veri tutarsızlığı kontrol örneği

Yapılan mantıksal tutarlılık kontrolü sonrasında 1.9716.781 olan tekil kayıt sayısı 1.848.890 tekil kayda düşmüştür. Bu aşamada karşılaşılan en büyük sorun verilerin tekrar etme sayısındaki büyük farklılıktır. Buna dair bir örnek Şekil 3’te verilmiştir.

Yeni_Sıralı	Toplam
526094045 76 eti g kg kraker pizza	138462
500 526069500 eti ezme g kg lifalif yulaf	135180
106968492 750 adet dokuma erkek kiper kot pantolon smart trendy ürün v vision	1
069286507 adet çocuk gömlek hırka k kız mont örme snowball ürün	1

Şekil 20: Eğitim veri setinde ifade sıklığı

Bu durum aynı zamanda birliktelik kuralları üzerinde etkilidir. İfadeler mevcut tekrar sayıları ile birliktelik kuralına dahil edildiğinde çok tekrar eden kayıtların etkinliği artacak ve az tekrar eden ifadelerin gücü azalacaktır. Bu durum aynı zamanda nadir görülen COICOP değerleri için de etkili olacaktır. Diğer yandan her ifadenin sadece bir kere modele dahil edilmesi hiyerarşilerin kaybolmasına sebep olacak ve genel model performansını olumsuz etkileyecektir. Bu iki durum dikkate alındığında bulunan çözüm ifadelerin tekrar sayılarının yeniden ağırlıklandırılmasıdır. Bu sayede hem hiyerarşik yapı korunacak hem de çok tekrar edenlerin az tekrar edenler üzerindeki etkisi azaltılacaktır.

Birliktelik Kuralı Analizi

Birliktelik kuralları, veri madenciliğinde sıklıkla kullanılan yöntemlerden biridir. Günümüzde “Tavsiye Motoru” olarak aldandırılan yöntemin arkasında birliktelik kuralı yöntemi vardır. Bir ürün satın alırken, “bu ürünü alan bunu da aldı” şeklindeki uyarının arkasında birliktelik kural analizinden elde edilen sonuç yatmaktadır (Babadağ, 2003).

Birliktelik kuralları analizi problemi ilk olarak Agrawal, Imielinski ve Swami tarafından 1993 yılında ele alınmış olup ve veri madenciliğinde kullanılan ilk tekniklerden birisidir (Agrawal vd., 1993). Birliktelik Kuralları Analizi için Apriori, Carma, Sequence, GRI, Eclat, FP-Growth vb algoritmalar olsa da bunlar içerisinde en yoğun kullanılan algoritma Apriori’dir.

Apriori Algoritması

Apriori Algoritması önceki adımdan bilgi aldığı için “prior” kelimesinden türetilmiştir (Agrawal ve Srikant, 1994). Bu algoritma tekrarlama mantığı üzerine geliştirilmiştir (Han ve Kamber, 2006). Birliktelik kuralı analizi, tüm sık geçen öğelerin bulunması ve sık geçen bu öğelerden güçlü birliktelik kurallarının üretilmesi olmak üzere iki aşamalıdır.

Destek (Support), belirli bir kuralın verilerde ne sıklıkla ortaya çıktığını gösterir. Matematiksel olarak, destek bir kuralın gerçekleşme olasılığıdır. Örneğin, "Dondurma ve çikolata satın alma" kuralının destek değeri, bu iki ürünün birlikte satın alındığı işlemlerin toplam işlem sayısına oranıdır. N

$$\text{Destek} = \frac{\text{Frekans}(X,Y)}{N}, \quad N: \text{Toplam işlem sayısı}$$

Güven (Confidence), bir olayın diğer bir olayla birlikte gerçekleşme olasılığını gösterir. Matematiksel olarak, A olayı gerçekleştiğinde B olayının gerçekleşme olasılığıdır. İlgili (Lift), iki olay arasındaki ilişkiyi ölçer ve olayların bağımsız olup olmadığını belirler. Eğer ilgili 1'den büyükse, olaylar arasında bir ilişki var demektir. 1 ise ilişki yoktur ve 1'den küçükse ilişki negatif olarak adlandırılır.

$$\text{Güven} = \frac{\text{Frekans}(X,Y)}{\text{Frekans}(X)}$$

$$\text{İlgili} = \frac{\text{Frekans}(X,Y)}{(\text{Frekans}(X) * \text{Frekans}(Y))}$$

Birliktelik Kuralı ile Kod Atama

Python'daki mevcut kütüphaneler transactional veri setini desteklememektedir. Öte yandan bizim veri setimizin büyüklüğü ve kelime sayısındaki çokluk nedeniyle transactional veri setiyle çalışmak bir zorunluluk haline almıştır. Bundan dolayı birliktelik kuralı çalışma prensibine sağdik kalarak transactional veriye uygun kendi fonksiyonlarımızı yazdık. Kuralların hepsi bu fonksiyonların çalışmasıyla elde edilmiştir. Diğer taraftan sürecin hızı ve kalitesi açısından tekrar eden kuralların elimine edilmesi gerekmektedir. Örnek:

pişirme → 05612031100

Yukarıda verilen kural dikkate alınır, aşağıdaki kurallara ihtiyaç yoktur.

pişirme, kağıt → 05612031100

pişirme, yağlı → 05612031100

Çünkü “pişirme” kelimesi “05612031100” kodunu alması için destek ve güven değerlerine tek başına sağlamaktadır. Söz konusu 2 diğer kuralın elimine edilmesi skorlama sürecinin hızlanmasına katkı sağlayacaktır.

Yukarıda verilen örneğe benzer şekilde kurallar elimine edildiğinde birbiri ile kesişmeyen 42.142 kural elde edilmiştir.

Skorlama

Eğitim verisinden kural setleri elde edildikten sonra kurallar kullanılarak mevcut veri setindeki ifadelerin skorlanması yapılır. Bir ifade birden fazla kural ile çakışabilir. Burada dikkate alınması gereken durum aynı ifadeye karşı farklı tahmin değerlerinin olup olmamasıdır. Skorlama sonucunda 3 durumun ortaya çıkma ihtimali vardır. Bunlar sırayla aşağıda verilmiştir.

Veri setindeki ifadeye karşılık tek bir kodun tahmin değerinin elde edilmesi,

Veri setindeki ifadeye karşılık birden fazla tahmin değerinin elde edilmesi,

Veri setindeki ifadeye karşılık hiçbir kodun elde edilememesi.

Tablo 8: Kod atama 1. madde örnek

Madde_Kod	Yeni	Reg_Tahmin	Açıklama	Kod
1111010101	carrefour baldo pirinç 1 kg 405010305	1111010101	baldo kg	1.1
1111010101	besintaş baldo pirinç 2 5 kg 382202506	1111010101	baldo 5, baldo kg	1.1
1111010101	onur bereket bak pirinç baldo 2 5 kg 417890250	1111010101	baldo 5, baldo kg	1.1

Tablo 2'nin 2.satırını incelenirse, aynı ifade hem “baldo, 5” hem de “baldo, kg” ifadelerinden dolayı “1111010101” kodunu almıştır. Kod aynı olmasına rağmen ifade farklı kurallara takılmıştır.

Tablo 9: Kod atama 2.madde örnek

Madde_Kod	Yeni	Reg_Tahmin	Açıklama	Kod
3121060600	ürün erkek klasik örme hırka gömlek mont d y yaş 22 w adet 113574549	03121060600 03122171400	, d gömlek y , hırka klasik w	1.2
12130020102	signal beyaz now 75 ml performans dış fırça 637592089	12130020102 12130020502	, signal 75 , dış fırça	1.2
12130102604	ipek duş jel ylang 400 ml 1 551011622	12130102604 12130130101	, duş jel , ipek 1	1.2

Tablo 3 incelendiğinde yine bir ifadenin birden fazla kurala takıldığı görülür ancak kural sonunda elde edilen tahmin değerleri birbirinden farklıdır. Bu durumda destek ve güven değerlerinin çarpılması ile elde edilen skor değerinin yüksekliğine göre tahmin önceliği belirlenir.

Performans

Tüm veri seti elde edilen 42.142 kural setinden faydalanılarak skorlanmış ve aşağıdaki sonuçlar elde edilmiştir. Tablo 4'deki sonuçlarda “Ağırlıksız Oran” kolonu tekil ifadeleri temsil etmektedir, “Ağırlıklı Oran” kolonu ise tekrarlı kayıt sonuçlarını göstermektedir.

Tablo 10: Skorlama performans sonuçları

KOD ATAMA	Sayı	Sonuc	Ağırlıksız Oran	Ağırlıklı Oran
-	305,183	77,356,292	15.9%	14.9%
1.1	1,519,609	434,639,425	79.3%	83.7%
1.2	91,989	7,294,888	4.8%	1.4%
Toplam	1,916,781	519,290,605	100.0%	100.0%

Tablo 4 incelendiğinde kayıtların (“Ağırlıklı Oran”) %85’inden fazlasına bir kod atandığı görülür. Atanamayan kodların içinde ilk aşamada mantıksal tutarsızlık nedeniyle ayrılan kayıtların da olduğu dikkate alındığında bu oranın aslında daha yüksek olduğunu söylemek mümkündür. Diğer taraftan, atama yapılan kayıtların %98.3’ünün tek bir kod aldığı görülür. Bu durum kural setinin tutarlılığı olarak yorumlanabilir.

Tek kod alan kayıtlar alandan gelen kod değerleriyle karşılaştırıldığında Tablo 5 elde edilmiştir.

Tablo 11: Tek kod alan kayıtlar inceleme

	Kayıt Sayı	Doğru Eşleşme	Oran
Ham Veri	1,519,609	1,492,839	98.2%
Ağırlıklı Veri	434,639,425	430,191,097	99.0%

Tablo 5 incelendiğinde tek tahmin değeri alan kayıtların alandan gelen kod değerleriyle uyum oranı %99 olarak gerçekleşmiştir. Buradan da, elde edilen kurallardaki kelimelerin gerçekten anahtar ifadeler olduğu ve alanı iyi temsil ettiği söylenebilir.

TARTIŞMA VE SONUÇ

Çalışmada yapılanlar aşağıdaki şekilde sıralanabilir.

Sonuç

Türkçe dil problemlerine yönelik alternatif öneriler: Mevcut kütüphanelerin Türkçe dilindeki yetersizliğinden dolayı, kendi geliştirdiğimiz düzenli ifadeler (regular expression) kullanılarak veri üzerinden doğal işleme yapılmıştır (Friedl, 2006).

Çalışma süresi ve hızdaki performans: Daha hızlı işlem yapabilmek için veri yapısı uzunlamasına (transactional) oluşturulmuştur. Birliktelik kuralı ile ilgili olan mevcut kütüphaneler bu yapıyı desteklemediği için kendi yazdığımız fonksiyonlar kullanılmıştır.

Ağırlıklandırma ile çok tekrar eden kayıtların etkisinin azaltılması: Çok tekrar eden ifadelerin kurallar üzerinde aşırı etkili olmasını engellemek için yeniden ağırlıklandırma yapılmıştır. Bu sayede, veri setindeki mevcut hiyerarşilerin korunması da sağlanmıştır.

Hedef kategori sayısının çok olduğu durumlarda mevcut yöntemlerin yetersiz kalması: Klasik sınıflama yöntemleri, hedef değişkenin çok kategorili olduğu durumlarda, nadir olan kategoriler için düşük performans göstermektedir. Birliktelik kuralının bu şekilde kullanımı çok sayıda kategorinin tahmin edilmesinde önemli katkı sağlamaktadır.

Diğer taraftan birliktelik kuralı analizi, denetimsiz algoritma olmasına rağmen bu çalışmada denetimli olarak kullanılmıştır.

Öneriler

Veri setindeki kalitenin artırılması: Çalışmada en büyük zaman verinin temizlenmesi aşamasında harcanmaktadır. Tutarlı veri setiyle çalışmak tüm sürecin kalitesini etkilemektedir. Dolayısıyla analiz için gelen veri setinin daha tutarlı olması tüm süreci olumlu etkileyecektir.

Kayıt sayısı az tekrar eden kategoriler için ters ağırlıklandırma ile özel kuralların bulunması: Nadir gözlenen durumların doğru tahmin edilmesi gerektiği durumlar için ekstra ağırlık verilmesi gerekebilir bu senaryoya ilişkin ters ağırlıklandırma yöntemleri denenebilir.

Gridsearch ile birliktelik kuralındaki parametreler için daha zengin alternatiflerin sunulması: Birliktelik kuralındaki parametre seçimi (destek ve güven değeri) şartlara ve duruma göre değişiklik gösterebilir. (Genel performansın artırılması veya nadir görülen değerlerin doğru tahmin edilmesi gibi). Söz konusu durumlar için ayrı ayrı kural setleri oluşturulup bunların birlikte değerlendirilmesi durumunda gridsearch önem kazanacaktır.

Tahmin edilemeyen kayıtlar için cümle benzerliği algoritmalarının kullanılması: Tüm yukardaki denemeler sonrasında halen tahmin edilemeyen kayıtlar için cümle benzerliği algoritmaları sürece dahil edilerek performansın değişimi incelenebilir.

Farklı yöntemlerden elde edilen sonuçların birlikte incelenmesi: Yukarıda bahsedilen yöntemlere WordVec gibi algoritmaların sonuçları da eklenerek toplu model değerlendirmesi (heyet algoritmalar mantığı) yapıp sonuçlar incelenebilir.

Referanslar

Agrawal, R., Imielinski, T., & Swami, A. (1993). Mining association rules between sets of items in large databases. *ACM SIGMOD Record*, 22(2), 207-216. <https://doi.org/10.1145/170039.170072>

Babadağ, K. K. (2003). *Veri madenciliği yaklaşımı ve veri kalitesinin artması için kullanılması*. Uzmanlık tezi, Devlet İstatistik Enstitüsü

Friedl, J. E. F. (2006). *Mastering regular expressions* (3rd ed.). O'Reilly Media.

Jurafsky, D., & Martin, J. H. (2021). *Speech and language processing* (3rd ed.). Pearson.

Lopez, F. & Romero, V. (2014). *Mastering Python Regular expressions*. Packt Publishing.

MEDOS: Kısa Dönemli İstatistiklerin Mevsimsellikten Arındırılması için Bir R-Shiny Uygulaması

Muhammed Fatih TÜZEN^{1*}

¹Yöntem Araştırmaları Daire Başkanlığı / Türkiye İstatistik Kurumu, Ankara, Türkiye

Özet

Mevsimsel düzeltme, ekonomik analizlerde zaman serisi verilerindeki mevsimsel dalgalanmaları gidermek için kullanılan kritik bir yöntemdir. Türkiye İstatistik Kurumu (TÜİK), bu amaçla kısa dönemli istatistikler için mevsim ve takvim etkilerinden arındırılmış veriler üretir. Bu süreçte, veriler ARIMA modelleriyle analiz edilerek mevsimsel etkilerden arındırılır ve toplamsal ya da çarpımsal modeller kullanılarak ayrıştırılır. TÜİK, bu işlemleri daha verimli ve standart hale getirmek için MEDOS (Mevsimsel Düzeltme Otomasyon Sistemi) adlı bir R Shiny uygulaması geliştirmiştir. MEDOS, mevsimsel düzeltme işlemlerini merkezi bir sistemle otomatikleştirerek veri işleme sürecini hızlandırır, insan hatalarını en aza indirir ve sonuçların tutarlılığını artırır. Uygulama, JWSACruncher aracı ile JDemetra+ kullanarak önceden belirlenmiş mevsimsel düzeltme modellerini uygular ve veri analizi, rapor oluşturma gibi işlemleri tek merkezden yürütür. Ayrıca, MEDOS'un esnek yapısı sayesinde yeni istatistikler kolayca sisteme entegre edilebilir. Bu sayede TÜİK, zaman serisi analizinde verimlilik ve doğrulukta yeni bir standart oluşturmuştur.

Anahtar Kelimeler: Mevsimsel Düzeltme, Otomasyon, R, Shiny, JDemetra+

GİRİŞ

Mevsimsel düzeltme, zaman serisi verilerindeki mevsimsel değişkenlikleri ortadan kaldırarak, verilerin ardındaki gerçek eğilimlerin daha net bir şekilde gözlemlenmesini sağlar. Bu işlem, ekonomik ve istatistiksel analizlerde büyük önem taşır. Türkiye İstatistik Kurumu (TÜİK), pek çok kısa dönemli istatistikte mevsim ve takvim etkilerinden arındırılmış veriler sunmaktadır. Bu verilerin oluşturulması sürecinde, her yıl düzenli olarak modeller JDemetra+⁴ yazılımı ile belirlenir. Belirlenen modeller yıl boyunca kullanılarak ilgili birimler tarafından mevsim ve takvim etkilerinden arındırılmış nihai sonuç tabloları üretilir. Bu sonuçların oluşturulmasındaki süreçte arka planda birçok hesaplama adımı yatmaktadır ve bu adımlar üretilen istatistiğin konu başlığına göre değişmektedir. Ayrıca sonuçların her bir zaman serisi değişkeni düzeyinde anlık olarak tablo ya da grafiksel olarak görüntülenmesi, önceki dönemlere göre olan büyümelerinin hem rakamsal hem de görsel olarak analiz edilmesi, özet istatistikler elde edilmesi ve sonuçların çeşitli şekillerde dışarı aktarılması da bu sürecin parçalarını oluşturmaktadır. Bu doğrultuda üretici birimlerin de talepleri dikkate alınarak bu karmaşık ve zaman alıcı süreci daha verimli hale getirebilmek için MEDOS (Mevsimsel Düzeltme Otomasyon Sistemi) adlı R Shiny tabanlı bir uygulama geliştirilmiştir.

MEDOS, mevsimsel düzeltme sürecini merkezileştirerek, sürecin daha hızlı ve standart bir şekilde işlenmesine olanak tanır. JDemetra+ yazılımı kullanılarak önceden belirlenen modellerle veriler otomatik olarak düzeltilir. Uygulama, sonuçların görsel analizinden, NACE kodları ve diğer sınıflandırmalara göre raporlanmasına kadar geniş bir araç seti sunarak, analistlerin veri işlemeye ayırdığı zamanı azaltır ve veri yorumlamaya odaklanmalarını sağlar. Sistemin esnek yapısı, yeni istatistiklerin sisteme entegre edilmesine olanak tanıyarak TÜİK'in değişen veri ihtiyaçlarına hızlıca yanıt verebilmesine imkan sağlar. Bu çalışmanın amacı, MEDOS'un işleyişini ayrıntılı olarak inceleyerek, mevsimsel düzeltme analizlerinde verimlilik ve doğruluğun artırılmasını sağlamaktır.

YÖNTEM

Mevsimsel Düzeltme

Mevsimsel düzeltme ekonomide politika yapma, konjonktür analizi/modelleme ve öngörü çalışmaları için anahtar bir araçtır. Bu amaçla uzun yıllardan bu yana İstatistik Ofisleri, resmi istatistikleri mevsim ve takvim etkisinden arındırılmış olarak yayımlamaktadır. Verilerde yıldan yıla yinelenen ve yılın bazı ay/dönemlerinde ortaya çıkan periyodik artış ya da azalışlar mevsimsel etkiler olarak adlandırılırken, ay/dönem/yıl içindeki takvim

⁴ JDemetra+, zaman serilerinin mevsimsel düzeltilmesi için geliştirilmiş, kullanıcı dostu bir yazılım ürünüdür. Herkese açık, açık kaynaklıdır ve 2015 yılından bu yana Avrupa'da resmi istatistiklerin mevsimsel düzeltilmesi için önerilen araç olmuştur.

kompozisyonuna bağlı olarak oluşan etkiler takvim etkileri olarak adlandırılmaktadır. Kısa dönemli istatistikler sıklıkla mevsimsel hareketlerin etkisi altındadır. Mevsimsel etkiler zaman serilerinin diğer karakteristik özelliklerinin (mevsimsel olmayan özellikleri: trend ve dönüm noktası gibi) izlenmesine imkan vermez. Bu yüzden mevsimsel düzeltilmiş göstergeler, yanıltıcı mevsimsel değişiklikler olmaksızın ekonominin gerçek hareketlerinin izlenmesine olanak tanır.

Mevsim ve takvim etkisinden arındırma mevsim ve takvim etkilerinin istatistiksel yöntemler kullanılarak tahmin edilmesi ve eğer anlamlı ise veriden ayrıştırılması işlemidir. Mevsimsel düzeltme, gözlenen bir zaman serisinin, gözlemlenemeyen bileşenlerine ayrıştırılıp bu bileşenlerden biri olan mevsimsel bileşenin tahmin edilerek seriden arındırılmasıdır. Diğer bir ifade ile mevsimsel etkilerin serilerden yok oluncaya kadar filtrelenmesidir. Mevsim ve takvimden kaynaklanan etkiler, geçici nitelikte olduklarından, verinin genel eğiliminin gözlemlenmesini engellemektedir. Mevsimsel hareketler içeren veride, belirli bir dönemde meydana gelen değişikliğin, verideki gerçek artış veya azalıştan mı, yoksa mevsimsel etkilerden mi kaynaklandığını anlamak oldukça güçtür. Oysa karar verme ve planlama faaliyetleri için gerekli olan kısa dönemli göstergelerin, istenen dönemler arasında sağlıklı karşılaştırma yapmaya olanak sağlayacak şekilde üretilmesi büyük önem taşımaktadır. Kısa dönemli göstergelerde aylık/dönemlik ve yıllık değişimlerin sağlıklı bir şekilde yorumlanabilmesi için, bir önceki aya/döneme göre yapılacak karşılaştırmalarda mevsim ve takvim etkilerinden arındırılmış göstergelerin kullanılması, bir önceki yılın aynı ayına/dönemine göre yapılacak karşılaştırmalarda ise takvim etkisinden arındırılmış göstergelerin kullanılması daha anlamlı olacaktır.

Mevsimsel Düzeltme Modellerinin Belirlenmesi

Kısa dönemli istatistiklerin mevsim ve takvim etkisinden arındırılması konusunda TÜİK'te tek bir sorumlu birim bulunmaktadır. Bu birim aynı zamanda MEDOS'un yönetiminden de sorumludur. Dolayısıyla hem mevsimsel düzeltme modellerinin süreci hem de uygulaması merkezi olarak sistematik ve standart bir bakış açısıyla yürütülmektedir. Mevsim ve takvim etkilerinden arındırma süreci, her yılın sonunda bir sonraki yılın model, takvim etkisi ve aykırı değer kompozisyonunun belirlenmesi ile başlamaktadır. Belirlenen bu kompozisyon sabit tutularak, yıl boyunca her yayım dönemi için model parametreleri ve filtreler yeniden tahmin edilmektedir. Dolayısı ile mevsim ve takvim etkilerinden arındırma için kullanılan faktörler her dönem için yeniden belirlenmektedir.

Arındırılmamış zaman serilerinin modellerinin belirlenmesinde, veriler mevsimsel etkilerden arındırılmadan önce ön arındırma yapılmaktadır. Ön arındırma aşamasında, veri setine uygun dönüşümler uygulanır, aykırı değerler tespit edilir, eksik veriler tahmin edilir ve takvim etkileri istatistiksel olarak anlamlı ise veriden temizlenir. Ön-arındırmada kullanılacak ARIMA modelleri, birçok teşhis (diagnostic) istatistikleri dikkate alınarak ve aynı zamanda cimrilik kriteri (az parametre) kullanılarak belirlenir. Bu aşamaların ardından elde edilen veriler, doğrusallaştırılmış hale getirilir. Doğrusallaştırılan veriler, simetrik filtreler aracılığıyla daha hassas analizler için bileşenlerine ayrıştırılır. Bu süreçte ARIMA modelleri kullanılarak veri ileri ve geri yönlü genişletilir. Ayrıştırma modelinin tipi, veriyi oluşturan çeşitli bileşenlerin (temel olarak trend-konjonktür, mevsimsel ve düzensiz bileşenler) ham veriyi oluşturmak üzere nasıl bir araya geldiğini belirler ve buna göre ayrıştırır. Ayrıştırma işlemi için toplamsal ya da çarpımsal modeller kullanılmaktadır.

Mevsim ve takvim etkilerinden arındırma işlemi, TRAMO-SEATS (Gomez ve Maravall, 1996) yöntemi ile gerçekleştirilmektedir. Kullanılan yöntem, Avrupa Birliği İstatistik Ofisi (EuroStat)'nin tavsiyeleri çerçevesinde belirlenmiştir. Bu yöntemin uygulanmasında Belçika Merkez Bankası tarafından Almanya Merkez Bankası (Deutsche Bundesbank), Fransa İstatistik Ofisi (Insee) ve Avrupa Birliği İstatistik Ofisi işbirliğiyle Avrupa İstatistik Sistemi (ESS) Yönergelerine uygun olarak geliştirilen tarafından geliştirilen JDemetra+ yazılımı kullanılmaktadır. Mevcut durumda 12 farklı konu başlığında istatistiğin toplamda 2000'e yakın sayıda zaman serisi değişkeni için her yılın başında modeller belirlenir ve bu modeller kullanılarak her dönem mevsim ve takvim etkilerinden arındırılmak işlemi gerçekleştirilir. Bu sayılara yeni istatistiki konu başlıkları da ihtiyaç oldukça eklenmektedir.

Mevsimsel Düzeltme Otomasyon Sistemi

Mevsimsel Düzeltme Otomasyon Sistemi (MEDOS) başlangıçta birkaç istatistik konu başlığı ile başlayan ve zaman içerisinde ihtiyaç oldukça yeni istatistik konu başlıklarının eklenmesi ile TÜİK içerisinde mevsim ve takvim etkilerinden arındırılmış zaman serilerini üretmek ve yayımlamak isteyen birimler için kullanışlı bir araç haline gelmiştir. Verilerin sistem içerisine aktarılması ile başlayan süreç içerisinde veri hazırlama, veri işleme, modellerin tetiklenmesi, gerekli çıktı tablolarının farklı zaman serisi bileşenleri için hazırlanması, kontrol süreçleri ve tüm bunların tek bir sayfa üzerinden görüntülenmesi ve analizi gibi alt süreçler yer almaktadır. Bu alt süreçlerin bir kısmı bütün istatistik konu başlıkları için ortak iken bazı alt süreçler istatistik konu başlığına göre farklılık göstermektedir. Bu nedenle MEDOS tüm bu süreçlerin tek bir noktadan yönetilmesi, hızlı bir şekilde sonuç üretmesi ve süreçleri standart hale getirmesi açısından oldukça kullanışlı ve önemli bir araçtır.

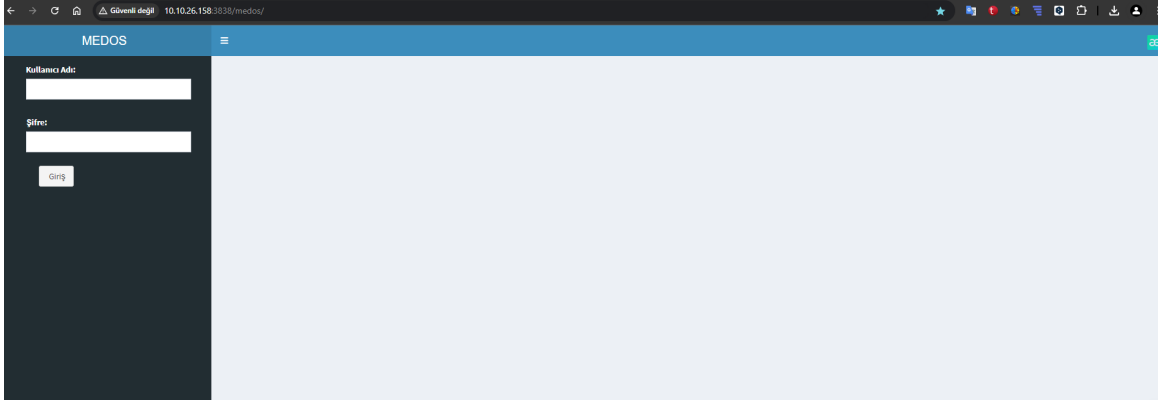
Bu aracın oluşturulması için R programlama ve Shiny paketi kullanılmıştır. R programı, istatistiksel analizler ve veri bilimi alanında güçlü ve esnek bir araç olarak öne çıkmaktadır. Açık kaynaklı olması, dünya genelindeki akademik ve profesyonel topluluklar tarafından sürekli geliştirilmesine olanak tanımaktadır. Özellikle zaman serisi analizleri, regresyon modellemeleri ve veri görselleştirme gibi karmaşık istatistiksel işlemler için geniş bir fonksiyon yelpazesi sunar (Wickham ve Grolemond, 2017). R, Shiny gibi interaktif uygulama geliştirme paketleri sayesinde, veri analiz sonuçlarının kolayca paylaşılabılır ve kullanıcı dostu ara yüzlerle sunulmasını sağlamaktadır (Xie vd., 2021). Ayrıca, R'nin zengin paket ekosistemi ve güçlü grafik kütüphaneleri, büyük veri setleriyle etkili bir şekilde çalışmayı mümkün kılar. Bu nedenle, R programı özellikle veri bilimi projelerinde esnekliği, hız ve doğruluğu artıran bir platform olarak yaygın bir şekilde tercih edilmektedir.

Shiny ise R programlama dilinin gücünü web tabanlı etkileşimli uygulamalar aracılığıyla genişleten bir pakettir. Shiny, kullanıcıların istatistiksel analiz ve veri görselleştirme sonuçlarını dinamik ve kullanıcı dostu bir arayüzde sunmalarını sağlar. Kod bilgisi sınırlı olan kullanıcılar dahi, Shiny uygulamaları sayesinde karmaşık veri setlerini görselleştirip analiz edebilir. Ayrıca, Shiny'nin HTML, CSS ve JavaScript ile entegrasyonu, uygulamaların görsel olarak zengin ve işlevsel olmasına olanak tanır. Shiny, veri bilimi projelerinde etkileşimli dashboard'lar oluşturarak analistlere ve karar vericilere anlık geri bildirim sağlayan güçlü bir araç haline gelmiştir. Bu esneklik, veri analizlerinin daha hızlı ve etkili bir şekilde paylaşılmasını sağlarken, aynı zamanda analiz sonuçlarını görselleştirme ve anlamlandırma süreçlerini de büyük ölçüde kolaylaştırmaktadır (Chang vd., 2021).

MEDOS'un ürettiği arındırılmış sonuçlar için belirlenen modeller JDemetra+ masaüstü yazılımı kullanılarak oluşturulur ve MEDOS içerisinde ise JWSACruncher aracı ile tetiklenerek ilgili sonuçlar üretilir. JWSACruncher, JDemetra+ yazılımının bir bileşeni olarak geliştirilmiş ve zaman serisi verilerinin mevsimsel düzeltilmesi için kullanılan komut satırı tabanlı bir araçtır. JDemetra+ gibi, Avrupa İstatistik Ofisi (Eurostat) tarafından resmi istatistiklerde mevsimsel düzeltme için önerilen bu araç, büyük miktarda veriyi otomatikleştirilmiş bir şekilde işlemek üzere tasarlanmıştır. JWSACruncher, JDemetra+'ın arka planda çalışan altyapısını kullanarak, kullanıcıların manuel müdahale gerektirmeden toplu işlemler yapmasına olanak tanır. Bu araç, özellikle yüksek hacimli veri setleri üzerinde hızlı ve tutarlı mevsimsel düzeltme işlemleri gerçekleştirmek isteyen analistler ve kurumlar için idealdir. API entegrasyonu sayesinde, JWSACruncher diğer sistemlerle kolayca entegre edilebilir, böylece mevsimsel düzeltilmiş verilerin raporlanması ve analiz edilmesi süreçlerinde hız ve verimlilik sağlar (Eurostat, 2015). Özellikle kamu istatistik kurumları ve büyük veri setleriyle çalışan organizasyonlar, bu araç sayesinde işlem sürelerini önemli ölçüde azaltabilir ve sonuçların tutarlılığını artırabilir.

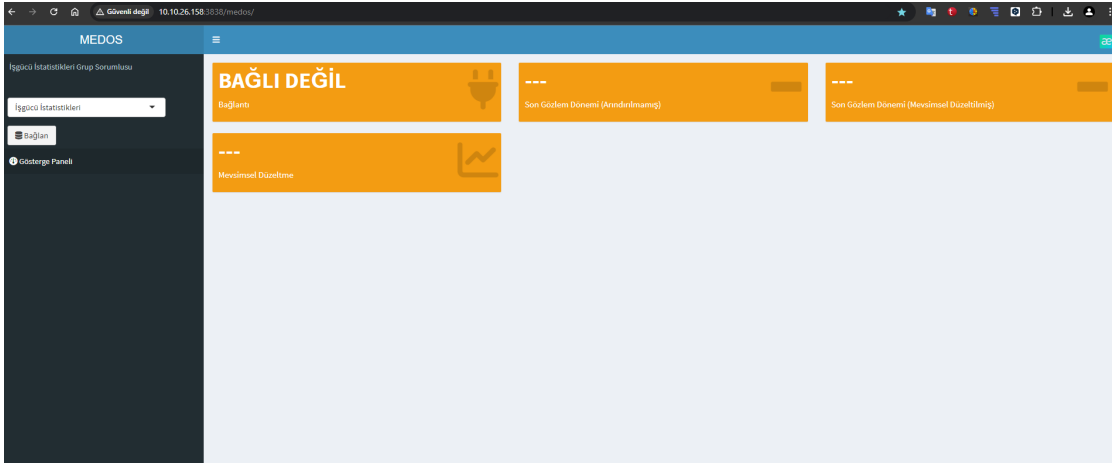
MEDOS uzak makine üzerine kurulu Shiny Server üzerinde çalışan ve sadece TÜİK ağı üzerinden erişilebilen bir Shiny uygulamasıdır. Shiny Server, R dilinde yazılmış Shiny uygulamalarını web üzerinden erişilebilir hale getiren bir yazılım altyapısıdır. Sonuçların üretilmesi için kullanılan modeller ve gerekli dosyalar sistem yöneticisi tarafından Shiny Server üzerine her yılın başında yüklenmekte ve yıl boyunca kullanılmaktadır. MEDOS kullanıcıları bir tarayıcıya uzak makine adresini girerek uygulama sayfasına erişirler. Sonrasında uygulama sayfasını açmak kendi kullanıcı adı ve şifrelerini kullanmaktadırlar. Kullanıcı adı ve şifreler sistem yöneticisi tarafından kullanıcılara sağlanır. Kullanıcılar uygulamaya giriş yaptıktan sonra tek bir buton ile sistemi çalıştırarak cari döneme ilişkin

arındırılmamış, mevsim ve takvim etkisinden arındırılmış, takvim etkisinden arındırılmış ve trend-cycle sonuçlarını ve bunların büyüme tablolarını istenilen sınıflama detaylarında elde edebilmektedirler. Şekil 1’de MEDOS giriş sayfası ekran görüntüsü verilmiştir.



Şekil 21. MEDOS giriş ekranı

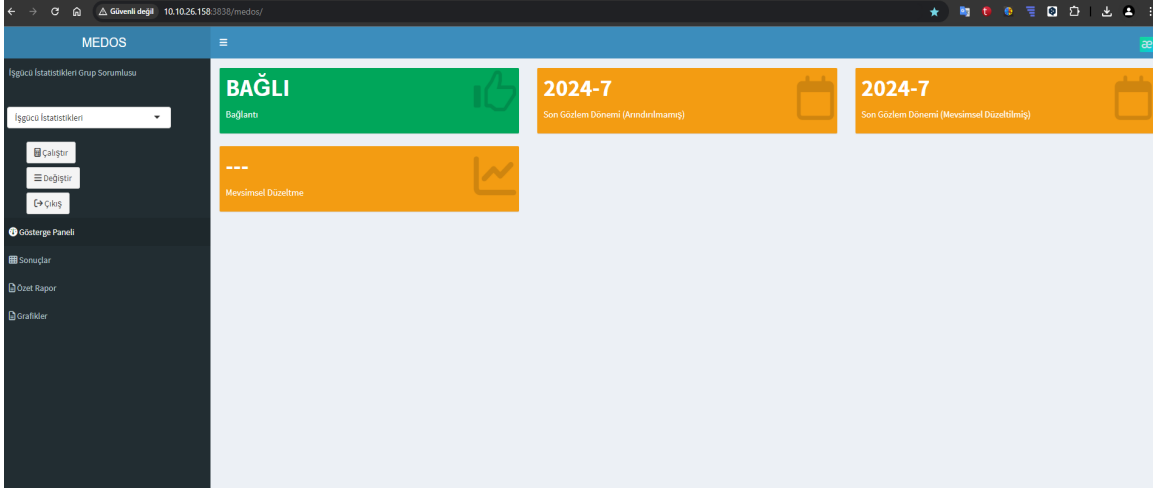
Kullanıcı uzak makine adresindeki 3838 portu üzerinden MEDOS uygulamasına erişebilmektedir. Varsayılan olarak Shiny Server, uygulamaları sunarken 3838 portunu kullanır. Bu, Shiny uygulamasının yerel bir sunucuda veya bir web sunucusunda çalıştırılması durumunda, kullanıcıların tarayıcılarından uygulamaya erişim sağlayabilmeleri için kullanılan port numarasıdır. Bir kullanıcı sisteme giriş yaptıktan Şekil 2’ verilen ara yüz ekranı ile karşılaşır.



Şekil 22. MEDOS kullanıcı gösterge paneli

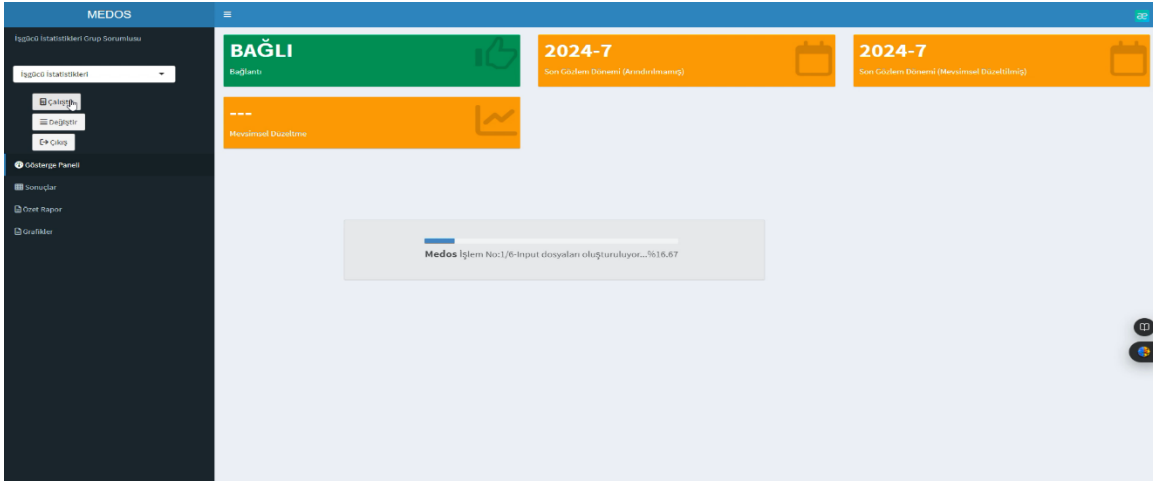
Gösterge paneli MEDOS sisteminin çalışabilir olduğunu denetleyebilmek için bazı göstergelerin yer aldığı bir paneldir. Burada veri tabanı bağlantısı kontrolü, arındırılmamış istatistikler için son gözlem dönemi, mevsimsel düzeltilmiş istatistikler için son gözlem dönemi ve mevsimsel düzeltme işleminin başarılı bir şekilde gerçekleşip gerçekleşmediğinin kontrolünü yapan bazı kutucuklar görülmektedir.

Bağlantı kutucuğu, mevsimsel düzeltilecek ham veriyi veri tabanından başarılı bir şekilde alınıp alınmadığını kontrol etmek amaçlı konulmuştur. İki tür gösterim içermektedir. Eğer "BAĞLI DEĞİL" yazısı kutucukta görünüyorsa veri tabanından veri henüz çekilmemiştir veya bağlanma komutu verildikten sonra veri tabanından ilgili veri çekilememiştir. Eğer "BAĞLI" yazısı kutucukta görünüyorsa bağlanma komutu verildikten sonra veri tabanından veri başarılı bir şekilde çekilmiş anlamına gelmektedir. "Bağlan" butonuna basıldığında Şekil 3'teki ekran görüntüsü açılacaktır.



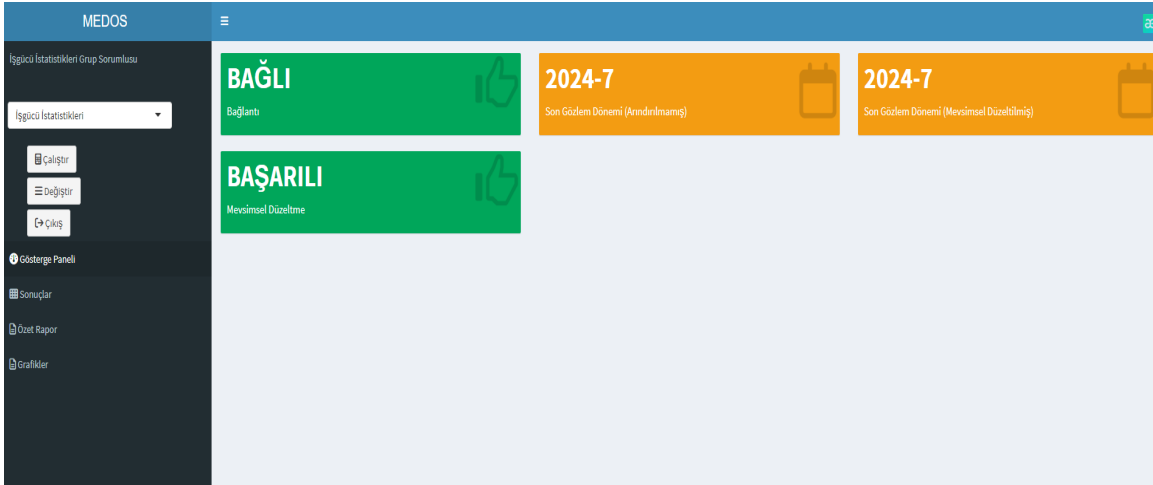
Şekil 23. MEDOS bağlantı ekranı

Bu görüntüde önemli olan iki buton "Çalıştır" ve "Değiştir" butonlarıdır. "Değiştir" butonu eğer bir kullanıcı birden fazla çalışmaya MEDOS içerisinden ulaşabiliyorsa çalışmayı değiştirmek için tıkladığında diğer çalışmalara erişebilmesi için konulmuştur. Diğer yandan bağlantı kutucuğunda yeşil renk ile "BAĞLI" yazmaktadır. Bu veritabanına bağlantının başarılı şekilde gerçekleştiğini göstermektedir. Gösterge panelindeki diğer göstergelerden arındırılmamış son gözlem dönemi (bu örnek için) "2024-7" şeklinde görünmektedir. Bu ifade veri tabanına üretici birim tarafından atılan son gözlemin verisini içermektedir. Mevsimsel düzeltilmiş son gözlem dönemi ise en son çalıştırılan mevsim düzeltilmiş verinin tarihini içermektedir. Mevsimsel düzeltme kutucuğunda ise "---" ifadesi yer almaktadır. Bu ifade de MEDOS açıldıktan ve veri tabanına bağlandıktan sonra hiç mevsimsel düzeltme yapılmadığını ifade etmektedir. "Çalıştır" butonuna basıldığında ekrana bir bilgi kutucuğu açılır ve sürecin ilerlemesinin izlenmesi sağlanır. Şekil 4'te buna ilişkin ekran görüntüsü verilmiştir.

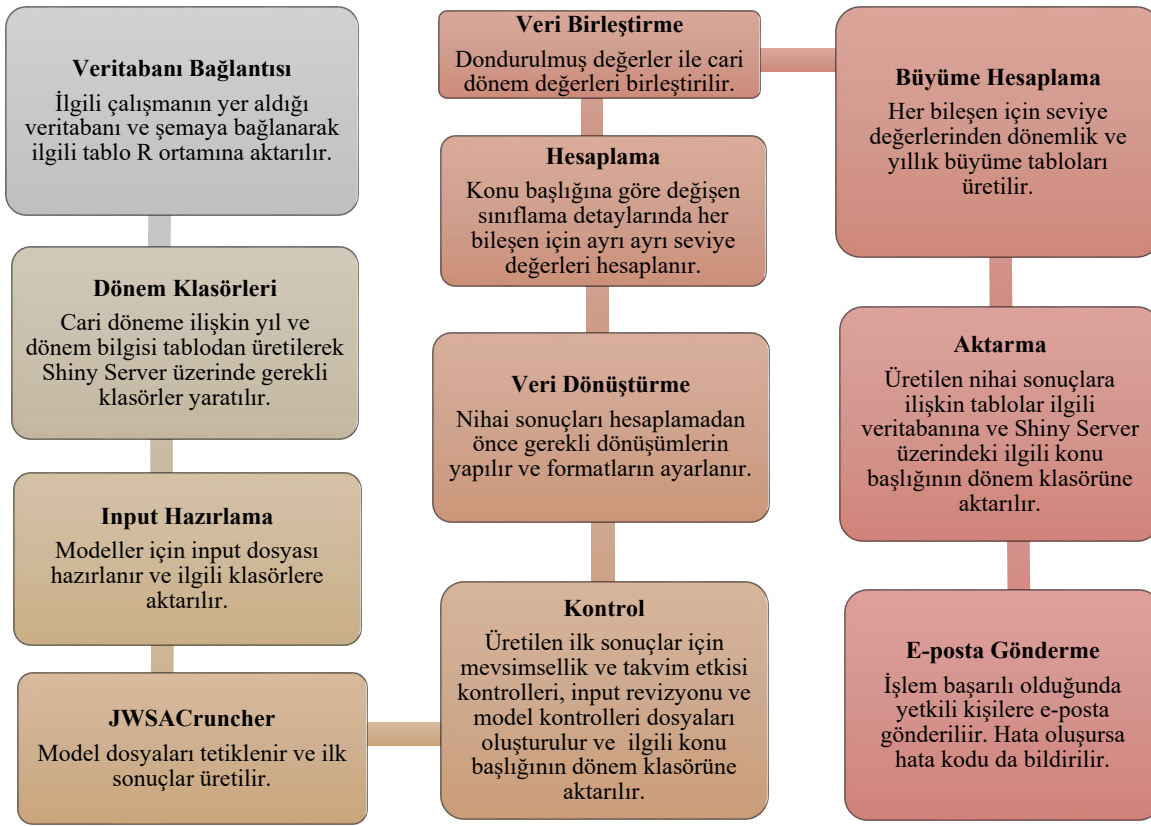


Şekil 24. MEDOS süreç bilgi kutucuğu

Süreç başarılı bir şekilde gerçekleştikten sonra ise Şekil 5'te yer alan mevsimsel düzeltme sonrası gösterge paneli ortaya çıkmaktadır. "Çalıştır" butonuna basılması ile başlayan ve sistemin başarılı bir şekilde tamamlanmasına kadar arka planda çalışan süreç adımları ise şekil 6'da belirtilmiştir.



Şekil 25. MEDOS başarılı mevsimsel düzeltme ekranı



Şekil 26. MEDOS çalışma süreci

Sistem başarılı bir şekilde çalıştıktan sonra kullanıcı bu ara yüz üzerinden sonuçları tablolar veya grafikler üzerinden değişken detayında inceleyebilir. Bunu yanında tabloları ve modellere ilişkin teşhis sonuçlarını bilgisayarını da indirebilir. Ara yüzün sol yanında yer alan "Sonuçlar" butonuna basıldığında Şekil 7'de verilen görüntü elde edilecektir.

Date	TOPLAM_15_ARTI_NUFUS	TOPLAM_15_ARTI_ISG	TOPLAM_15_ARTI_IST	TOPLAM_15_ARTI_ISSIZ	TOPLAM_15_ARTI_IDO	TOPLAM_15_ARTI_IKO	TOPLAM_15_ARTI_IST_ORAN	TOPLAM_15_ARTI_ISSZ_ORAN
2023-07	65,446.00	35,310.00	31,873.00	3,437.00	30,137.00	54.00	48.70	9.70
2023-08	65,494.00	35,316.00	32,071.00	3,245.00	30,178.00	53.90	49.00	9.20
2023-09	65,542.00	35,321.00	32,165.00	3,155.00	30,222.00	53.90	49.10	8.90
2023-10	65,589.00	34,851.00	31,985.00	2,865.00	30,738.00	53.10	48.80	8.20
2023-11	65,637.00	35,036.00	31,961.00	3,075.00	30,601.00	53.40	48.70	8.80
2023-12	65,683.00	35,259.00	32,113.00	3,146.00	30,424.00	53.70	48.90	8.90
2024-01	65,728.00	34,930.00	31,523.00	3,407.00	30,796.00	53.10	48.00	9.80
2024-02	65,763.00	35,513.00	32,263.00	3,250.00	30,250.00	54.00	49.10	9.20
2024-03	65,799.00	35,361.00	32,252.00	3,109.00	30,438.00	53.70	49.00	8.80
2024-04	65,834.00	35,244.00	32,266.00	2,978.00	30,590.00	53.50	49.00	8.50

Şekil 27. MEDOS Sonuçlar ekranı

Bu görüntü bize arındırılmamış, mevsim ve takvim etkilerinden arındırılmış seviye, takvim etkilerinden arındırılmış seviye, eğilim-döngü (trend-cycle) seviye, mevsim ve takvim etkilerinden arındırılmış büyüme, takvim etkilerinden arındırılmış büyüme, eğilim döngü büyüme tablolarını göstermektedir. Örneğin Şekil 7 arındırılmamış seviye değerlerini değişken detayında göstermektedir. Diğer sekmeler ise yukarıda bahsedilen diğer detaylarda sonuçları içermektedir. Aynı zamanda her bir sayfanın altında "İndir" butonu yer almaktadır. Bu buton bize programda üretilen istatistikleri kendi bilgisayarımıza indirme imkanını vermektedir.

Ekranın sol kısmında yer alan "Özet Rapor" kısmına gelindiğinde ise program, üretim aşamasında kullandığı ve ürettiği verilerle alakalı bazı özet raporları sunmaktadır. Şekil 8'de buna ilişkin ekran görüntüsü verilmiştir.

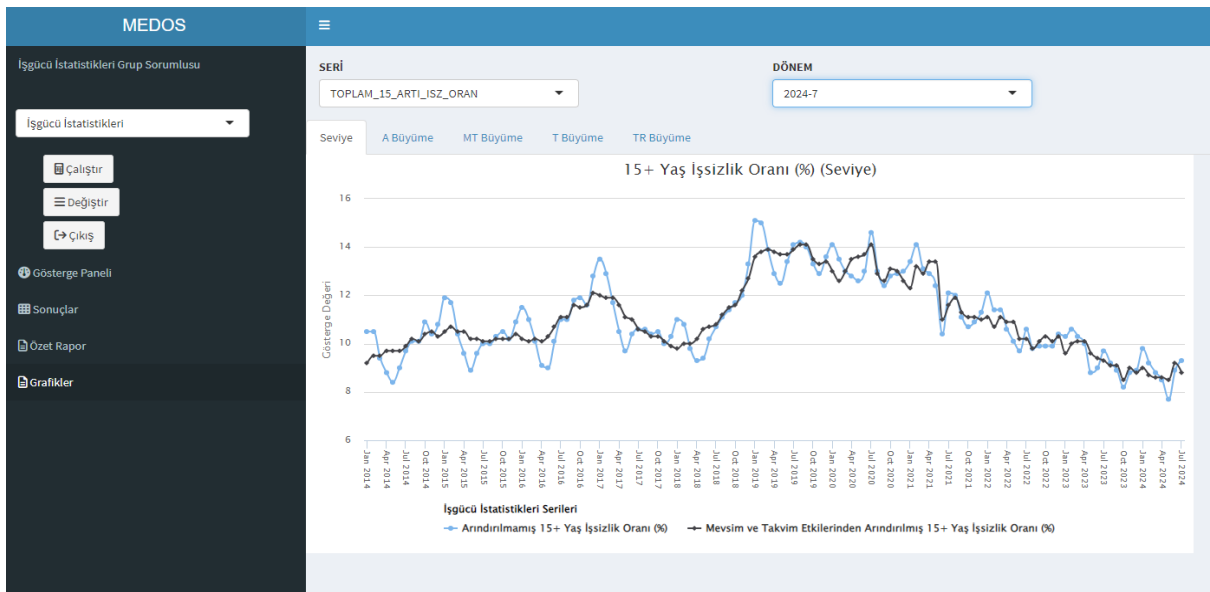
MEDOS	
<p>İşgücü İstatistikleri Grup Sorumlusu</p> <p>İşgücü İstatistikleri</p> <p>Çalıştır</p> <p>Değiştir</p> <p>Çıkış</p> <p>Gösterge Paneli</p> <p>Sonuçlar</p> <p>Özet Rapor</p> <p>Grafikler</p>	<p>ISG</p> <p>2024-7</p> <p>SERİ SAYISI</p> <p>Doğrudan: 22</p> <p>Dolaylı: 73</p> <p>Mevsim ve Takvim: 4</p> <p>Mevsim: 11</p> <p>Takvim: 0</p> <p>Orjinal: 7</p> <p>NEGATİF DEĞER VARLIĞI</p> <p>Arındırılmamış: 0 (% 0)</p> <p>Mevsim ve Takvim Etkilerinden Arındırılmış: 0 (% 0)</p> <p>Takvim Etkilerinden Arındırılmış: 0 (% 0)</p> <p>Diagnostic İndir</p> <p>İndir</p>

Şekil 28. MEDOS özet rapor

Yukarıdan aşağıya bakılacak olursa ilk kutucukta programın mevsimsel düzeltmek için kullandığı verinin kısaltmasını göstermektedir. Örneğimizde bu veri "ISG" işgücü istatistiklerini ifade eden kısaltmadır. Yine aynı kutucukta ise mevsimsel düzeltme uyguladığı dönemi göstermektedir. Örneğimizde bu "2024-7" dönemidir. Yukarıdan aşağıya ikinci kutucuğa bakacak olursak bize seri sayılarıyla alakalı bazı istatistikler sunmaktadır. Bu

istatistikler doğrudan yöntemle elde edilen seri sayısı, dolaylı yöntemle elde edilen seri sayısı, mevsim ve takvim etkisi barındıran seri sayısı, sadece mevsim etkisi barındıran seri sayısı, sadece takvim etkisi barındıran seri sayısı, ne mevsim ne de takvim etkisi barındıran seri sayısı yani hiçbir mevsim ve takvim etkisi barındırmayan seri sayısını göstermektedir. Örnek vermek gerekirse mevsim ve takvim etkisi barındıran seri sayısının örneğimizde 4 olduğu gösterilmektedir. Yukarıdan aşağıya bakıldığında son sıradaki kutucuk ise bize arındırılmamış serilerde, mevsim ve takvim etkilerinden arındırılmış serilerde ve sadece takvim etkilerinden arındırılmış serilerde negatif değerlerin olup olmadığını göstermektedir. Örneğimize bakılacak olursa her 3 tür seri için de negatif değer varlığı görülmektedir. Program tüm bu istatistikleri ürettikten sonra bize csv olarak bilgisayarıma indirme imkanı sunmaktadır.

Ekranın sol kısmında yer alan “Grafikler” kısmına tıkladığında program bize seri ve dönem bazında grafikler üretebilmektedir. Buna ilişkin ekran görüntüsü Şekil 9’da verilmiştir.



Şekil 29. MEDOS grafikler

MEDOS grafikleri arındırılmamış yani "Seviye", arındırılmamış büyüme yani "A Büyüme", mevsim ve takvim etkisinden arındırılmış yani "MT Büyüme", takvim etkisinden arındırılmış büyüme yani "T Büyüme", eğilim-döngü büyüme yani "TR Büyüme" bazında üretebilmektedir. Ayrıca istenilen seri yani değişken ve döneme göre interaktif olarak grafiksel analiz yapmak da mümkündür.

TARTIŞMA VE SONUÇ

Bu çalışma, ekonomik analizlerde büyük önem taşıyan mevsimsel düzeltme sürecini ele alarak, Türkiye İstatistik Kurumu'nun (TÜİK) geliştirdiği MEDOS (Mevsimsel Düzeltme Otomasyon Sistemi) aracılığıyla bu sürecin nasıl merkezileştirildiğini ve otomatik hale getirildiğini incelemiştir. Mevsimsel dalgalanmaları ortadan kaldırarak gerçek trendleri daha doğru bir şekilde ortaya çıkarmak, veri bilimi ve istatistik dünyasında önemli bir rol oynamaktadır. TÜİK'in MEDOS uygulaması, R ve Shiny teknolojileri ile geliştirilmiş olup, karmaşık zaman serisi analizlerini otomatik hale getirerek insan hatasını en aza indirmekte ve işlem süresini büyük ölçüde kısaltmaktadır. Yapay zeka ve gelişmiş algoritmaların entegrasyonu sayesinde, istatistiksel analizlerde hız ve doğruluk artarken, analistlerin daha stratejik işlere odaklanmasına olanak sağlanmaktadır.

Bu sonuçlar, veri bilimi ve istatistik alanında yapay zeka destekli otomasyon sistemlerinin ne kadar etkili olduğunu bir kez daha göstermektedir. MEDOS'un kullanımı, sadece zaman serisi analizlerinde standartlaşma ve verimliliği artırmakla kalmamış, aynı zamanda esnek yapısı sayesinde TÜİK'in gelecekteki ihtiyaçlarına hızla adapte olmasına olanak tanımıştır. Bu, yapay zeka ve veri bilimi teknolojilerinin kamu kurumlarında daha yaygın kullanımının, hem süreçleri kolaylaştıracağını hem de karar vericilere daha güvenilir ve güncel veriler sunacağını ortaya koymaktadır.

Sonuç olarak, TÜİK'in MEDOS sistemi, R ve Shiny teknolojileri ile desteklenen yenilikçi bir çözüm sunarak, veri bilimi ve istatistiksel analizlerde yeni bir standart belirlemiştir. Bu sistem, diğer kamu kurumlarına da veri işleme ve analiz süreçlerinde rehberlik edebilecek bir örnek teşkil etmektedir.

Referanslar

Gomez, V. ve Maravall, A., (1996), “Programs TRAMO (Time series Regression with Arima noise, Missing observations and Outliers) and SEATS (Signal Extraction in Arima Time Series) Instructions for the User”, Documento de Trabajo 9628, Servicios de Estudios, Banco de Espana.

Xie, Y., Cheng, J., & Tan, X. (2021). R Markdown: The Definitive Guide. Chapman and Hall/CRC.

Wickham, H., & Grolemund, G. (2017). R for Data Science: Import, Tidy, Transform, Visualize, and Model Data. O'Reilly Media.

Chang, W., Cheng, J., Allaire, J., Xie, Y., & McPherson, J. (2021). Shiny: Web Application Framework for R.

Eurostat. (2015). JDemetra+ User Guide. <https://ec.europa.eu/eurostat>.

Yapay Zeka Destekli Veri Sınıflandırma ve Yönetim Modeli

Tülin Demiralp^{1*}, Ahmet Künyeli²⁺, Necati Kaya³ ve Maryat Coşkun⁴

^{1,2,3,4}Barikat İnternet Güvenliği Bilişim Ticaret A.Ş. Ar-Ge Merkezi/ Ankara, Türkiye

Özet

Bilgisayar tabanlı veri artışı ile birlikte, bilginin depolanması, yönetimi, işlenmesi, güvenliği ve düzenlenmesi güçleşmekte, mobil uygulamaların, sensörlerin ve bulut bilişim çözümlerinin artmasından kaynaklı, devletlerin ve organizasyonların veri üretme ve arşivleme oranları da yükselmektedir. Pek çok sektörde işlemlerin tamamıyla elektronik ortama taşınması, organizasyonların depoladıkları hassas veri büyüklüğü ve çeşitliliğini artırmış, veriyi toplayan, araştıran ve analiz eden yeni araçlar vasıtasıyla yapısal olmayan veriden de faydalanmalarını olanaklı kılmıştır. Ancak gizlilik, bütünlük ve şeffaflık gibi önemli unsurların sağlanamadığı veri yönetim yöntemleri kişisel, kurumsal ve ulusal güvenlik risklerini beraberinde getirmektedir. Bu risklerin minimize edilmesi amacıyla özellikle kişisel verinin korunumunun sağlanması için, Türkçe kurumsal dokümanlar içerisinde doğal dil işleme yöntemlerine dayalı şekilde sınıflandırma yapan çözümler geliştirilmektedir. Bu çalışmanın amacı hassas verinin analizinde yapısal ve yapısal olmayan verinin keşfedilmesi, korunması gereken verinin sınıflandırılması ve hatalı veri yönetiminden kaynaklanabilecek güvenlik risklerinin değerlendirilmesini temel alan bir veri yönetim döngüsünü sağlamak amacıyla, derin öğrenme modellerine dayanan algoritmalarla veriyi sınıflandıran, risk puanlamasını gerçekleştiren ve KVKK uyumluluk gerekliliklerinin yerine getirilmesinde kullanılmak üzere bir veri güvenliği yönetim modeli elde edilmesidir. Çalışma kapsamında, kimlik belgesi, adli sicil kaydı, özgeçmiş, sertifika, işe giriş, idari şartname ve teknik şartnameden oluşan 7 tipte toplam 2392 Türkçe kurumsal doküman barındıran bir veri seti hazırlanmıştır. Veri ön işleme, tokenizasyon, kelime köklerinin bulunması, veri temizleme ve kelimelerin vektör değerlerinin bulunması işlemlerinin ardından eğitim setindeki dokümanlar kullanılarak BERT Doğal Dil İşleme modelinin eğitilmesi ve yeni bir model ağırlık matrisinin elde edilmesi sağlanmıştır. Eğitim sonrasında 7 tip doküman seti için modelin Kesinlik ve Duyarlılık değerlerinin ağırlıklı harmonik ortalaması (F1 Puanı) 0.6 olarak elde edilmiştir. Veri işleme çalışmaları sırasında kişisel veriler için gizlilik, bütünlük ve erişilebilirlik açısından 1'den 5'e kadar puanlama yapılarak, modelin veriyi sınıflandırmasının ardından risk seviyesini hesaplaması da sağlanmıştır.

Anahtar Kelimeler – Veri güvenliği, Kişisel verinin korunması, Veri gizliliği, Veri sınıflandırma

GİRİŞ

Günümüzde dijitalleşmenin hız kazanmasıyla birlikte, kişisel verilerin elektronik ortamda depolanması ve işlenmesi yaygınlaşmıştır. Özellikle 25 Mayıs 2018'de yürürlüğe giren Avrupa Birliği Genel Veri Koruma Tüzüğü (GDPR) ile birlikte, hem kişisel hem de kurumsal hassasiyet taşıyan dokümanların korunması hayati bir gereklilik haline gelmiştir (Safari, 2016). Dijital ortamda toplanan bu verilerin güvenliği, sadece bireylerin mahremiyetini korumak açısından değil, aynı zamanda kurumların itibarını ve hukuki sorumluluklarını yerine getirmeleri açısından da önem arz etmektedir.

Koruma altına alınmayan hassas veriler, siber saldırılar, veri ihlalleri, yetkisiz erişimler ve yanlış kullanımlar gibi ciddi riskleri barındırmaktadır. Bu tür olaylar, kurumların yüksek cezalarla karşılaşmasına, itibar kaybına ve maddi zarara uğramalarına yol açmaktadır. Aynı zamanda, bireyler açısından kimlik hırsızlığı, dolandırıcılık ve kişisel verilerin sızması gibi olumsuz sonuçlar doğurabilmektedir. Bu nedenle, dijitalleşme ile birlikte artan kurumsal dokümanların korunması gerekliliği, yalnızca bir güvenlik meselesi değil, aynı zamanda hukuki ve etik bir zorunluluk olarak karşımıza çıkmaktadır.

Kurum ve kuruluşların işlediği ve depoladığı doküman çeşitliliği ve hacminin artmasıyla birlikte, doküman yönetiminin el yordamıyla yapılması da zorlaşmıştır. Dokümanların kategorilerine göre otomatik olarak sınıflandırılması ve buna uygun envanter oluşturma kurallarının belirlenerek güvenli bir şekilde saklanması kritik bir gereklilik haline gelmiştir.

Bu çalışmada, hassas veri içeren kurumsal dokümanların doküman kategorilerine göre sınıflandırılması için bir yapay sinir ağı modeli kademeli (incremental) eğitim metoduyla eğitilmiştir. Model ilk olarak kategorisi önceden bilinen ve buna göre etiketlenmiş doküman setiyle denetimli eğitime tabi tutulmuştur. Kategorisi önceden bilinmeyen dokümanlar için bir denetimsiz hiyerarşik kümeleme (unsupervised hierarchical clustering) çalışması

gerçekleştirilmiş ve bu çalışma sonucu elde edilen yeni kategoriler ve bu kategorilere atanmış dokümanlar ilk veri setine eklenerek sinir ağı tekrar eğitilmiştir. Çalışmada yapay sinir ağı modelinin sınıflandırdığı dokümanlar için bir risk skorlama çerçevesi de oluşturulmuştur.

VERİ SETİ VE VERİ ÖN-İŞLEME

Sınıflandırma modelinin eğitiminde, Barikat'ın kurumsal veri tabanında depolanan ve kişisel hassas bilgileri içeren dokümanlar kullanılmıştır. Bu kapsamda, 75 adet kimlik belgesi, 114 adet adli sicil kaydı, 1.191 adet özgeçmiş, 796 adet sertifika, 44 adet işe giriş sözleşmesi, 100 adet idari şartname ve 72 adet teknik şartname olmak üzere yedi farklı doküman tipinden oluşan toplam 2.392 doküman derlenmiştir. Veri seti, yapılandırılmış (standart formlar, Word dokümanları, PDF dosyaları) ve yapılandırılmamış (veri tabanı tabloları, grafikler, resim dosyaları, free-text formatındaki metin dosyaları, vb.) dokümanlardan oluşmaktadır.

Denetimli eğitim süreci için veri setinin oluşturulması aşamasında dokümanlar, kategorilerine göre el yordamıyla etiketlenerek indekslenmiş, yapılandırılmamış dokümanlar için farklı indeksleme süreçleri işletilmiştir. Yapılandırılmış dokümanlar dosya adı, dosya adresi, dokümanı oluşturan kullanıcı ismi, dokümanın en son değiştirildiği/oluşturulduğu tarih, vb. içerikten bağımsız üst bilgilere göre, yapılandırılmamış dokümanlar ise veri tabanı, şema, tablo ismi, vb. üst bilgilere göre indekslenmiştir. Veri üst bilgileri kategorik yapıda olduğundan yapay zekâ algoritmaları tarafından kullanılabilmesi için kodlanarak sayısal değerlere çevrilmiştir. Bu dönüşüm sırasında her bir değişkenin 0-1 arasında bir değer alabilmesi için 1 no'lu formül kullanılmıştır.

$$\text{(değer-MIN)/(MAX-MIN)} \quad (1)$$

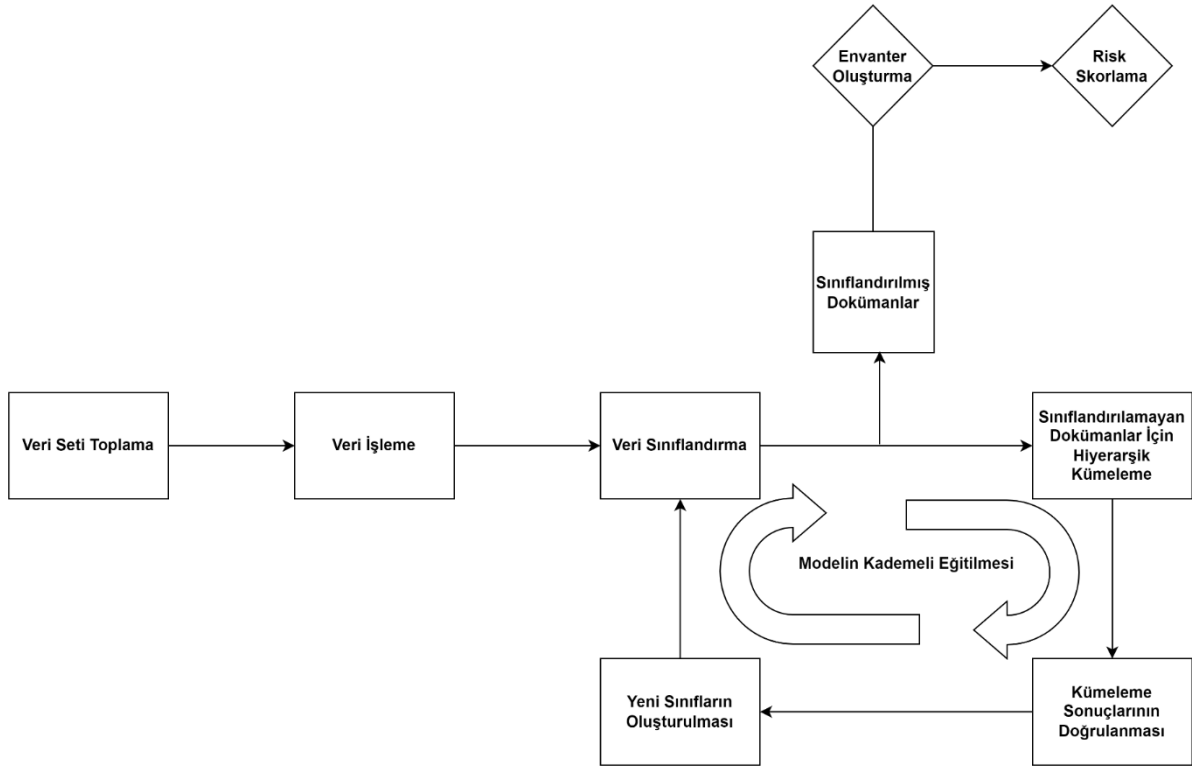
Etiketleme ve indeksleme çalışmalarının sonrasında metin verileri üzerinde ön işleme yapılmış; bu aşamada özel karakterler, sayılar ve gereksiz sözcükler (stop words) kaldırılmış, kelimeler küçük harfe dönüştürülmüştür.

KADEMELİ EĞİTİM METODOLOJİSİ

Kurumsal dokümanların sınıflandırılması, doküman yığınlarının büyüklüğü ve sürekli güncellenen dinamik yapıları nedeniyle karmaşık bir problem haline gelmiştir. Bu probleme bir çözüm olarak bu çalışmada, kademeli (incremental) öğrenme ve metin madenciliği yöntemlerinden biri olan artırımsal hiyerarşik kümeleme yöntemi (agglomerative hierarchical clustering) birleştirilerek dinamik doküman yığınlarının etkili bir şekilde sınıflandırılması için bir çerçeve önerilmiştir.

Kademeli eğitim, bilginin zaman içinde birikmesi ve yönetilmesi sürecini ifade etmektedir (Siddiqui ve Park, 2021) ve önceki eğitim verilerinden öğrenilen bilgiyi unutmadan yeni bir eğitim veri setinden yeni bilgi öğrenmeyi amaçlamaktadır (Chen vd., 2007).

Kademeli eğitim metodolojisi, sınıflandırma modelinin yeni kategoriler ve özniteliklerle sürekli olarak güncellenebilmesine olanak tanımayı ve dinamik veri setleri karşısında modelin doğruluğunu yüksek tutmayı amaçlamaktadır. Çalışmada uygulanan kademeli eğitim akışı Şekil 1'de gösterilmiştir.



Şekil 1. Yapay sinir ağının kademeli eğitim akışı

İlk Doküman Seti ile Eğitim

İlk aşamada, 7 farklı kategoriye göre etiketlenmiş dokümanlardan oluşan, yukarıda bahsedilen veri ön işleme adımları ile temizlenen ve formatlanan doküman seti (D_0) modelin ilk eğitim sürecini başlatmak için kullanılmıştır.

Özniteliklerin Çıkarılması ve Vektörleştirilmesi

Metin sınıflandırma görevlerinde, metni temsil eden anlamlı özniteliklerin çıkarılması kritik bir adımdır (Dzisević ve Šešok, 2019). Ham metnin yapay sinir ağı tarafından işlenebilir bir formatta temsil edilmesi, bunun için kelimelerin birbirleriyle olan ilişkilerini ve bağlamlarını göz önünde bulunduran ve sınıflandırma algoritmalarının kullanabileceği öznitelik vektörlerine dönüştürülmesi gerekmektedir. Öznitelik çıkarma ve vektörleştirme işlemi, metindeki önemli bilgi noktalarının temsilini ve modelin sınıflandırma performansını doğrudan etkilemektedir (Wang ve Zhang, 2013). Bağlamsal olarak doğru öznitelikler çıkarılmadığında, modelin performansı düşebilmekte ve yanlış sınıflandırmalar ortaya çıkabilmektedir.

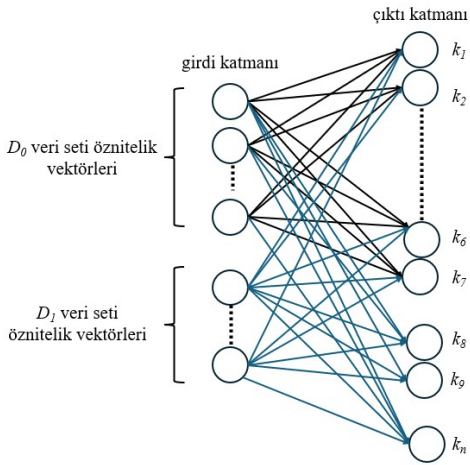
Çalışmada, dokümanlardan öznitelik çıkarma ve vektörleştirme işlemi için ilk olarak iki alternatif metot denenmiştir. Metin gömme (embedding) modelinin oluşturulması amacıyla, indekslenen dokümanların korpusunda (corpus) bulunan her bir özgün kelime için bir vektör oluşturulmasında Word2Vec yöntemi uygulanmıştır. Word2Vec, kelimeleri sabit boyutlu vektörlerle temsil eden ve bu vektörleri kelimelerin genel kullanımına dayalı olarak öğrenen bir modeldir (Zhang ve Wang, 2020). Bu yöntemle, dokümanlardaki kelimeler birbirine benzeyen diğer kelimelerle olan ilişkilerine göre vektörlere dönüştürülmüştür. Kelimelerin vektörleştirilmesi için kullanılan ikinci yöntem ise FastText'tir. FastText kelimeleri *n-gram* parçalarına bölen ve bu şekilde her kelimeyi alt parçalardan (subword) öğrenerek metin temsili oluşturan bir dil modelidir (Dharma vd., 2022). FastText yöntemi ile metin içeriklerinde bulunan her bir kelime, büyüklüğü 300 olan ve sayısal değerlerden oluşan vektör dizilerine dönüştürülmüştür.

Ancak Word2Vec ve FastText modellerinin bağlamdan bağımsız çalışmaları ve her kelimeye sabit birer vektör atamaları nedeniyle metinlerdeki bağlamsal bilgiyi yeterince yansıtmadıkları görülmüştür. Bu durum, öznitelik vektörlerinin bağlamsal ilişkilerden yoksun olmasına ve dolayısıyla metinlerdeki anlam derinliğini yeterince yansıtamamalarına neden olmaktadır. Bu nedenle, öznitelik vektörlerinin bağlamsal bilgiye dayandırılması amacıyla BERT (Bidirectional Encoder Representations from Transformers) dil modeli denenmiştir. BERT, kelimelerin bağlam içindeki diğer kelimelerle olan ilişkilerini dikkate alarak dinamik ve bağlamsal vektörler üretmektedir

(Hoffman ve Schütze, 2019). BERT'in sunduğu bağlamsal temsiller ile, yalnızca kelimenin kendisini değil, kelimenin dokümandaki diğer unsurlarla olan ilişkisi de dikkate alınarak vektörleştirme işlemi gerçekleştirilmiştir. BERT'in bu avantajları, sınıflandırma performansında belirgin bir iyileşme gözlemlenmesine yol açmıştır. Bu nedenle, yapay sinir ağının eğitiminde BERT modeli kullanılarak oluşturulan öznitelik vektörlerinin kullanılmasına karar verilmiştir.

Yapay Sinir Ağının Eğitimi

Metin sınıflandırma görevi için bir sinir ağı yaklaşımı geliştirilmiştir. Sinir ağı sınıflandırıcıları, sinir hücreleri ağlarından oluşmaktadır. Ağda, girdi katmanı öznitelik terimlerini içerirken, çıktı katmanları önceden tanımlanmış hedef kategorileri temsil emektedir. Nöron birimleri arasındaki bağlantılar, atanan ağırlıklarla gösterilmektedir. Şekil 2'de gösterildiği üzere metin dokümanlarının öznitelik vektörleri, girdi katmanı nöronlarına atanmakta, bu birimlere aktivasyon fonksiyonu (sigmoid) uygulanarak ağ boyunca iletilmektedir. Çıktı katmanlarındaki değer, girdi olarak verilen metin dokümanlarının kategorisini belirlemektedir. Sinir ağını eğitmek için ileri besleme (feed forward) ve geri yayılım (back propagation) yöntemleri kullanılmaktadır.



Şekil 2. Kademeli yapay sinir ağı modeli

Yapay sinir ağı modelinin eğitiminde, sektör bağımsız olarak tüm organizasyonlarda depolanan ve işlenen adli sicil kaydı, kimlik, özgeçmiş, işe giriş sözleşmesi, sertifika gibi kişisel bilgilerin yer aldığı İnsan Kaynakları ve Mali İşler departmanlarında yer alan dokümanlar ile ihale ve tedarik süreçlerinde yer alan idari şartnameler ve teknik şartnamelerden oluşmak üzere 7 farklı kategoride etiketlenmiş D_0 veri setinden elde edilen öznitelik vektörleri kullanılmıştır. Geliştirilen modelin, hazırlanan veri setleri ile eğitilmesi sonucunda, organizasyonel yapıya özelleşmiş ve öğrenme kazanımlarının taşınmasını olanaklı kılan bir sınıflandırma yöntemi elde edilmiştir.

Yeni Doküman Setlerinin Entegrasyonu

Önceden tanımlanmış doküman kategorilerine göre etiketlenmiş eğitim setleri ile eğitilen yapay sinir ağının sınıflandırma performansı önceden tanımlanmış kategorilere girmeyen yeni doküman türleri ile karşılaştığında düşmektedir. Yapay sinir ağının yeni doküman yığınları ile karşılaştığında bunları sınıflandırabilmesi için yeni dokümanların tekrar kategorize edilmesi ve etiketlenmesi gerekmektedir. Yeni dokümanların el yordamıyla kategorize edilmesi zaman alan ve etkin olmayan bir yöntemdir. Bu probleme bir çözüm olarak sınıflandırıcı yapay sinir ağının kategorisi önceden belli olmayan ve etiketlenmemiş yeni doküman yığınları ile tekrar eğitiminde kullanılacak veri setinin elde edilmesi amacıyla yeni doküman setine (D_1) denetimsiz doküman kümeleme (unsupervised document clustering) yöntemi uygulanmıştır.

Denetimsiz Doküman Kümeleme Yöntemi

Gerçek dünya senaryolarında dokümanların el yordamıyla etiklenebilmesi için gerekli kategori ön bilgisi çoğunlukla mevcut değildir veya yetersizdir. Ayrıca, verilerin el yordamıyla etiketlenmesi genellikle maliyetlidir. Bu tip durumlarda kümeleme, denetimli öğrenme yaklaşımlarına göre daha uygun bir seçenek olarak öne çıkmaktadır (Bouguettaya vd., 2015).

Çalışmanın bu aşamasında D_1 doküman setindeki dokümanlar önceden belirlenen herhangi bir kategoriye el yordamıyla atanmamış, bunun yerine, dokümanlar arasındaki benzerlikler ile ilgili bilgilere dayalı kümeleme algoritmaları kullanılarak bir doküman kategorizasyonu elde edilmiştir. Hacimsel olarak büyük miktardaki veri yığınının tek bir adımda kümelenmesi hatalı sonuçlara yol açabileceği için D_1 doküman setinin kategorizasyonu için iki aşamalı bir kümeleme çalışması gerçekleştirilmiştir. İlk kümeleme çalışması üst bilgi (metadata), ikinci kümeleme ise içerik bilgisine göre yapılmıştır.

Üst bilgiye dayalı kategorizasyon için, verileri benzerliklerine göre gruplara ayıran *K-Means* kümeleme algoritması uygulanmıştır (Xiao ve Yu, 2012). Algoritmanın temel prensibi, verileri önceden belirlenen küme merkezine (centroid) yakınlığına göre sınıflandırmaktır (Singh ve Bhatia, 2011). Optimum küme sayısını belirlemek ve veri kümeleri içerisindeki tutarlılığı doğrulanmak amacıyla *silüet katsayısı* (silhouette coefficient) tekniği kullanılmıştır (Rousseeuw, 1987). 2 no'lu formül ile gösterilen silüet katsayısı, bir nesnenin kendi kümesine ne kadar iyi uyduğunu ve diğer kümelerden ne kadar farklı olduğunu ölçmektedir (Dinh vd., 2019; Subbalakshmi vd., 2015).

$$s(i) = a(i)b(i)/\max\{a(i),b(i)\} \quad (2)$$

İçerik bilgisine göre yapılan kümelemede ise hiyerarşik kümeleme (hierarchical clustering) algoritması kullanılmıştır. Hiyerarşik kümeleme algoritması, verilerin nasıl gruplandığını ve bu grupların hangi benzerlik seviyelerinde değiştiğini gösteren dendrogram adı verilen bir ağaç yapısı oluşturur. Kümeleme işlemi, başlangıçta en benzer verilerin birleştirilip daha büyük kümeler oluşturulmasıyla ilerler. Yani, en benzer öğeler bir araya getirilerek daha büyük gruplar oluşturulur ve bu süreç, tüm veriler tek bir büyük kümede toplanana kadar devam eder (Jain vd., 1999).

Hiyerarşik kümeleme, özellikle gerçek dünya verilerinin üstel büyümesi nedeniyle veri analitiğinde büyük öneme sahiptir. Etiketlenmemiş ve kategori bilgisi önceden bilinmeyen veri yığınlarının kategorizasyonunda hesaplama maliyetini düşüren etkin bir yöntemdir (Bouguettaya vd., 2015). Hiyerarşik kümeleme, bireysel noktalardan en üst düzey kümeye doğru artırarak (agglomerative) ya da en üst kümeyi atomik veri nesnelere bölerek (top-down) iki farklı şekilde uygulanabilir. Bu çalışmada, hesaplama maliyetinin azaltılması amacıyla ve hem departman içi dokümanların birbirine benzediği hem de süreç bazında kullanılan dokümanların birbirinden farklı olduğu durumlarda daha doğru yakınsamalar vermesi nedeniyle artırimsal yaklaşım kullanılmıştır.

İçerik bilgilerinin kümelenmesi sürecinde vektörleştirme için, veri analizi ve makine öğrenimi için yaygın olarak kullanılan Scikit-learn (sklearn) kütüphanesi tercih edilmiştir. Bu kütüphane, metin verilerinin sayısal temsillerini oluşturmak amacıyla çeşitli vektörleştirme yöntemleri sunmaktadır. İlk aşamada, içerik verileri vektörleştirilmiş ve bu veriler, kümeleme algoritmalarının uygulanmasına uygun bir formatta işlenmiştir.

Vektörleştirme işlemi sonrasında, elde edilen veriler kullanılarak kümeleme işlemi gerçekleştirilmiş ve kümeler arasındaki mesafeler hesaplanmıştır. Mesafe hesaplama işlemi, benzer içeriklerin gruplandırılmasını sağlayarak her bir kümenin içindeki elemanların birbirine ne kadar yakın olduğunu göstermektedir.

Bu aşamalar sonucunda oluşan farklı üst bilgi kümelerinde bulunan benzer dokümanların birleştirilebilmesi için dokümanlar merkez (centroid) değerlerine göre listelenmiş ve sıralama yapılarak yakınlıkları incelenmiştir. Doküman üst bilgilerine göre farklı kümelerde olan, ancak centroid değerleri birbirine yakın olan ve kullanıcı değerlendirmesine göre anlamlı bir yakınsaklığa sahip dokümanlar aynı kümede birleştirilmiştir.

İki aşamalı kümeleme ile elde edilen kategorizasyon çıktıları, kullanıcı değerlendirmesine (user review) sunularak kümelemenin doğru yapıldığı yapılmadığı ile ilgili geri bildirim alınmıştır. Kullanıcının anlamlı bulduğu kümeler modelin tekrar eğitilmesinde kullanılacak veri setine dahil edilmiş, anlamlı bulunmayan kümeler ise tekrar kümeleme işlemine tabi tutulmuştur.

D_1 doküman setine daha önce D_0 doküman setine uygulanan köklendirme (stemming) ve lemmatizasyon (lemmatization) işlemleri uygulanmıştır. D_1 doküman seti için de Bölüm 3.1.1'de bahsedildiği şekilde BERT dönüştürücüsü yardımıyla öznitelik vektörleri elde edilmiştir. Elde edilen yeni öznitelik vektörleri yeni doküman kümeleri için her özneliğin ilgili kategori ile ilişkilendirilmesini sağlamaktadır. D_1 doküman seti için elde edilen öznitelik vektörleri, D_0 doküman seti için elde edilen öznitelik vektör uzayına eklenmiş, böylece hem mevcut hem de yeni doküman kategorileri aynı özellik uzayında temsil edilmiştir.

Yapay Sinir Ağı Modelinin Yeniden Eğitimi

Yapay sinir ağı modeli, yeni gelen doküman yığınlarını da sınıflandırabilmesi için yukarıda bahsedilen artırimsal hiyerarşik kümeleme çalışması sonucu elde edilen yeni dokümanları temsil eden öznitelik vektörleri ve yeni doküman kategorileri ile yeniden eğitilmiştir.

RİSK SKORLAMASI

Risk skorlaması yapılırken siber güvenliğin temelini oluşturan gizlilik (confidentiality), bütünlük (integrity) ve erişebilirlik (availability) prensipleri temel alınmıştır. İlk olarak her bir doküman için her bir prensip ile ilgili skorlaması yapılmıştır. Dokümanın gizlilik risk skorunun belirlenebilmesi için dokümanın, daha önceden tanımlanan, *gramer* adı verdiğimiz TCKN, IBAN, sicil no, vb., belirli bir formata sahip kişisel veri yapılarını barındırıp barındırmadığı, RegEX desen tanımlama yöntemi ile keşfedilmiştir. Keşfedilen her bir *gramer*'e 1-5 skalasında bir gizlilik skoru verilmiştir. Her bir dokümana, içinde geçen *gramer*'lerin gizlilik skorlarına göre 1-5 arası gizlilik risk skoru verilmiştir. Tabiidir ki bir dokümanın içinde farklı gizlilik skoruna sahip *gramer*'ler bulunabilmektedir. Bu durumda, en yüksek gizlilik skoruna sahip *gramer*'in gizlilik skoru dokümanın gizlilik risk skoru olarak atanmıştır. Bütünlük ve erişebilirlik risk skorları dokümanın içeriği ile ilgili olmayıp, bağlandığı ve indekslendiği ortam ile ilgili olduğundan, her bir ortam analistin yargısına göre 1-5 skalasında skorlanarak, dokümanlara buldukları ortamın bütünlük ve erişebilirlik risk skorları atanmıştır.

Gizlilik, erişebilirlik ve bütünlük risk skorları ile risk gerçekleşme olasılıklarının çarpımı sonucu elde edilen en büyük değer dokümanın genel risk seviyesini oluşturmaktadır (3 no'lu formül). Her bir riske ilişkin gerçekleşme olasılığı tarihsel veri göz önünde bulundurularak güvenlik uzmanı tarafından belirlenmektedir.

$$Risk = Max (Risk(gizlilik, bütünlük, erişilebilirlik)) \times Olasılık (gizlilik, bütünlük, erişilebilirlik) \quad (3)$$

SONUÇLAR VE TARTIŞMA

Tekrar eğitilen yapay sinir ağı modelinin performans göstergeleri Tablo 1'de verilmiştir. Yapay sinir ağı modelinin dokümanları %60 doğrulukla sınıflandırdığı gözlenmiştir.

Tablo 1. Yapay sinir ağı kademeli eğitim sonuçları

Sınıflar	Kesinlik (Precision)	Duyarlılık (Recall)	F1 Skor	Destek Değeri
Adli Sicil Kaydı	0.71	1	0.83	17
Özgeçmiş	0.91	0.5	0.64	179
Kimlik	0.37	1	0.54	11
Sertifika	0.66	0.6	0.63	119
İşe Giriş Sözleşmesi	0.75	0.43	0.55	7
İdari Şartname	0.19	1	0.31	15
Teknik Şartname	0.40	0.55	0.46	11
Doğruluk (Accuracy)			0.60	359
Makro Ortalama	0.57	0.72	0.57	359
Ağırlıklı Ortalama	0.75	0.60	0.63	359

Kümeleme ve yapay sinir ağının eğitimine farklı kurumlardan elde edilen farklı kategorilerde ve içerikte doküman setleri ile devam edildikçe modelin sınıflandırma kapsamı ve performansının artacağı değerlendirilmektedir. Böylece model, farklı sektörlerde faaliyet gösteren kurum ve kuruluşların doküman sınıflandırma görevleri için daha yüksek bir sınıflandırma performansı ile kullanılabilir olacaktır.

Çalışmanın özgün değerini, özgün bir veri seti ile kurumsal dokümanların sınıflandırılması için bir sinir ağı eğitim çerçevesinin oluşturulması, denetimsiz hiyerarşik kümeleme yönteminin etiketsiz kurumsal dokümanların sınıflandırılmasında kullanılması, veri güvenliği prensipleri bağlamında bir risk skorlama metodolojisinin sunulması oluşturmaktadır.

Referanslar

- Bouguettaya A, Yu Q, Liu X, Zhou X, Song A, 2015. Efficient agglomerative hierarchical clustering. *Expert Systems with Applications*, 42(5): 2785-2797.
- Chen Z, Huang L, Murphey YL, 2007. Incremental learning for text document classification. *International Joint Conference on Neural Networks*. 12-17 August 2007, Page: 2592, Orlando, Florida, USA.
- Dharma EM, Gaol FL, Warnars HLHS, Soewito BENFANO, 2022. The accuracy comparison among word2vec, glove, and fasttext towards convolution neural network (cnn) text classification. *Journal of Theoretical and Applied Information Technology*, 100(2): 349-359.
- Dinh, DT, Fujinami T, Huynh VN, 2019. Estimating the optimal number of clusters in categorical data clustering by silhouette coefficient Knowledge and Systems Sciences: 20th International Symposium, KSS 2019. November 29 – December 1 2019, Page 1, Da Nang, Vietnam.
- Dzisevič R, Šešok D, 2019. Text classification using different feature extraction approaches. *Open Conference of Electrical, Electronic and Information Sciences (eStream)*, 25 April 2019, Page: 1, Vilnius, Lithuania.
- Hofmann V, Pierrehumbert JB, Schütze H, 2020. Dynamic contextualized word embeddings. *arXiv preprint arXiv:2010.12684*.
- Jain AK, Murty MN, Flynn PJ, 1999. Data clustering: a review. *ACM computing surveys (CSUR)*, 31(3): 264-323.
- Rousseeuw PJ, 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20: 53-65.
- Safari BA, 2017. Intangible privacy rights: How europe's gdpr will set a new global standard for personal data protection. *Seton Hall Law Review*, 47(3): 809-849.
- Siddiqui ZA, Park U, 2021. Progressive convolutional neural network for incremental learning. *Electronics*, 10(16): 1879.
- Singh RV, Bhatia MS, 2011. Data clustering with modified K-means algorithm. *International Conference on Recent Trends in Information Technology (ICRTIT)*. 3-5 June 2011, Page: 717, Chennai, India.
- Subbalakshmi C, Krishna GR, Rao SKM, Rao PV, 2015. A method to find optimum number of clusters based on fuzzy silhouette on dynamic data set. *Procedia Computer Science*, 46: 346-353.
- Wang XG, Lo SM, Zhang HP, 2013. Influence of feature extraction duration and step size on ANN based multisensor fire detection performance. *Procedia Engineering*, 52: 413-421.
- Xiao Y, Yu J, 2012. Partitive clustering (K-means family). *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 2(3): 209-225.
- Zhang T, Wang L, 2020. Research on text classification method based on word2vec and improved TFIDF. *4th International Conference on Intelligent, Interactive Systems and Applications (IISA2019)*, 26 – 30 December 2019, Page: 199, Mumbai, India

Teşekkür

Yazarlar, araştırmanın gerçekleştirilmesinde desteğini esirgemeyen Barikat İnternet Güvenliği Bilişim ve Ticaret A.Ş.'ye ve Ar-Ge faaliyetlerine destek olan Türkiye Bilimsel ve Teknolojik Araştırma Kurumu (TÜBİTAK)'na şükranlarını sunmayı borç bilir.

Covid-19 ve Deprem Süreçlerinde Uzaktan Eğitim: Öğrenci Beklentileri ve Deneyimlerinin Makine Öğrenimi ile Karşılaştırılması

Döndü Soytürk^{1}, Osman Keçeçi¹ ve Şahika Gokmen^{2,3}**

^{1,2} Ekonometri Bölümü, İktisadi ve İdari Bilimler Fakültesi, Hacı Bayram Veli Üniversitesi, Ankara, Türkiye.

³ Statistics Department, Uppsala University, Uppsala, Sweden.

Özet

Uzaktan eğitim, öğretici ve öğrenenin farklı mekânlarda olduğu, özel ders tasarımı ve teknolojilerini gerektiren planlanmış, açık erişimli bir öğrenme yöntemi olarak tanımlanmaktadır. Türkiye'de uzaktan eğitim, özellikle Covid-19 salgını sürecinde uygulanmaya başlanan bir eğitim yöntemi olarak karşımıza çıkmıştır. Covid-19 sürecinde üniversite öğrencileri yaklaşık 1,5 yıl boyunca uzaktan eğitim görmüş ve bu sürecin çıktıları literatürde birçok araştırmacı tarafından incelenmiştir. Bu çalışmaların ortak noktası; sistemin eksik ve olumlu yönlerinin ortaya konulmasıyla daha iyi bir eğitim ve öğretim sürecine katkı sağlayabileceğini vurgulamış olmalarıdır. Buna ek olarak, Covid-19 sürecinin ani başlangıcıyla birçok üniversitenin uzaktan eğitim süreçlerine önceden hazırlık yapma fırsatı bulamadığı da belirtilmiştir. Bu ani başlayan uzaktan eğitim süreci Haziran 2021'de sona ermiş olsa da, birçok üniversite uzaktan eğitim ile sertifika ve diploma programları açmaya devam etmiştir. Diğer yandan, Türkiye'de 6 Şubat 2023 tarihinde meydana gelen iki büyük deprem, pek çok ilde büyük hasar ve yıkıma neden olmuş, bu sebeple Yükseköğretim Kurulu (YÖK), 2022-2023 bahar döneminde üniversitelerin tekrar uzaktan eğitime geçmesi kararı almıştır. Bu aşamada üniversitelerin geçmiş deneyimlerine dayanarak uzaktan eğitim sürecine daha hızlı adapte olmaları beklenmiştir. Üniversiteler, bu süreçte uzaktan eğitim sistemlerini geliştirmek ve iyileştirmek adına çeşitli adımlar atmışlardır. Buna karşılık literatürde yapılan incelemelerde, 6 Şubat depremi sonrası uzaktan eğitim sistemini pandemi dönemindeki uzaktan eğitimle karşılaştıran bir çalışmaya rastlanmamıştır. Mevcut çalışmaların büyük bir kısmının tekil olarak belirli bölümleri veya öğrencileri hedef aldığı görülmektedir. Bu nedenle, bu çalışma ile Ankara Hacı Bayram Veli Üniversitesi İktisadi ve İdari Bilimler Fakültesi ile Ankara Üniversitesi Fen Fakültesi'nin ortak örnekleme üzerinden, öğrencilerin deprem sonrası uygulanan uzaktan eğitim sürecine yönelik algılarının ve deneyimlerinin Covid-19 uzaktan eğitim sürecine göre nasıl değiştiğinin incelenmesi amaçlanmıştır. Bu sayede öğrencilerin son süreçte daha gelişmiş olduğu varsayılan uzaktan eğitim sürecinden beklentilerinin ve bu eğitim tarzına yaklaşımlarının daha iyi anlaşılabilirliği düşünülmektedir. Bu doğrultuda saha araştırmasıyla gözlenerek düzenlenen veriler makine öğrenmesi araçlarından uygun karar ağacı algoritmasıyla incelenmiştir. Bu aşamada karar ağacı algoritmalarının tercih edilmesinin nedeni, öğrencilerin uzaktan eğitim sürecindeki deneyimlerini belirleyen temel faktörleri ortaya çıkarmak ve incelemektir. Çalışmada elde edilen en önemli bulgu, öğrencilerin deprem sonrası uygulanan uzaktan eğitimin Covid-19 dönemi uzaktan eğitim sürecine göre daha verimli yürütüldüğünü düşündüren en önemli faktörün dersteki beklentilerinin karşılanmasıdır. Elde edilen bulguların, eğitim kurumlarının ve ilgili idarelerin uzaktan eğitim süreçlerini iyileştirmesi ve öğrenci deneyimini geliştirmesi adına yararlı veriler sunabileceği düşünülmektedir.

Anahtar Kelimeler: Uzaktan Eğitim, Makine Öğrenmesi, Karar Ağacı, Random Forest

DİPNOT: Söz konusu çalışma adı geçen yazarlar tarafından yürütülen TÜBİTAK 2209- A projesinden üretilmiştir.

1.GİRİŞ

Uzaktan eğitim fiziksel olarak öğrencilerin buldukları yerlerde olmasını gerektirmeksizin, öğrenci ve öğretmenlerin güncel iletişim teknolojilerinden yararlanılarak bir internet ortamında eşzamanlı ya da eşzamanlı olmadan değişik şekillerde ders işledikleri, planlı bir öğretim şeklidir. Bu yöntemle öğretmenler, öğrenciler ve bu iki tarafı birbirine bağlayan köprü vasıtası gören merkez tarafından etkileşim sağlanmaktadır (Bakıoğlu ve Can, 2014) Uzaktan eğitim, belirli merkezlerden yürütülen, bireyin kendi kendine öğrenmesini amaçlayan, eğitsel içeriklerin öğrenenlere özel olarak hazırlanmış araçlarla ve çeşitli ortamlar yoluyla sağlandığı bir öğretim yöntemidir (Banar ve Fırat, 2015). Uzaktan eğitimde, her bir öğrenci kendi hızında öğrenir, bilgiye hızlı ve kolay bir şekilde erişme imkânı verir ve yüz yüze eğitimde gereken ek harcamaları ortadan kaldırır (Andsoy vd., 2012, Demir, 2014, Çiftçi, 2015) Uzaktan eğitimle daha geniş kitlelere, daha ekonomik ve kısa zamanda ulaşmak mümkün olabilmektedir (Banks vd., 2007). Geleneksel eğitimden yararlanmayan kişiler için, vazgeçilmez bir eğitim sistemidir. Uzaktan eğitim, bireylere kendi kendilerine öğrenme imkânının sağlandığı, geleneksel eğitime göre daha esnek ve birey koşullarına uyarlanabilir bir eğitimidir. Uzaktan eğitimle eğitim hizmeti götürmekteki sınırlılıkların kısmen ya da tümüyle ortadan kaldırılarak eğitim imkânlarının daha geniş kitlelere ulaştırılması amaçlanmaktadır (Uluğ ve Kaya, 1997).

Dünya'daki birçok örgün eğitim kurumu içerisinde, uzaktan eğitim kurumları yer almaktadır. Uzaktan eğitim yapan bu kurumlar, öğretme etkinliğinden çok öğrenmenin yapısalcı yönünü daha etkin olarak ortaya çıkarmışlar ve uzaktan eğitim teknoloji alt yapısının gelişmesi sayesinde öğrenme sürecine katkıda bulunmuşlardır (Çıgıllık ve Bayrak, 2015). Tarihi 1800'lü yıllara kadar uzanan uzaktan eğitim Türkiye'de Anadolu Üniversitesi'nde Açık Öğretim Fakültesi'nin kurulması ve 1982-83 eğitim-öğretim yılında öğrenci alımına başlaması ile yükseköğretimde yerini almıştır (Akdemir, 2011). 1990'larda internetin doğuşuyla uzaktan eğitim yeni bir boyut kazanmıştır. Özellikle 2000'lerden itibaren ise çevrimiçi eğitim platformlarının gelişimiyle birlikte bu eğitim türü büyük bir ivme kazanmıştır. Son on yılda Türkiye'de Anadolu Üniversitesi dışında farklı üniversitelerin açık ve uzaktan öğrenme alanında yatırımlar yapması ve programlar açmasıyla birlikte açık ve uzaktan öğrenmenin yaygınlaşması hızlanmış, yeni teknolojilerden yararlanılmaya başlanmıştır (Turkan ve Cihan, 2015). Uzaktan eğitim yalnızca üniversite eğitimi değil, meslek sahibi bireylerin de uzaktan eğitim programlarına dahil edildiği görülmektedir. Deneyimsiz öğretmenlerin mesleklerinin başlangıcında, deneyimli öğretmenlerin ise kendilerini geliştirmek için sürekli eğitimi amacıyla kullanılabilirlerdir. 2020 yılında COVID-19 pandemisinin başlaması, dünya çapında uzaktan eğitimin hızla yayılmasına yol açmıştır. Salgının kontrol altına alınması amacıyla, okullar ve üniversiteler kapatılmış ve eğitim, zorunlu olarak uzaktan yapılmaya başlanmıştır. Türkiye'de de Mart 2020 itibarıyla yüz yüze eğitime ara verilerek uzaktan eğitim sistemine geçilmiş, Millî Eğitim Bakanlığı (MEB) ve Yükseköğretim Kurulu (YÖK) tarafından hızlıca dijital platformlar oluşturulmuştur. Bu uzaktan eğitim uygulamaları sayesinde birbirlerinden kilometrelerce uzaklıkta farklı ortamlarda olan öğretmen ve öğrenciler, birbirleriyle görüntülü ve sesli olarak etkili iletişim kurabilmektedirler (İşman, 2011). Bu dönemde uygulanan uzaktan eğitimin öğrenci ve öğretmenler açısından avantajları ve dezavantajları olmuştur. Erişim ve esneklik, zaman yönetimi ve teknoloji kullanımı, bu eğitim modelinin sunduğu başlıca avantajlar arasındayken; erişim kısıtlamaları, etkileşim yetersizliği, motivasyon eksiklikleri ve disiplin sorunları ise karşılaşılan başlıca dezavantajlar olarak öne çıkmıştır.

Literatürde bu dönemi inceleyen ve uzaktan eğitimin olumlu ve olumsuz yanları araştıran çalışmalar mevcuttur. Bu çalışmalar arasında özellikle Erzen ve Ceylan (2020), Kırmızıgül (2020) ve Çetin ve Akduman'ın (2022) çalışmaları örnek gösterilebilir. Bu çalışmalar incelendiğinde; Erzen ve Ceylan'ın (2020) COVID-19 salgını ile birlikte gelen uzaktan eğitim sürecinin uygulanmasındaki sorunlara odaklandığı; Kırmızıgül'ün (2020) ise COVID-19 salgınının getirdiği eğitim sürecinin hem olumlu hem de olumsuz yanlarına odaklandığı görülmüştür. Çetin ve Akduman (2022) ise COVID-19 salgını sonrası uygulanan uzaktan eğitimde öğretmen ve öğrenci görüşlerini toplayarak değerlendirmiş ve literatüre farklılık katmıştır. Diğer yandan bu çalışmaların uygulamaları incelendiğinde temel istatistiksel analizlerden öteye gitmediği görülmüştür.

COVID-19 pandemisi sürecinde hızla deneyimlenen uzaktan eğitim sistemi, 6 Şubat 2023'te yaşanan deprem felaketi sonrasında yeniden uygulamaya alınmıştır. Okullardaki yapısal zararlar ve olası riskler nedeniyle, Millî Eğitim Bakanlığı (MEB) ve Yükseköğretim Kurulu (YÖK) tarafından pandemi döneminde geliştirilen dijital eğitim platformlarının kullanılmasına karar verilmiştir. Bu süreçte, pandemi edinilen deneyimlerin ve yapılan altyapı iyileştirmelerinin uzaktan eğitimin daha verimli ve etkili işlenmesine katkı sunduğu düşünülmektedir. Öğrenci ve öğretmenlerin uzaktan eğitime aşinalıklarının da bu sürecin daha sorunsuz ilerlemesine yardımcı olabileceği düşünülmektedir. Literatürde bu kapsamda incelendiğinde, 6 Şubat deprem felaketi sonrasında iki üniversitede uygulanan uzaktan eğitim sistemlerini karşılaştıran kapsamlı bir araştırmanın henüz yapılmadığı görülmektedir. Bununla birlikte bu alanda benzer konularda gerçekleştirilmiş bazı araştırmalar bulunmaktadır. Fıratlı (2023), 6 Şubat 2023 depremleri sonrasında zorunlu uzaktan eğitime geçiş nedenlerine göre üniversite öğrencilerinin uzaktan eğitim algılarının farklılaşmasını incelemiştir. Özkaya vd. (2023) ise çevrimiçi eğitim uygulamasının İktisadi ve İdari Bilimler Fakültesi (İİBF) öğretim üyeleri üzerindeki etkilerini ele almıştır. Mete vd. (2024), meslek yüksekokullarındaki uzaktan eğitimi, öğrenci ve öğretim elemanları açısından değerlendirmiştir. Bu çalışmaların ortak noktası, belirli bir fakültenin öğrencileri veya öğretim üyeleri açısından uzaktan eğitim sistemine yönelik bakış açısının incelenmesidir. Mevcut araştırmalarda, öğrencilerin eğitim alanlarındaki farklılıklara yeterince önem verilmediği ve çeşitli okulların sistemlerinin karşılaştırılmasına olanak tanımadığı gözlemlenmektedir. Bununla birlikte Keçeci vd.'nin (2024) çalışması 6 Şubat Deprem Felaketi Ardından Uygulanan Uzaktan Eğitim Sistemine Bakışın İncelenmesi ve Değerlendirilmesi: Ankara Hacı Bayram Veli Üniversitesi İktisadi ve İdari Bilimler Fakültesi ve Ankara Üniversitesi Fen Fakültesi Karşılaştırması şeklinde ele alarak bu konuda literatüre farklılık katmıştır. Söz konusu çalışma, bu çalışmanın devamı niteliğinde olup, makine öğrenmesi sınıflanmasına odaklanmaktadır. Bu bağlamda, bu araştırmanın amacı, deprem felaketi sonrası uzaktan eğitim süreçlerini, farklı disiplinlere odaklanan iki fakülte üzerinden incelemektir. Özellikle, yakın konumda yer alan Fen Fakültesi'ndeki uygulamalı derslerle Sosyal Bilimler Fakültesi'ndeki teorik dersler arasındaki farklılıklar ve benzerlikler, bu çalışmanın ana odağını oluşturmaktadır. Uygulamalı derslerin daha fazla etkileşim ve laboratuvar uygulamaları içermesi, Fen Fakülteleri'nde yürütülen uzaktan eğitim süreçlerini Sosyal Bilimler Fakülteleri'ne kıyasla daha farklı hale getirmektedir. Bunun

yanı sıra, Sosyal Bilimler Fakültesi'nde verilen teorik derslerin uzaktan eğitim platformlarında nasıl sunulduğu ve bu derslerin öğrenci katılımı üzerindeki etkileri de dikkate alınması gereken önemli unsurlardır. Bu amaçla, Ankara'daki iki devlet üniversitesinden anket yoluyla toplanan veriler, veri görselleştirme yöntemleriyle analiz edilmiştir. Bu sayede, her iki fakültede uygulanan uzaktan eğitim sistemlerinin güçlü ve zayıf yönleri belirlenmeye çalışılmıştır. Elde edilen bulguların, gelecekteki eğitim politikaları ve uzaktan eğitim uygulamalarının geliştirilmesine önemli katkılarda bulunması beklenmektedir. Bu bağlamda, Ankara'daki iki devlet üniversitesinden anket yöntemiyle toplanan veriler, makine öğrenmesi yöntemleri kullanılarak analiz edilip karar ağaçları ile karşılaştırmalar yapılmış ve her iki fakültede yürütülen uzaktan eğitim sistemlerinin güçlü ve zayıf yönleri belirlenmeye çalışılmıştır. Elde edilen bulguların, gelecekteki eğitim politikalarının ve uzaktan eğitim uygulamalarının geliştirilmesine önemli katkılar sağlaması beklenmektedir.

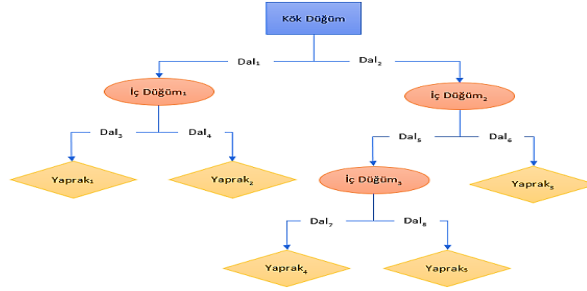
Çalışma dört ana bölümden oluşmaktadır. İlk bölümde Giriş kısmı sunulmakta, ikinci bölümde Makine Öğrenmesi, Karar Ağaçları ve veri toplama süreçleri ele alınmaktadır. Üçüncü bölümde bulgulara yer verilirken, son bölümde ise tartışma ve öneriler paylaşılmaktadır.

2. MAKİNE ÖĞRENMESİ ile KARAR AĞAÇLARI

Makine öğrenmesi, bir bilgisayarın belirli bir görevi veya problemi insan müdahalesi olmadan, veri üzerinden öğrenerek çözmeye çalışan bir yapay zekâ alt dalıdır (Mackenzie, 2015). Makine öğrenmesi, örnek verileri veya deneyimleri kullanarak bir performans kriterini en iyi ve verimli şekilde kullanmak için algoritmalarından yararlanmaktadır (Alpaydin, 2020) Günümüzde çok sayıda büyük verinin bulunması, bu verilerin çözümlenmesinde bilgisayar teknolojilerine ihtiyaç duyulmasını arttırmıştır. Bu noktada makine öğrenmesi üzerine yapılan araştırmalar da artmıştır. Bu artışla birlikte hesaplamalı öğrenme teorisi, yapay sinir ağları, istatistik ve örüntü tanıma gibi araştırma alanları arasında bağlantı kurularak bu alanlarla birlikte çalışmaya başlanmıştır. Böylece makine öğrenmesi teknikleri yüz tanıma gibi daha geleneksel sorunların yanı sıra veri tabanlarında bilgi keşfi, dil işleme ve robot kontrolü gibi problemlere uygulanmaya başlanmıştır (Dietterich 1997; Çoban, 2016). Makine öğrenmesi algoritmaları, geçmiş verilere dayanarak öğrencilerin gelecekteki derslerde ve konulardaki başarılarını tahmin etme yeteneğine sahiptir. Bu tahminler, öğrencilere destek sağlamak veya erken müdahaleler yapmak için kullanılabilir. Böylece, öğrencilerin başarı oranlarını artırmak ve eğitim kalitesini yükseltmek amacıyla eğitim bilimleri alanında makine öğrenmesi tekniklerinden faydalanmak mümkün hale gelir. Verinin türüne göre uygulanması gereken çeşitli makine öğrenmesi algoritmaları bulunmaktadır. Algoritmalar özelliğine göre kümeleme, sınıflandırma, tahmin ve kestirim yapabilme özelliklerine sahiptirler. Yapılandırılmış veriler için genellikle denetimli öğrenme algoritmaları (örneğin, regresyon, karar ağaçları veya destek vektör makineleri) tercih edilirken, yapılandırılmamış veri türleri (metin, görüntü, ses gibi) için derin öğrenme ve doğal dil işleme yöntemleri daha uygun olabilir. Özetle, verinin yapısına, formatına ve içeriğine bağlı olarak uygun makine öğrenmesi algoritmaları seçilerek analizlerde en doğru sonuçlara ulaşılması hedeflenir. En yaygın makine öğrenmesi algoritmaları k-en yakın komşu, Bayes sınıflandırıcısı, karar ağaçları, lojistik regresyon, destek vektör makineleri ve yapay sinir ağlarıdır. Literatürde birçok farklı alanda çeşitli makine öğrenmesi algoritmaları kullanılmaktadır. Bu bölümde, bu çalışmada kullanılan makine öğrenmesi algoritması Karar Ağacı açıklanmasına odaklanılmıştır.

Makine öğrenmesi yöntemleri içerisindeki sınıflama tekniklerinden biri de karar ağacıdır. Çeşitli seçenekler ve bunların arasındaki ilişkiyi gösteren grafiksel bir yapıdır. Şekil 1'de genel yapısı gösterilen karar ağacı dal, düğüm (kök düğüm ve iç düğüm) ve yapraklardan oluşmaktadır. Karar ağacının yapısındaki düğümler nitelikleri temsil eder. Dallar ve yapraklar ise ağacın birer elemanlarıdır. Kök düğüm, ağaç yapısının en üst seviyesinde yer alır ve bir ağaç yapısında yalnızca bir kök düğüm vardır. Ağacın en altındaki elemanlar yapraklardır. Kök düğüm ve yapraklar arasında kalan elemanlar ise dal olarak isimlendirilir (Alpaydin, 2000). Bu kök düğüm, kök düğümün ayrı sınıflarını veya düğümün ölçeği boyunca belirli aralıkları temsil eden çok sayıda dalda iki bölüme ayrılır. Her bölümdede, bölünen değişkenin sınıfları veya aralığı bakımından yanıtı olan bir soru sorulmaktadır. Bu soru örneğin, "erkek mi kadın mı?" olabilir. Bunun gibi sorular, ikiye bölünmüş karar ağacı oluşturmak için kullanılır. Karar ağaçları birden çok bölme ile de oluşturulabilir. Her bir bölünmede sorulan sorular, sonuçta ortaya çıkan vakaların bölünmelerde ne kadar uniform olması gerektiğini yansıtan bazı belirsizlik ölçüleri açısından tanımlanır. Her dal, diğer değişkenlerin

sınıfları veya aralıkları kullanılarak daha da bölünür. Her bölüntüde bölünen düğüme ana düğüm, bölünmüş olduğu düğümlere de alt düğüm adı verilir. Bu işlem, kesme kuralı gerçekleşinceye kadar devam eder (Nisbet vd., 2009:241).



Şekil 1: Karar ağacı yapısı

Kaynak: Albayrak, Y. D. D. A. S., & Yılmaz, Ö. G. Ş. K. (2009). Veri madenciliği: Karar ağacı algoritmaları ve İMKB verileri üzerine bir uygulama. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 14(1), 31-52.

Karar ağaçlarında her dalın belirli bir gerçekleşme olasılığı bulunmaktadır. Bu nedenle, düğümlerden dallara veya dallardan düğümlere ulaşma olasılıkları hesaplanabilir. Karar ağaçlarında kurallar (IF-THEN rules), kökten yaprağa doğru ilerleyerek oluşturulmaktadır. Sınıflandırma işlemi için kullanılan karar ağaçları, veri setinin eğitim ve test olarak ikiye ayrılmasını gerektirir. Eğitim verileri, öğrenme sürecinde kuralların oluşturulması için kullanılırken, test verileri bu kurallar doğrultusunda sınıflandırılır ve kuralların başarısı değerlendirilir. Başarılı sonuçlar elde edilmesi durumunda kurallar, yeni verilerin sınıflandırılması amacıyla kullanılabilir (Sarmanova ve Albayrak, 2013). Karar ağaçları ile yapılan sınıflandırmada eğitim verileri tek sınıf içeren alt kümeler bölünür. Bu işlemin sonucunda karmaşık ve büyük bir ağaç yapısı elde edilir. Bu durumda ağaç yapısının gereksiz karmaşıklığını azaltmak için ‘budama’ yapılır. Budama, oluşturulan ağaç yapısındaki alt kümeler yaprak atanması ile ağaç yapısının daha küçük bir yapıya indirgenmesidir (Cheng vd. 1988; Friedl ve Brodley, 1997). Karar ağacı hem karar vermeyi hem de sınıflama sürecinde öne çıkan faktörlerin belirlenmesini sağlayan bir tür sınıflama modelidir. Sınıflandırma analizi için geliştirilmiş pek çok karar ağacı algoritması bulunmaktadır. Bölümün devamında çalışma boyunca kullanılan karar ağacı algoritması olarak seçilen Random Tree Algoritması açıklanmıştır.

Karar ağacı kendi başına bir uygulama olmakla birlikte makine öğrenmesi algoritmalarının içerisinde de farklı algoritmalarla da birlikte çalışabilir. Biz bu çalışma için en sık kullanılan “Random Forest” sınıflama algoritmasını tercih ettik. Random Forest sınıflama algoritma adımlarını kısaca aşağıdaki gibidir:

Adım 1: Eğitim setinden rastgele K veri noktalarını seçin.

Adım 2: Seçilen veri noktalarıyla (Alt Kümeler) ilişkili karar ağaçlarını oluşturun.

Adım 3: Oluşturmak istediğiniz karar ağaçları için N sayısını seçin.

Adım 4: 1. ve 2. Adımları tekrarlayın.

Bu adımlar tekrar ettikten sonra bu sınıfların belirleyici olan ağaç ve dallara otomatik olarak bir tarafından karar ağacı diyagramı da elde edilmiş olmaktadır.

Karar ağaçları içerisinde Random Tree algoritması, her düğümden belirli sayıda rastgele seçilmiş özellikleri kullanarak ağaç yapısını oluşturmaktadır (Breiman, 2001). Bu algorithmada budama işlemi uygulanmamakta ve sınıf olasılıklarını tahmin etmek için veri setine dayanan bir seçenek bulunmaktadır. Random Forest algoritmasında birden fazla karar ağacı oluşturulmakta ve bu ağaçlar bir araya getirilerek karar ormanları ortaya çıkmaktadır. Bu bakımdan Random Forest algoritması ile tahmin üzerindeki her bir özelliğin diğerlerine göre önemini ölçmek oldukça kolaydır (Ballı, 2021).

3. VERİ TOPLAMA ve BULGULAR

Araştırmada veri toplama aracı olarak anketler kullanılmıştır. Çalışma, Ankara Hacı Bayram Veli Üniversitesi İktisadi ve İdari Bilimler Fakültesi ile Ankara Üniversitesi Fen Fakültesi’nde 2023-2024 bahar dönemini takip eden lisans öğrencilerinden oluşan bir örneklem üzerinde gerçekleştirilmiştir. Sosyal Bilimler Fakültesi’nde yaklaşık 2100, Fen Fakültesi’nde ise yaklaşık 2200 öğrenci bulunmaktadır. Bu öğrenciler çalışmanın kitlesini oluşturmaktadır. %95 güven düzeyinde belirlenen orantılı örneklem hacimlerinin üzerinde bir katılımla, Sosyal Bilimler Fakültesinden 512 ve Fen Bilimleri Fakültesi’nden 442 öğrenci olmak üzere toplam 954 kişilik bir katılım sağlanmıştır. Anket çalışması, yüz yüze saha görüşmeleri ve Google Anketler üzerinden yürütülmüştür. Anket toplamda 24 sorudan oluşmakta olup,

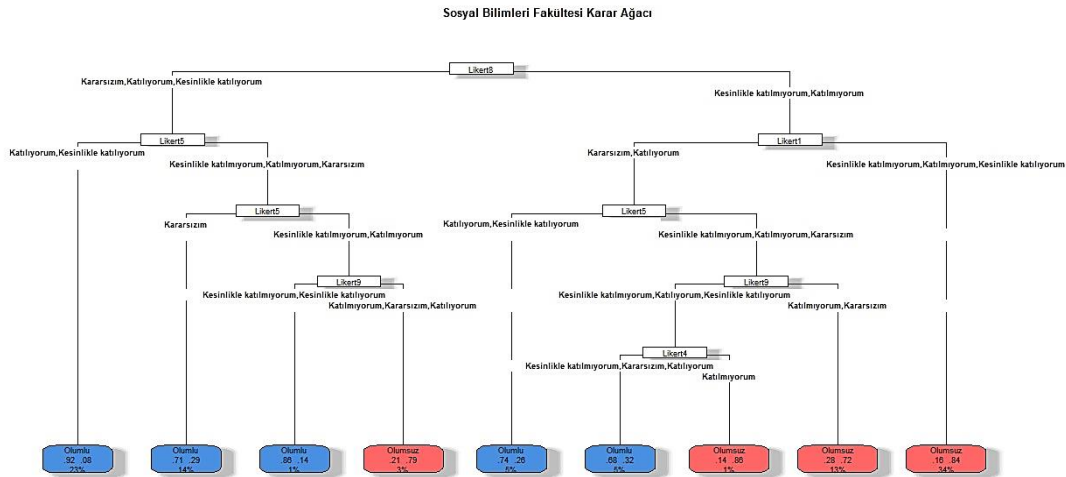
bu soruların on tanesi Likert ölçeğinde, altı tanesi ise evet-hayır formatındadır. Likert soruları, öğrencilerin uzaktan eğitim deneyimlerine yönelik genel kalite ve memnuniyet değerlendirmesini amaçlamaktadır. Ayrıca, COVID-19 pandemisinde uygulanan uzaktan eğitimle 6 Şubat deprem felaketi sonrası uygulanan uzaktan eğitimi karşılaştırmalarına yönelik bir soru da yer almaktadır. Çalışmanın ana konusunu oluşturan sınıflayıcı değişken ise, öğrencilerin üniversitenin 6 Şubat depremi sonrası sunduğu uzaktan eğitimin yeterliliği hakkındaki değerlendirmelerinden elde edilmiştir. Bu kapsamda, öğrencilere uzaktan eğitimi 0 ile 10 arasında bir puanla değerlendirmeleri istenmiş; ardından bu puanlar, iki ana kategoriye ayrılarak sınıflayıcı değişken oluşturulmuştur. Böylelikle, öğrencilerin verdikleri puanlar doğrultusunda, uzaktan eğitimin "daha iyi" veya "daha kötü" olduğu şeklinde iki kategorili bir sınıflandırma yapılmıştır.

Bu değişkenin sınıflayıcı ve belirleyici özelliklerini ortaya çıkarmak amacıyla, 10 adet Likert sorusu kullanılmıştır. Bu sorular, öğrencilerin değerlendirmelerini ölçerek modelin önemli niteliklerini daha net şekilde belirlemeye katkı sağlamaktadır. Söz konusu sorular Tablo 1’de verilmektedir.

Tablo 1: Öğrenci Memnuniyet Anketinde Yer Alan Likert Sorular.

LİKERT1	Öğrenci Bilgi Sistemi Öğrenci İhtiyaçlarını Karşılıdı.
LİKERT2	Gerekli Bilgilere(Sınav Tarihleri, Ders Kayıtları, Ödevler Vb.) Kolaylıkla Erişebildim.
LİKERT3	Ders İçeriklerini Daha Etkili Öğrenebildim.
LİKERT4	Ders Veren Öğretim Üyeleri Öğretim Sürecini Yönlendirme Ve Rehberlik Etme Konusunda Başarılılardı.
LİKERT5	Öğretim Materyalleri Yeterli, Güncel Ve Bilgi Düzeyi Açısından Uygundu.
LİKERT6	Dersin Yönetimi İçin Kullanılan "Canlı Ders Sistemi"(ZOOM Vb.) Öğrenci İhtiyaçlarını Karşılıdı.
LİKERT7	Dersi Veren Öğretim Üyeleri Canlı Ders Konusunda Hazırlıklı Ve Donanımlılardı.
LİKERT8	Aldığım Eğitim Beklentimi Karşılıdı.
LİKERT9	Öğrenci Başarısının Ölçüm Ve Değerlendirmesi Sağlıklı Şekilde Yapıldı.
LİKERT10	Teknolojik Altyapı Güvenilir Ve Sorunsuz Çalıştı.

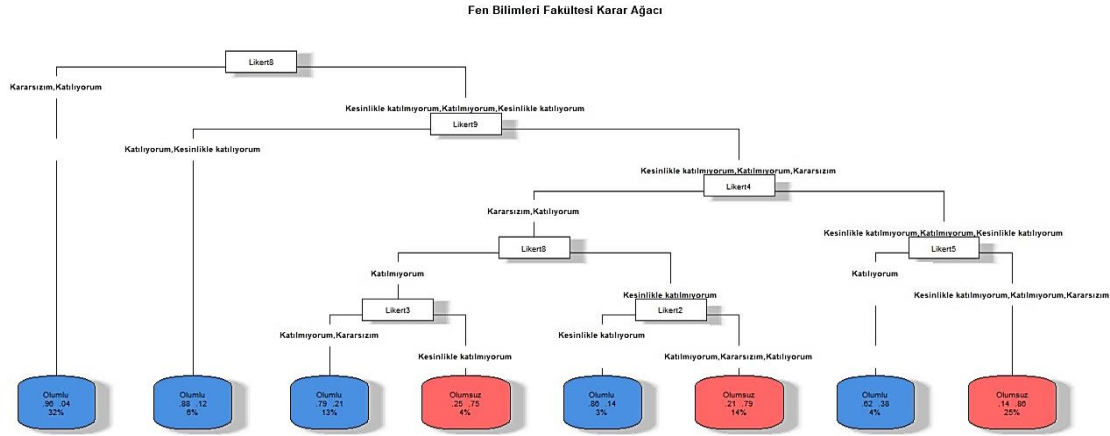
Veri seti, modelin performansını değerlendirmek ve farklı veri kümeleri karşısındaki doğruluğunu artırmak amacıyla %80 eğitim ve %20 test oranında ikiye bölünmüştür. Bu bölme işlemi, modelin eğitim sürecinde örneklerin çoğunu görebilmesini sağlarken, kalan %20’lik test verisi üzerinden modelin doğruluğu ve güvenilirliği ölçülmüştür. Bu yöntem, modelin yeni ve görülmemiş verilere karşı ne kadar başarılı olduğunu değerlendirmede önemli bir rol oynamaktadır.



Şekil 3: Sosyal Bilimler Fakültesi Karar Ağacı

Şekil 3’de Sosyal Bilimler Fakültesi öğrencilerinin ders değerlendirmelerine yönelik oluşturulmuş bir karar ağacı yer almaktadır. Random Forest yöntemi ile oluşturulan bu ağaç, öğrencilerin derslere yönelik memnuniyet düzeylerini ve çeşitli kriterlere göre değerlendirmelerini yansıtmaktadır. Ağaç dallarında, öğrencilerin belirli ifadelerle katılım dereceleri Likert ölçeği kullanılarak ayrıntılı bir şekilde incelenmiştir. Sosyal Bilimler Fakültesi’nde en önemli faktörler olarak öne çıkan değişkenler; beklentilerin karşılanıp karşılanmadığı (Likert8), kaynakların güncelliği ve tatmin ediciliği (Likert5), ve Öğrenci Bilgi Sistemi’nin işlevselliği (Likert1) olmuştur. Bu üç faktör birlikte değerlendirildiğinde, öğrencilerin çevrimiçi eğitimden tatmin olup olmalarının bilgi sisteminin işlevselliğine, öğretim materyallerinin kalitesine ve eğitim beklentilerinin karşılanmasına bağlı olduğu görülmektedir. İyi çalışan bir bilgi sistemi, kaliteli öğretim materyalleri ve öğrencilerin beklentilerini karşılayan bir eğitim deneyimi, çevrimiçi eğitimde yüksek memnuniyet sağlamaktadır. Ancak, bu faktörlerden herhangi birinde aksaklık yaşanması, öğrencilerin genel eğitim deneyimini olumsuz etkileyebilir. Örneğin, bilgi sistemi yeterli olsa bile materyaller yetersiz veya güncel

değilse, öğrenciler eğitimden istedikleri verimi alamamaktadır. Benzer şekilde, materyaller yeterli olsa bile bilgi sistemine erişimde yaşanan aksaklıklar, öğrencilerin öğrenme sürecini sekteye uğratabilmektedir. Sosyal Bilimler Fakültesi özelinde yapılan bu analiz, çevrimiçi eğitimin öğrencilerin beklentilerini karşılama düzeyini ortaya koymuştur. Özellikle sosyal bilimler öğrencisi sıklıkla kaynak ve okuma üzerinden çalışmalarını sürdürdüğü için öğretim materyallerinin güncelliği ve yeterliliği ile Öğrenci Bilgi Sistemi'nin işlevselliği, öğrenci memnuniyeti üzerinde önemli bir etkiye sahiptir. Bu nedenle, çevrimiçi eğitimin etkinliğini artırmak için bilgi sisteminin işlevsel hale getirilmesi ve materyallerin güncel ve yeterli olmasının sağlanması önem arz etmektedir. Yapılacak bu iyileştirmeler, öğrenci memnuniyetini artırarak, Sosyal Bilimler Fakültesi'nde uzaktan eğitimin kalitesini yükseltecektir.



Şekil 4: Fen Bilimler Fakültesi Karar Ağacı

Fen Bilimleri Fakültesi'nde uygulamalı derslerin ağırlıklı olması, çevrimiçi eğitim sürecinde öğrenciler için özel zorluklar yaratmaktadır. Uygulamalı dersler, teorik derslere kıyasla öğrencilerin sadece bilgi edinmekle kalmayıp, aynı zamanda deneyim odaklı bir öğrenim sürecine ihtiyaç duydukları ve uygulamalı becerilerin geliştirilmesinin önemli olduğu bir yapıya sahiptir. Ancak çevrimiçi eğitim ortamında bu ihtiyaçların karşılanması zorlu hale gelmektedir. Laboratuvar çalışmaları, ekipman kullanımı ve deney yapma gibi etkinliklerin fiziksel olarak gerçekleştirilememesi veya yeterince etkili bir şekilde yapılamaması, bu zorlukların başında gelmektedir. Bu durum, ders notlamaları ve öğretim üyeleriyle olan rehberlik beklentilerini de farklı bir perspektiften ele almayı gerektirmektedir. Uygulamalı derslerde öğrencilerin becerilerini gözlemlemek ve değerlendirmek, fiziksel etkileşime dayandığı için çevrimiçi ortamda bu süreç daha karmaşık bir hal alabilir. Ders notlamasında yaşanan bu zorluklar, öğrencilerin adil ya da verimli bir değerlendirme aldıkları hissini zedeleyebilir. Çevrimiçi ortamda yalnızca teorik bilgiler üzerinden yapılan değerlendirmeler, uygulamalı yetkinliklerin ölçülmesi yerine daha yüzeysel bir performans değerlendirmesine yol açabilir. Bu nedenle, uygulamalı derslerde öğrencilerin başarılarını daha etkili bir şekilde değerlendirebilmek için esnek ve yenilikçi değerlendirme yöntemlerinin geliştirilmesi oldukça önemlidir. Öğretim üyelerinin rehberliği de uygulamalı derslerde kritik bir öneme sahiptir. Öğrencilerin öğretim üyelerinin rehberliğine yönelik beklentileri, teorik derslere göre daha yüksektir. Laboratuvarlarda veya uygulama alanlarında öğretim üyelerinin, öğrencilerle birebir etkileşim kurarak anlık geri bildirimde bulunmaları, öğrencilerin becerilerini geliştirmeleri için hayati bir rol oynamaktadır. Ancak çevrimiçi ortamda bu etkileşimlerin sınırlı kalması, öğrencilere rehberlik sağlama sürecini zorlaştırmakta ve bu durum, öğrencilerin uygulamalı derslerde öğretim üyelerinden yeterli rehberlik alamadıkları algısına yol açabilmektedir. Bu zorlukların üstesinden gelebilmek için, çevrimiçi platformlar üzerinden daha fazla etkileşim sağlayacak canlı oturumların düzenlenmesi, öğrencilere özel geri bildirimlerin sunulması ve rehberliğin daha bireysel ve yoğun bir şekilde yapılması önem taşımaktadır.

Sonuç olarak, Fen Bilimleri Fakültesi'nin çevrimiçi uygulamalı derslerinde öğrencilerin değerlendirme ve rehberlik beklentilerini karşılamak, çevrimiçi eğitim sürecinin en önemli zorluklarından biri olarak öne çıkmaktadır. Bu derslerde, öğrenci memnuniyetini ve öğrenme verimliliğini artırmak amacıyla öğretim üyelerinin rehberlik süreçlerine daha fazla zaman ayırmaları ve uygulamalı becerileri ölçmeye yönelik yenilikçi değerlendirme yöntemlerini benimsemeleri, çevrimiçi eğitimdeki kaliteyi artıracaktır.

Tablo 2: Sosyal Bilimler Fakültesi Ve Fen Bilimleri Karar Ağacı Metriği

Sosyal Bilimler Fakültesi		Fen Bilimleri Fakültesi	
Metrik	Modelden Hesaplanan Değer	Metrik	Modelden Hesaplanan Değer
Doğruluk	0.8105	Doğruluk	0.8597
%95 Güvenlik Aralığı	(0.7739, 0.8436)	%95 Güvenlik Aralığı	(0.8238, 0.8907)
No Information Rate (NIR)	0.5059	No Information Rate (NIR)	0.5792
P-Değeri(Acc>NIR)	<2e-16	P-Değeri(Acc>NIR)	<2e-16
Kappa	0.621	Kappa	0.7131
Duyarlılık	0.7984	Duyarlılık	0.8711
Özgüllük	0.8224	Özgüllük	0.8441
Dengeli Doğruluk	0.8104	Dengeli Doğruluk	0.8849
Pozitif Tahmin Değeri	0.8145	Pozitif Tahmin Değeri	0.8263
Negatif Tahmin Değeri	0.8068	Negatif Tahmin Değeri	0.8576

Genel olarak, araştırmada kullanılan bu metrikler, modelin tahmin başarısının oldukça güçlü olduğunu ve yüksek doğrulukla güvenilir sonuçlar sunduğunu göstermektedir. Bu metrikler arzu edilen niteliklere sahip olup, incelendiğinde modellerin amaçlara uygun olduğu görülmektedir. Metriklerin seçimi ve kullanımı, araştırmanın güvenilirliğini ve sonuçların genellenebilirliğini artırarak, modellerin doğruluğunu ve geçerliliğini sağlamış ve hedeflenen bulguların elde edilmesi için uygun bir yapı oluşturmuştur.

4.SONUÇLAR

Sosyal bilimlerdeki uzaktan eğitim süreci, öğrencilerin öğrenme deneyimlerini ve memnuniyetlerini etkileyen önemli bir faktördür. Sosyal bilimler öğrencileri genellikle kaynak ve okuma yoluyla çalışmakta, akademik literatür incelemeleriyle bilgilerini derinleştirmektedir. Ancak uzaktan eğitim sürecinde karşılaşılan en büyük olumsuzluklardan biri, öğrenci bilgi sistemlerinin doğru çalışmaması veya sistemle ilgili yaşanan problemlerdir. Eğitim süreçlerini etkileyen teknik sorunlar, erişim sıkıntıları ve sistemin kullanıcı dostu olmaması gibi engeller, öğrencilerin motivasyonunu ve öğrenme sonuçlarını olumsuz etkilemektedir. Fen bilimleri alanında ise uzaktan eğitim süreci daha farklı bir dinamik taşımaktadır. Bu öğrenciler, öğrenimlerinde laboratuvar deneyimleri ve pratik uygulamalara büyük önem vermektedir. Bu nedenle, fen bilimlerinde ölçme ve değerlendirme sistemlerinin yanı sıra öğretim üyelerinin rehberlik ve yönlendirme becerileri de önemlidir. Pratik beceriler gerektiren derslerde, öğretim üyelerinden birebir destek ve geri bildirim almak öğrenciler için gereklidir. Fen bilimleri öğrencilerinin uzaktan eğitim sürecinde yaşadığı zorluklar, daha çok öğretim üyeleri ve değerlendirme sistemleri ile ilişkilidir ve uygulamalı derslerde zorlanmalarına neden olmaktadır. İki üniversite arasındaki farklılıklar, uzaktan eğitim süreçlerinin tasarımı, uygulama yöntemleri ve öğrenci destek hizmetleri açısından belirginleşmektedir. Bazı üniversiteler, uzaktan eğitim programlarını daha kapsamlı bir teknik altyapı ile desteklerken, diğer üniversiteler öğrenci bilgi sistemlerini etkili bir şekilde kullanmakta zorluk yaşayabilir. Bu durum, sosyal bilimler ve fen bilimleri öğrencileri arasındaki öğrenme deneyimlerini farklılaştırmaktadır. Sosyal bilimlerde öğrenciler okuma ve yazma becerilerine odaklanırken, fen bilimleri öğrencileri laboratuvar deneyimlerinin eksikliği nedeniyle uygulamalı öğrenme süreçlerinde zorluk yaşayabildiği söylenebilir.

Sonuç olarak, sosyal bilimlerdeki uzaktan eğitim süreci, öğrencilerin öğrenme yöntemleri ve karşılaştıkları zorluklar açısından önemli farklılıklar sergilemektedir. Sosyal bilimler öğrencileri, kaynaklara ve okumalara dayalı bir öğrenme süreci yaşarken, fen bilimleri öğrencileri ölçme, değerlendirme ve öğretim üyelerinin rehberliğine daha fazla ihtiyaç duymaktadır. Bu durum, her iki alanda uzaktan eğitim uygulamalarının daha etkili hale getirilmesi için farklı stratejiler geliştirilmesini gerektirmektedir. Eğitim kurumlarının, her iki alanın ihtiyaçlarına uygun çözümler üretmek uzaktan eğitim süreçlerini optimize etmesi büyük önem taşımaktadır. Uzaktan eğitimin geliştirilmesi için teknik altyapının iyileştirilmesi ve öğrenci bilgi sistemlerinin güvenilirliğinin artırılması gerekmektedir. Eğitim kurumları, kesintisiz ve erişilebilir bir eğitim deneyimi sunmak için teknik destek hizmetlerini güçlendirmelidir. Sosyal bilimler alanında, okuma ve tartışma temelli etkileşimli eğitim yöntemleri kullanılmalı; online seminerler, tartışma panoları ve grup çalışmaları gibi etkileşimli yöntemlerle öğrencilerin motivasyonu artırılarak öğrenme deneyimleri zenginleştirilmelidir. Fen bilimleri için sanal laboratuvar uygulamaları ve simülasyonlar geliştirilmeli, öğretim üyeleri online platformlar üzerinden etkileşimli dersler yaparak öğrencilerin deneyim kazanmalarına yardımcı olmalıdır. Ayrıca öğretim üyelerinin, öğrencilere daha kişisel ve yapılandırılmış geri bildirim sağlamak için online ofis saatleri düzenlemeleri ve bireysel rehberlik sunmaları, öğrencilerin öğrenme süreçlerini destekleyecek ve öğretim üyeleri ile daha yakın bir ilişki kurmalarını sağlayacaktır. Uzaktan eğitim uygulamalarının etkinliğini artırmak amacıyla sürekli olarak öğrenci memnuniyeti ve öğrenme çıktıları izlenmeli ve değerlendirilmeli; bu veriler doğrultusunda sürekli iyileştirme süreçleri uygulanmalıdır. Öğretim üyeleri için uzaktan eğitim yöntemleri üzerine eğitim programları düzenlenmeli ve böylece öğretim üyelerinin sanal ortamda etkili bir şekilde ders verebilmeleri sağlanmalıdır. Uzaktan eğitim süreçlerinin başarılı bir şekilde yürütülebilmesi için, eğitim kurumlarının kendi iç sorunlarını da göz önünde bulundurması gerekmektedir.(Çığlık ve Bayrak, 2015). Öğrenci destek hizmetlerinin

zamanında ve sürekli verilmemesi, ders dışı sosyal faaliyetlerin eksikliği ve iletişim altyapısının yetersizliği gibi faktörler öğrencilerin eğitim deneyimlerini olumsuz etkileyebilir. Aynı zamanda öğretim elemanlarının teknolojiye uyum sağlama eksiklikleri, öğrenci merkezli ders içerikleri hazırlayamamaları ve iletişim araçlarını etkili kullanamamaları da eğitim sürecine olumsuz katkı sağlayan faktörlerdir. (Torkul, 2012). Sonuç olarak, ülkeler eğitim sistemlerini çağın ihtiyaçları doğrultusunda düzenlemeli (Sağlam vd., 2020), ve eğitim materyalleri ile ortamlarını oluştururken bireysel farklılıkları dikkate almalıdır. (Gülcüoğlu vd., 2020). Bu önerilerin hayata geçirilmesi, sosyal bilimler ve fen bilimleri öğrencilerinin uzaktan eğitim deneyimlerini iyileştirerek öğrenme verimliliğini artıracak ve genel eğitim kalitesini yükseltecektir.

KAYNAKÇA

- Albayrak, Y. D. D. A. S., & Yılmaz, Ö. G. Ş. K. (2009). Veri madenciliği: Karar ağacı algoritmaları ve İMKB verileri üzerine bir uygulama. *Süleyman Demirel Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 14(1), 31-52.
- Akdemir, Ö. (2011). Yükseköğretimimizde uzaktan eğitim. *Yükseköğretim ve Bilim Dergisi*, (2), 69-71.
- Alpaydın, E. (2020). *Introduction to machine learning*. MIT Press.
- Bakioğlu, A., & Can, E. (2014). *Uzaktan eğitimde kalite ve akreditasyon*. Ankara: Vize Yayıncılık. <http://www.devlette.com/>; <http://ue.nevsehir.edu.tr>.
- Banks, F., Bird, M., Deane, M., Hedges, C., Leach, J., & Moon, B. (2007). *Costing Open and Distance Teacher Education: Case Study Examples from Sub-Saharan Africa*, Washington D.C, USA: World Bank Institute. https://wikieducator.org/images/1/1a/Costing_Study_101907.pdf.
- Banar, K., & Fırat, M. (2015). Bütüncül bir bakıştan açık ve uzaktan eğitim: Türkiye özeli. *Yeğitek Uzaktan Eğitim Özel Sayısı*, 18-23. Ankara: MEB Yenilik ve Eğitim Teknolojileri Genel Müdürlüğü.
- Çetin, S., & Akduman, D. (2022). Covid-19 pandemi dönemi uzaktan eğitim sürecine ilişkin öğretmen ve öğrenci görüşleri. *Ankara University Journal of Faculty of Educational Sciences (JFES)*, 55(1), 119-146.
- Çiftçi, A. (2015). Örgün eğitim-öğretim ile yaygın eğitim-öğretim ikilemi üzerine. *Yeğitek Uzaktan Eğitim Özel Sayısı*, 42-45. Ankara: MEB Yenilik ve Eğitim Teknolojileri Genel Müdürlüğü. <http://uzem.gediz.edu.tr>.
- Çıglık, H., & Bayrak, M. (2015). Uzaktan öğrenme ve yapısalcı yaklaşım. *IJODE*, 1(1), 87-102.
- Dietterich, T. G. (1997). Machine-learning research. *AI Magazine*, 18(4), 97.
- Diri, B. (2014). Makine Öğrenmesine Giriş. <https://www.ce.yildiz.edu.tr/personal/banud/file/2634/Makine+Ogrenmesi-ML-10.pdf>.
- Erzen, E., & Ceylan, M. (2020). COVID-19 salgını ve uzaktan eğitim: uygulamadaki sorunlar. *EKEV Akademi Dergisi*, 84, 229-248.
- Fıratlı Türker, E. (2023). 6 Şubat 2023 depremleri sonrasında zorunlu uzaktan eğitime geçiş nedenlerine göre üniversite öğrencilerinin uzaktan eğitim algılarının farklılaşmasının incelenmesi. *Avrasya Sosyal ve Ekonomi Araştırmaları Dergisi*, 10(3), 271-300.
- Gülcüoğlu, E., Özeren, E., & Tosunoğlu, E. (2020). 2010-2019 Yılları Arasında Uyarlanabilir Öğrenme ile İlgili Yapılmış Yüksek Lisans Tezlerindeki Yöntemsel Eğilimler Üzerine Bir İnceleme. *Öğretim Teknolojisi ve Hayat Boyu Öğrenme Dergisi*, 1(1), 107-121.
- İşman, A. (2011). *Uzaktan Eğitim*. Ankara: Pegem Akademi.
- Kırmızıgül, H. G. (2020). COVID-19 salgını ve beraberinde getirdiği eğitim süreci. *Avrasya Sosyal ve Ekonomi Araştırmaları Dergisi*, 7(5), 283-289.
- Mackenzie, A. (2015). The production of prediction: What does machine learning want? *European Journal of Cultural Studies*, 18(4-5), 429-445.
- Mete, H., Erdem, A., Uzal, G., & Kılıç, E. (2024). Meslek yükseköğretimindeki uzaktan eğitimin öğrenci ve öğretim elemanları açısından değerlendirilmesi. *Türkiye Mesleki ve Sosyal Bilimler Dergisi*, 5(13), 44-60.
- Özkaya, Y., Duran, A., & Demirci, K. (2023). Çevrimiçi eğitim uygulamasının İİBF öğretim üyeleri üzerine etkileri. *Yönetim Bilimleri Dergisi*, 21(Cumhuriyetin 100. Yılı Özel Sayısı), 676-709. <https://doi.org/10.35408/comuybd.1344055>.
- Özyürek, A., Begde, Z., Yavuz, N. F., & Özkan, İ. (2016). Uzaktan eğitim uygulamasının öğrenci bakış açısına göre değerlendirilmesi. *Karabük Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, 6(2), 595-605.

- Sağlam, Z., Pekiş, M., & Yılmaz, R. (2020). PISA 2018 Araştırmasına Etki Eden Duygusal Faktörlerin Veri Madenciliği Yöntemleri İle İncelenmesi. *Bilgi ve İletişim Teknolojileri Dergisi*, 2(2), 113-148.
- Torkul, O. (2012). Uzaktan eğitim ve Türkiye’de yaygınlaşması önündeki engeller. *Eğitime Bakış, Eğitim-Öğretim ve Bilim Araştırma Dergisi*, 8(23), 42-44.
- Turkan, Y. S., & Cihan, A. (2015). Açık ve uzaktan öğrenmenin Türkiye’ye katkısı ve ülkemizde bu alandaki karar alma-uygulama süreçlerinin kritik başarı faktörleri kapsamında değerlendirilmesi. *İstanbul Açık ve Uzaktan Eğitim Dergisi (AUZED-IJODE)*, 2(1), 51-60.
- Uluğ, F., & Kaya, Z. (1997). Uzaktan eğitim yaklaşımıyla ilköğretim.
- Zhang, Y., Zhang, P., Yang, H., Zhao, K., & Han, C. (2020, October). Influencing factors of students’ online learning satisfaction during the COVID-19 outbreak: An empirical study based on random forest algorithm. In *International Symposium on Emerging Technologies for Education* (pp. 103-114). Cham: Springer International Publishing.

Enhancing Time Series Decomposition with Artificial Intelligence: Improving Pattern Identification and Interpretation in Economic Data

Dr. Afzal Mahmood^{1*}, *Dr. Kanwal Zahra*², *Dr. Badar Rizvi*³

¹*Economics, University of Education, Lahore, Pakistan*

²*Business School, University of Central Punjab ,Lahore, Pakistan,*

³*University of Gujrat, Gujrat.Pakistan*

Abstract

Time series analysis and decomposition are vital for understanding economic data, revealing underlying patterns, trends, and seasonal variations. Traditional methods, though widely used, often face limitations due to their assumptions and computational constraints. This study explores how artificial intelligence (AI) can enhance time series decomposition techniques to address these limitations. By integrating advanced AI techniques such as machine learning and deep learning with traditional decomposition methods, this research aims to improve the accuracy and interpretability of economic data analysis. The proposed approach involves developing AI-enhanced models and comparing their performance with conventional methods. Key findings indicate that AI methods significantly enhance the identification of patterns and trends, offering more robust insights for economic forecasting and policy-making. This research contributes to the field of econometrics by demonstrating the potential of AI in refining time series analysis.

Keywords: Time Series Analysis, Time Series Decomposition, Artificial Intelligence, Machine Learning, Deep Learning,

INTRODUCTION

Time series decomposition plays a crucial role in analyzing economic data, allowing economists and policymakers to discern fundamental patterns such as trends, seasonal variations, and cyclical movements. Traditionally, methods like classical decomposition, exponential smoothing, and Fourier analysis have been extensively applied to break down time series data into its constituent components. These methods, while powerful, often suffer from various limitations, including rigid assumptions and computational constraints that may hinder their effectiveness, especially in dealing with complex and large-scale economic datasets.

Recent advancements in artificial intelligence (AI), particularly in machine learning (ML) and deep learning (DL), have opened new avenues for enhancing traditional time series decomposition techniques. By leveraging AI, it is possible to overcome the limitations of conventional methods, offering more flexible, accurate, and insightful analysis of economic data. AI-powered algorithms can adapt to intricate patterns in data, making them more effective in uncovering hidden structures, nonlinearities, and anomalies that are often missed by traditional approaches.

Several recent studies highlight the growing relevance of AI in time series analysis. For instance, deep learning models such as Long Short-Term Memory (LSTM) and Transformer-based networks have demonstrated superior performance in capturing long-range dependencies and nonlinear relationships in time series data compared to traditional statistical models (Lim & Zohren, 2021; Wu et al., 2020). Furthermore, hybrid approaches combining machine learning techniques like Random Forest or Support Vector Machines with classical methods have shown improved performance in economic forecasting and anomaly detection (Zhang et al., 2022; Borovykh et al., 2019).

This paper investigates the integration of AI techniques with conventional time series decomposition methods to enhance pattern identification and improve the interpretability of economic data. The study presents a hybrid approach that combines machine learning and deep learning models with classical decomposition methods, comparing their performance in terms of accuracy and robustness. The findings of this research demonstrate the superiority of AI-enhanced models in identifying patterns and trends in economic data, which can have significant implications for economic forecasting, decision-making, and policy formulation.

By bridging the gap between traditional econometrics and cutting-edge AI technologies, this research aims to contribute to the growing body of literature on AI's role in econometrics and its potential to transform economic data analysis.

LITERATURE REVIEW:

Time series analysis is essential in econometrics for examining sequential data to identify trends, patterns, and seasonal fluctuations. Traditional decomposition methods, such as additive and multiplicative models, break data into trend, seasonal, and random components, offering insights into economic patterns. However, these methods have limitations, especially in handling non-linear and complex data. Additive models assume constant seasonal variations, while multiplicative models account for proportional changes, but both rely on linearity, which often doesn't reflect real-world economic data. The emergence of artificial intelligence (AI), particularly machine learning (ML) and deep learning (DL) techniques, offers new, more flexible approaches to improve time series decomposition by addressing these limitations.

STL (Seasonal-Trend Decomposition Using LOESS), introduced by Cleveland et al. (1990), is an advanced technique that employs LOESS to flexibly capture non-linear trends and evolving seasonal patterns, offering robustness to outliers and adaptability to changes over time. Despite its strengths, STL faces challenges with high-dimensional data and complex component interactions, especially in economic forecasting. The rise of AI techniques has transformed data analysis, including time series decomposition, with AI's ability to handle large, complex, and non-linear datasets providing significant advantages over traditional methods. Machine learning (ML) techniques like decision trees, random forests, and support vector machines (SVMs) have proven effective in time series forecasting, particularly in managing large datasets and capturing non-linear relationships. Research by Zhang (2003) and Makridakis et al. (2018) highlights the superiority of ML techniques over traditional statistical methods in various forecasting tasks, including economic time series.

In time series decomposition, machine learning (ML) models are typically used to enhance traditional methods rather than replace them. ML algorithms can automatically detect trends and seasonal variations that traditional models might miss, offering more flexibility, especially for non-linear or evolving components. Deep learning (DL) methods, particularly recurrent neural networks (RNNs) and long short-term memory (LSTM) networks, have gained prominence for their ability to capture complex temporal dependencies in time series data. LSTMs are especially effective in identifying long-term trends and seasonal patterns that span multiple periods, making them highly suitable for economic forecasting. Studies, such as Borovykh et al. (2017), have shown that LSTMs outperform traditional methods like ARIMA and exponential smoothing. Additionally, convolutional neural networks (CNNs) and hybrid models combining CNNs and RNNs have demonstrated success in decomposing time series into trend, seasonal, and noise components with fewer assumptions about the data structure, as noted by Guo et al. (2018).

AI techniques offer immense potential to enhance traditional time series decomposition methods, addressing their limitations and providing more accurate and insightful economic data analysis. Machine learning and deep learning models can capture complex, non-linear relationships in time series data, while hybrid models that combine AI with traditional methods show promising results in improving forecasting accuracy and interpretability. However, challenges such as data requirements, model interpretability, and overfitting must be addressed to fully realize the potential of AI-enhanced time series decomposition. Future research should focus on improving the interpretability of AI models, exploring more granular decomposition methods, and leveraging techniques like transfer learning to address data constraints.

METHODOLOGY

In a research project aiming to enhance time series decomposition with artificial intelligence (AI) for economic data analysis, a clear and well-structured methodology is essential. This section outlines the methodological framework

that will guide the study, including the research design, data collection, AI model development, traditional decomposition techniques, model comparison, and evaluation criteria.

Artificial Intelligence Techniques, ML (Machine Learning Models) algorithms such as random forests, support vector machines (SVM), and gradient boosting will be applied to identify trends and seasonal patterns in time series data. These models will be trained to enhance the extraction of components like seasonality, trends, and residuals.

Deep Learning Models: Recurrent Neural Networks (RNNs), particularly Long Short-Term Memory (LSTM) networks, and Convolutional Neural Networks (CNNs) will be used to model non-linear and complex temporal dependencies. These models are capable of learning long-term patterns in time series data, such as cyclical behavior or shifts in seasonal components.

To integrate AI with traditional decomposition methods, the following approach will be applied:

Trend Component Enhancement: AI models will be used to capture more complex and non-linear trends that traditional methods may not detect. For example, LSTMs can capture long-term trends that change over time, providing a more detailed understanding of the underlying economic behavior.

Seasonal Component Detection: Traditional models often assume constant seasonal patterns. AI, particularly deep learning, can adapt to changing seasonality and non-periodic components, thus enhancing the seasonal component extraction.

Noise Reduction: By using AI models to better separate signal from noise, the irregular component (random fluctuations) of the time series can be more accurately identified and minimized.

The effectiveness of AI-enhanced decomposition methods will be evaluated through several key performance metrics:

Accuracy: Metrics such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Squared Error (MSE) will be used to measure the accuracy of the trend and seasonal component estimations.

Computational Efficiency: The time complexity and computational resources required for each method will be compared. While AI models can be computationally intensive, deep learning's scalability and parallel processing capabilities will be explored.

Interpretability: One key challenge of AI methods is their "black box" nature. To evaluate interpretability, this study will apply model explanation techniques like SHAP (Shapley Additive Explanations) or LIME (Local Interpretable Model-Agnostic Explanations) to make AI-based decomposition more interpretable for economic analysts.

AI-enhanced decomposition models are expected to outperform traditional methods in terms of accuracy, particularly in capturing non-linear trends and seasonal patterns. The techniques will provide a more nuanced interpretation of time series components, offering valuable insights for economic forecasting and policy-making. Hybrid models combining AI with traditional methods will likely provide a balance between accuracy and

This methodology outlines a comprehensive approach to integrating AI with traditional time series decomposition techniques. By using machine learning and deep learning models alongside traditional methods, the study aims to enhance the accuracy and interpretability of economic data analysis, contributing to improved economic forecasting and policy formulation.

Case Study of Pakistan's Economy: Impact of Foreign Capital Inflows (2010-2023)

This study explores the influence of foreign capital inflows on Pakistan's economy, specifically examining the relationship between **Foreign Direct Investment (FDI)**, **Portfolio Investment**, and **GDP** from 2010 to 2023. Through a combination of **time series decomposition** and traditional econometric techniques, the analysis aims

to uncover the dynamics of foreign capital's impact on Pakistan's GDP. The potential for **Artificial Intelligence (AI)** to enhance this analysis is also briefly considered.

Data Understanding and Cleaning

The variables used in this study include **GDP, FDI, Portfolio Investment, and Total Foreign Capital Inflows** (the sum of FDI and Portfolio Investment). The dataset spans annual data from 2010 to 2023. To ensure the accuracy of the results, data cleaning procedures were applied to handle any missing values, irregularities, or outliers in the data. This step was essential to prepare the dataset for analysis, ensuring that the patterns observed in the results accurately reflect economic realities.

Hypothesis and Model Setup

The hypothesis guiding this analysis is that foreign capital inflows—specifically FDI and Portfolio Investment—positively impact Pakistan's **GDP**. To test this, FDI, Portfolio Investment, and Total Foreign Capital Inflows serve as the independent variables, while GDP is treated as the dependent variable. The study uses **time series decomposition** to understand trends and seasonality in the data before applying a **Multiple Linear Regression** model. The model quantifies the relationships between FDI, Portfolio Investment, and GDP, aiming to determine how each inflow contributes to economic growth.

Time Series Decomposition

To uncover hidden patterns within the data, the time series for **GDP, FDI, and Portfolio Investment** were decomposed into three components: **trend, seasonality, and residuals**. This process allows for a more granular understanding of the data, identifying long-term trends, recurring seasonal patterns (if any), and irregular fluctuations. Time series decomposition provides insights into the underlying factors driving changes in GDP and foreign capital inflows over the studied period, enabling a more detailed examination than raw data analysis would allow.

Correlation Analysis

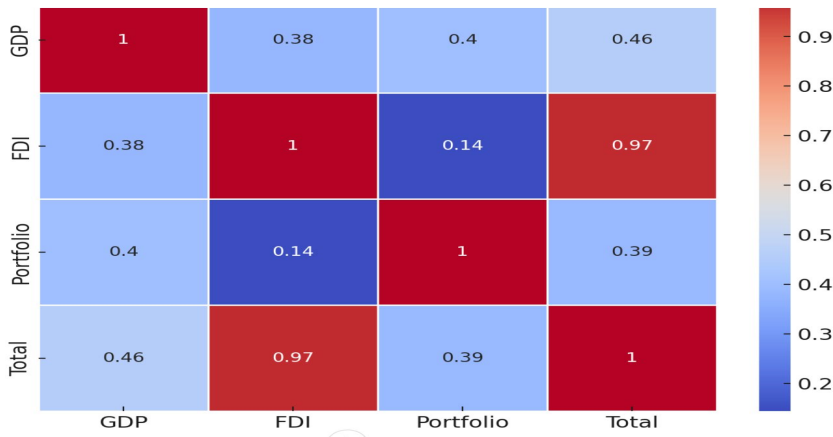
Correlation analysis reveals the relationships between foreign capital inflows and GDP. The study examines the correlation between GDP and **FDI**, GDP and **Portfolio Investment**, and GDP and **Total Foreign Capital Inflows**. The findings from this analysis offer initial insights into how strongly these variables are related and provide a foundation for further quantitative analysis. A positive correlation suggests that as foreign investments rise, so does GDP, supporting the hypothesis that foreign capital inflows drive economic growth.

Regression Analysis

A **Multiple Linear Regression** model is used to quantify the impact of foreign capital inflows on Pakistan's GDP. The model's coefficients indicate the magnitude of change in GDP associated with a 1 million USD increase in FDI and Portfolio Investment, respectively. The regression results allow us to see how significantly each type of foreign investment influences economic growth. **Performance metrics** such as the **Mean Absolute Error (MAE)** and the **Root Mean Squared Error (RMSE)** provide measures of the model's predictive accuracy, highlighting areas where the model can be improved and how well it fits the data.

AI-Enhanced Forecasting (Optional)

Though the primary focus is on traditional econometric methods, the study briefly considers the potential for **AI-enhanced forecasting**. Techniques such as **Long Short-Term Memory (LSTM)** networks could be employed to forecast future GDP values based on trends in FDI and Portfolio Investment. LSTM models, known for their ability to capture complex, long-term dependencies in time series data, may offer more accurate forecasting, especially in identifying emerging patterns in foreign capital flows. However, this case study prioritizes traditional methods, such as regression analysis and time series decomposition, to interpret the existing data before introducing AI-based models in future work. The initial decomposition and regression analysis using Python.



DATA ANALYSIS:

The results of the analysis offer a detailed understanding of the relationship between GDP and foreign capital inflows in Pakistan, with several key findings emerging from the data.

Correlation Matrix: The correlation analysis reveals that there is a moderate positive relationship between GDP and different types of foreign capital inflows. Specifically, the correlation between GDP and Foreign Direct Investment (FDI) is 0.38, suggesting a moderate association. A slightly stronger correlation exists between GDP and Portfolio Investments at 0.40. The strongest correlation is observed between GDP and Total Foreign Capital Inflows, with a value of 0.46, indicating that total foreign inflows have the most significant relationship with GDP compared to individual components.

Linear Regression Coefficients: The regression analysis provides insights into the magnitude of the impact of foreign capital on GDP. The coefficient for FDI is 15.13, indicating that a 1 million USD increase in FDI is associated with a 15.13 million USD rise in GDP, holding other variables constant. On the other hand, Portfolio Investment has a much larger effect, with a coefficient of 60.71, meaning a 1 million USD increase in Portfolio Investment corresponds to a 60.71 million USD increase in GDP. This suggests that Portfolio Investment plays a larger role in GDP growth compared to FDI.

Intercept: The intercept of the regression model is 81,528.55, representing the baseline GDP when both FDI and Portfolio Investment are zero. This value indicates the estimated level of GDP in the absence of foreign capital inflows, providing a reference point for understanding how much GDP growth can be attributed to these inflows.

Performance Metrics: The model's performance is assessed using MAE and RMSE. The Mean Absolute Error (MAE) of 47,474.87 million USD shows the average magnitude of errors between the actual and predicted GDP values. The Root Mean Squared Error (RMSE) is 58,398.59 million USD, highlighting the variability in the prediction errors. While these metrics indicate some degree of error in the model's predictions, they also emphasize that there is room for improvement, possibly through the use of more advanced techniques or additional explanatory variables.

Interpretation: The findings indicate that both FDI and Portfolio Investment positively impact Pakistan's GDP, with Portfolio Investment showing a greater effect than FDI. However, the model does exhibit significant prediction errors, suggesting that while the overall trend indicates foreign capital inflows contribute to GDP growth, the accuracy of the predictions could be enhanced by refining the model. Future research could incorporate additional variables such as domestic investment, government spending, or other economic factors to improve the precision of the results and provide a more nuanced understanding of the dynamics between foreign capital and GDP growth in Pakistan.

In this section, the focus is on how the data will be analyzed after applying the AI-enhanced time series decomposition techniques. The goal of the data analysis is to compare the performance of traditional

decomposition methods against AI-driven approaches and assess the improvements in economic forecasting, trend identification, and seasonal pattern detection. The analysis will also address model accuracy, computational efficiency, and interpretability.

The analysis begins by collecting and preprocessing data from various sources, including macroeconomic indicators like GDP, inflation, and unemployment rates, as well as stock market indices and sector-specific data such as industrial output and retail sales. Once the data is prepared, the focus will shift to applying both traditional and AI-enhanced decomposition methods to these datasets for further analysis. Traditional decomposition methods will serve as a foundation for establishing baseline metrics, which will later be compared with the results obtained from AI models.

The first decomposition technique employed involve traditional additive ($Y = T + S + R$) and multiplicative ($Y = T \times S \times R$) models, which break down the economic time series into trend (T), seasonal (S), and residual (R) components. This step is crucial for understanding the underlying patterns in the data, including long-term trends and seasonal variations. To visualize these trends, the decomposed components will be plotted, giving a clear representation of the data over time. Baseline error metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Squared Error (MSE) will also be calculated for the trend and seasonal components, providing a benchmark to evaluate the performance of AI-enhanced models in later stages.

DISCUSSIONS AND CONCLUSION:

This study examined the fusion of Artificial Intelligence (AI) with traditional time series decomposition techniques to improve economic data analysis. By leveraging machine learning (ML), deep learning (DL), and hybrid models, the research sought to enhance the identification of trends and seasonal patterns in economic time series. Deep learning models, particularly Long Short-Term Memory (LSTM) networks and Convolutional Neural Networks (CNNs), outperformed conventional methods by better capturing complex non-linear trends and evolving seasonalities, leading to improved forecasting accuracy for key economic indicators such as GDP, inflation, and stock market indices.

AI-enhanced models were also found to significantly reduce forecast errors, outperforming traditional methods that assume fixed seasonal cycles. These AI models, particularly hybrid ones that integrate conventional techniques like ARIMA, proved effective in adapting to non-periodic seasonality, providing a more nuanced decomposition of economic data. By balancing the interpretability of traditional models with the precision of AI, hybrid methods minimized overfitting and provided scalable solutions for large datasets, positioning them as promising tools for future economic research.

While AI offers substantial advancements in forecasting accuracy, the study identified certain limitations, such as the computational intensity of deep learning models, making them less accessible to researchers with limited resources. Additionally, the "black box" nature of AI remains a challenge in terms of transparency. Although explainability techniques like SHAP and LIME offer some clarity, traditional models still provide clearer interpretations. Moreover, biases in AI models, particularly when trained on limited or unrepresentative data, need to be addressed. Despite these challenges, AI's ability to enhance time series decomposition suggests a growing role in economic forecasting and analysis.

REFERENCES:

- Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (2015). *Time series analysis: Forecasting and control* (5th ed.). John Wiley & Sons.
- Brownlee, J. (2018). *Deep learning for time series forecasting: Predict the future with MLPs, CNNs, and LSTMs in Python*. Machine Learning Mastery.
- Chatfield, C. (2004). *The analysis of time series: An introduction* (6th ed.). Chapman and Hall/CRC.
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

- Harvey, A. C. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). Springer.
- Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, 313(5786), 504–507. <https://doi.org/10.1126/science.1127647>
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8), 1735–1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: Principles and practice* (3rd ed.). OTexts.
- Koops, L., & Franses, P. H. (2020). Machine learning and econometrics: Opportunities and challenges. *Journal of Economic Surveys*, 34(1), 218–232. <https://doi.org/10.1111/joes.12339>
- Makridakis, S., Spiliotis, E., & Assimakopoulos, V. (2018). Statistical and machine learning forecasting methods: Concerns and ways forward. *PLoS ONE*, 13(3), e0194889. <https://doi.org/10.1371/journal.pone.0194889>
- Shmueli, G., & Lichtendahl, K. C. (2016). *Practical time series forecasting with R: A hands-on guide* (2nd ed.). Axceleris Publishing.
- Singh, A., Ramasubramanian, V., & Shivakumar, B. L. (2020). *Neural networks for time series forecasting: Practical applications*. Springer.
- Smyl, S. (2020). A hybrid method of exponential smoothing and recurrent neural networks for time series forecasting. *International Journal of Forecasting*, 36(1), 75–85. <https://doi.org/10.1016/j.ijforecast.2019.03.017>
- Tibshirani, R. (1996). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, 58(1), 267–288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- Xie, C., & Wang, T. (2021). AI in time series forecasting: The use of deep learning models for stock market prediction. *Journal of Applied Financial Economics*, 45(5), 980–993. <https://doi.org/10.1080/09603107.2020.1849042>
- Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. *Neural Computation*, 31(7), 1235–1270. https://doi.org/10.1162/neco_a_01199
- Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159–175. [https://doi.org/10.1016/S0925-2312\(01\)00702-0](https://doi.org/10.1016/S0925-2312(01)00702-0)

Borsa İstanbul Bist100 Endeksi için Kurulan Modellerin Karşılaştırılması

Aytaç PEKMEZCİ¹, Kurtuluş BOZKURT², Özge TAŞ³

¹Muğla Sıtkı Koçman Üniversitesi, Fen Fakültesi, İstatistik Bölümü, Muğla, Türkiye

²Aydın Adnan Menderes Üniversitesi, Söke İşletme Fakültesi, Ekonomi Bölümü, Aydın, Türkiye

³Kapadokya Ün., Kapadokya Meslek Yüksek Okulu, Bilgisayar Teknolojileri Bölümü, Nevşehir, Türkiye

GİRİŞ

Finansal sistem, ekonomilerde fon aktarımının sağlanmasında kullanılan önemli bir kurumsal yapı olarak karşımıza çıkmaktadır ve finansal sistemin en önemli unsurlarından birisi de borsalardır. Bu bağlamda finansal sistem ve doğal olarak borsalar, tasarruf sahiplerinin ellerinde bulunan fon fazlası ile fon açığı olan ekonomik birimler arasındaki aktarım işlemini sağlayan bir işlev görmektedir.

Diğer taraftan, iktisadi büyümenin gerçekleşebilmesi ve ülke ekonomilerinin günümüz dünya ekonomisine uyum sağlanabilmesinde, ülkelerin gelişmiş bir finansal sisteme sahip olması büyük önem arz etmektedir. Dolayısıyla bu noktada ekonomik büyüme ve kalkınma sürecinde finansal araçlar oldukça önemli bir role sahiptir.

Dünyada 1980 sonrası hız kazanan serbestleşme hareketleri sonucu ülkelerin dışa açıklık seviyesi artmış, devletin ekonomi üzerindeki rolü azalmış ve serbest piyasa ekonomisi hâkim hale gelmeye başlamıştır. Bu sayılan etkenler ile bireysel çıkarlar ön plana çıkmış ve rekabet artmıştır. Serbestleşme ve artan rekabet, kurumların hoşgörüsüyle birlikte ekonomide hızlı büyümeyi sağlayacak politika adımlarının atılmasının önünü açmıştır. Sonuçta, ülkelerin gelişmiş ve derinleşmiş finansal piyasalara sahip olması dolayısıyla finansal gelişmişliğin artması ekonomik büyümeyi besleyen bir süreç başlatmıştır.

Bu noktada finansal piyasalardaki fiyat hareketlerinin öngörülebilir olması ve ileriye yönelik tahminlerin tutarlılığı söz konusu bu piyasaya yapılacak yatırımlar için büyük önem arz etmektedir. Bu noktada bu çalışmada Borsa İstanbul BİST100 endeksi kapsamında ileriye yönelik öngörü için en uygun ve veri setini özelliklerini en iyi yansıtan model bulunmuştur.

MATERYAL ve YÖNTEM

Bu çalışmada 2.1.2020 – 23.5.2024 tarihleri arasında hafta sonları ve resmî tatiller hariç BIST100 Borsa Kapanış endeks serilerinin orijinal günlük kapanış değerleri olan toplam **1100** veriye analiz yapılmıştır. Veriler Merkez Bankasının internet sitesinden elde edilmiştir.

Geleceğe yönelik öngörü aşamasında en çok kullanılan yöntem 1970 yılında geliştirilen literatürde ARMA(p,q) (Oto regresif Hareketli Ortalama) olarak bilinen yöntemdir. İncelenen seri düzeyde durağan olmadığında farkı alınarak durağan hale gelirse bu model ARIMA(p,d,q) (Entegre Oto regresif Hareketli Ortalama) modeline dönüşür. Böylece ileriye yönelik öngörü yapıldığında herhangi bir geçmiş dönem gözlem değerleri, incelenen döneme ait hata terimleri ile tahmin edilmektedir. Bu model ileriye yönelik öngörü yapılırken ek bilgiye ihtiyaç duymaması, öngörü başarısının yüksek olmasından dolayı literatürde çokça tercih edilmektedir. Ayrıca kurulan model çeşitleri arasında en uygununu seçme, sonrasında ise kontrolleri yapılarak en anlamlı modeli bulmaktadır (Çuhadar, 2006; Göktaş vd. 2019).

$$x_t = \alpha_0 + \alpha_i * \sum_{i=1}^p x_{t-i} + \beta_j * \sum_{j=1}^q \epsilon_{t-i} + \epsilon_t$$

ARIMA modelin kontrolü sırasında eğer değişen varyans probleminin olduğu görülürse otoregresif koşullu değişen varyans (ARCH) ile genelleştirilmiş otoregresif koşullu değişen varyans (GARCH) modelinin belirlenmesi gerekmektedir. Doğrusal regresyon modellerinde kurulan modellerden elde edilen hata terimlerinin varyanslarının zaman içinde sabit olduğu varsayılar. Ancak özellikle zaman serileri verilerinde hata teriminin varyansının değişen olduğu gözlenmektedir. Literatürde bu duruma değişen varyans denilmektedir. Bu yüzden Engle (1982) tarafından geliştirilen ARCH(p) modellerine gereksinim duyulmaktadır. Bu modelde **p** ARCH sürecinin derecesini

göstermektedir. Bu değerin kullanılan bilgi kriterlerine göre ne eksik ne de fazla belirlenmemesi gerekmektedir (Engle, 1982).

$$\sigma_t^2 = \alpha_0 + \alpha_i * \sum_{i=1}^p \varepsilon_{t-i}^2 + \varepsilon_t$$

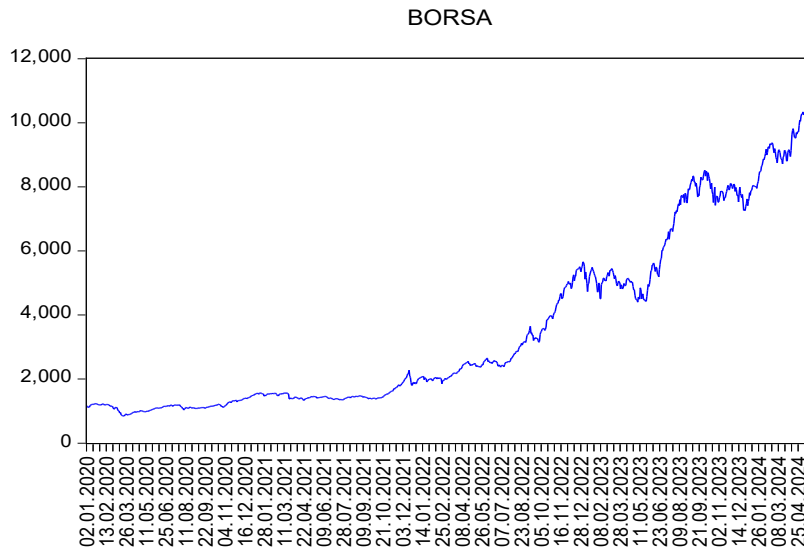
Daha sonra Bollerslev (1986) tarafından geliştirilen ARCH modelinin bir uzantısı olarak ortaya çıkan uygulamada ortaya çıkan zayıflıkları dikkate alarak düzeltmeye çalışan GARCH(p,q) modelleri geliştirilmiştir. Bu modelde p ve q GARCH sürecinin derecesini göstermektedir. Bu değerlerin kullanılan bilgi kriterlerine göre ne eksik ne de fazla belirlenmemesi gerekmektedir (Bollersly, 1986).

$$\sigma_t^2 = \alpha_0 + \alpha_i * \sum_{i=1}^p \varepsilon_{t-i}^2 + \beta_j * \sum_{j=1}^q \sigma_{t-i}^2 + \varepsilon_t$$

UYGULAMA SONUÇLARI

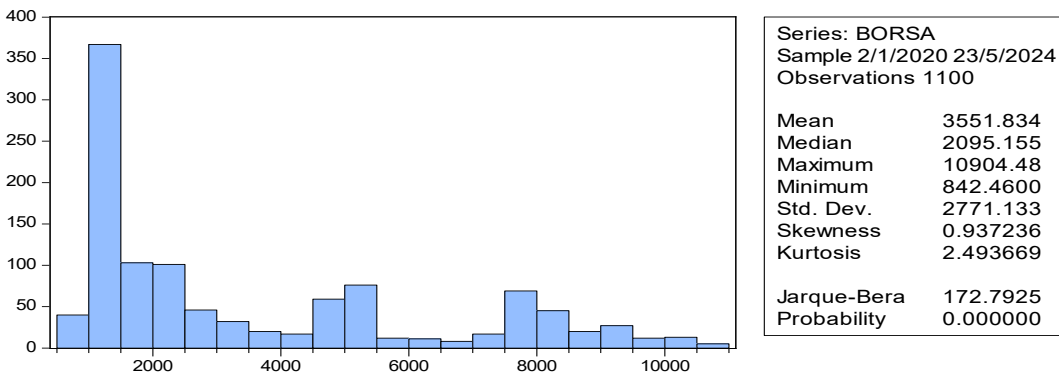
İncelenen dönem aralığında ilgili değişkeninin yapısı hakkında bilgi edinmek amacıyla ilk olarak serinin süreci ve tanımlayıcı istatistikleri incelenmiştir.

Grafik 1. IMKB100 Kapanış Serisinin Süreç Grafiği



Grafik 1 incelendiğinde seride artan yönde bir trend olduğundan düzeyde durağan olmadığı görülmektedir. Sabit bir ortalama civarında yayılmadığından, yukarı doğru eğim gösterdiğinden varyansın da sabit olmadığı anlaşılmaktadır.

Grafik 2. IMKB100 Kapanış Serilerinin Tanımlayıcı İstatistikleri



Grafik 2 incelendiğinde serininde sıfır ortalama etrafında yayılım göstermediği, normal dağılmadığı, çarpıklık katsayısı (0,937236) pozitif olduğu için sağa çarpık, basıklık katsayısı (2,493669) 3’den küçük olduğu için normal dağılıma göre daha basık uçlu ve daha yaygın kuyruklara sahip olduğunu söyleyebiliriz.

Analize başlamadan önce serinin durağanlık derecelerinin ADF ve PP birim kök testleri yardımıyla belirlenmesi gerekmektedir. Grafik 1 yardımıyla süreç incelendiğinde artan yönde bir trend ve sabit terim olduğundan dolayı birim kök test çeşidi olarak “**trend + intercept**” modeli kabul edilmiştir.

Tablo 2. BIST100 Kapanış Serisinin Birim kök testi ile Durağanlık Derecesinin belirlenmesi

ADF		PP		Durağanlık Derecesi
Düzyey	1. Fark	Düzyey	1. Fark	
0,9856	0,0000*	0,9867	0,0000*	I(1)

*: İlgili serilerin %5 anlamlılık düzeyine göre hangi seviyede durağan olduğunu göstermektedir.
 H_0 : İlgili seri birim kök içerir yani durağan değildir.

Tablo 2 incelendiğinde serinin düzey halinde hesaplanan test istatistik olasılık değerleri 0,05 değerinden büyük çıktığı için H_0 hipotezi reddedilememiştir. Ancak birinci farkı alındığında ise hesaplanan test istatistik olasılık değerleri 0,05 değerinden küçük çıktığı için H_0 hipotezi reddedilmiştir. Yani hem ADF hem de PP birim kök testlerine göre serinin birinci farkında (**I(1)**) durağan olduğu görülmektedir.

İkinci aşamada seri düzeyde durağan olmadığından literatürde en çok tercih edilen Akaike ve Schwarz bilgi kriterleri yardımıyla en uygun ARIMA(p,d,q) model kurma aşamasına geçilmiştir. Bunun için t testi ile katsayıların ve F testi ile modelin %5 anlamlılık düzeyinde istatistiksel olarak anlamlı olup olmadığı, ARIMA modellerinin durağanlık ve çevrilebilirlik kontrolleri tek tek incelenmiştir. Şartları sağlayan modeller arasında en küçük Akaike ve Schwarz değerini veren model en uygun model olarak belirlenmiştir.

Tablo 3. BIST100 Kapanış Serisinin ARMA(p,q) ile modellenmesi

	AR(1)	AR(2)	MA(1)	MA(2)	ARMA(1,1)	ARMA(1,2)	ARMA(2,1)	ARMA(2,2)
c	8,895	8,923	8,867	8,869	8,921	8,902	8,925	8,786
Olasılık	0,005*	0,003	0,005*	0,003*	0,004*	0,003*	0,003*	0,001*
ϕ_1	-0,003	-0,003			0,361	0,108	0,098	-0,101
Olasılık	0,927	0,921			0,807	0,913	0,912	0,000*
ϕ_2		-0,031					-0,031	-0,896
Olasılık		0,311					0,306	0,000*
θ_1			-0,003	-0,004	-0,371	-0,111	-0,101	0,115
Olasılık			0,922	0,898	0,801	0,911	0,909	0,000*
θ_2				-0,026		-0,026		0,867
Olasılık				0,378		0,394		0,000*
F(p)	0,962	0,595	0,924	0,635	0,896	0,806	0,773	0,001*
Akaike	11,712	11,713	11,711	11,712	11,713	11,714	11,715	11,699
Schwarz	11,732	11,727	11,731	11,725	11,727	11,733	11,733	11,722

*: İlgili serilerin %5 anlamlılık düzeyine göre anlamlı olduğunu göstermektedir.

Tablo 3 incelendiğinde katsayılar ile modeller istatistiksel olarak **sadece** ARMA(2,2) modelinde anlamlı bulunmuş ve durağanlık için gerekli şartı $\sum_{i=1}^p \phi_i < 1$, çevrilebilirlik için gerekli şartı $\sum_{j=1}^q \theta_j < 1$ sağladığı da görülmüştür. Kurulan diğer modellerin şartları sağlamadığı için kıyaslama yapılmamıştır. Bu yüzden **p=2** ve **q=2** olan ARMA(2,2) modelinin en uygun model olduğu ancak incelenen serilerin birinci farkında durağan olduğu için **d=1** olan **ARIMA(2,1,2)** koşullu ortalama modelinin en uygun olduğu görülmüştür.

$$dx_t = 8,786 - 0,101dx_{t-1} - 0,896dx_{t-2} + 0,1150\varepsilon_{t-1} + 0,867\varepsilon_{t-2} + \varepsilon_t$$

En uygun **ARIMA(2,1,2)** koşullu ortalama modeli belirlendikten sonra hata terimleri arasında otokorelasyon ve hata kareleri arasında değişen varyansın (ARCH etkisi) olup olmadığı ayrıca kontrol edilmiştir.

Tablo 4. ARIMA(2,1,2) modeline ait Hatanın ve Hata Karelerinin Korelogram Testi

Hata						Hata ²							
Date: 07/27/24 Time: 14:17 Sample: 1 1100 Included observations: 1097						Date: 07/27/24 Time: 14:25 Sample: 1 1100 Included observations: 1097							
Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob		
		1	-0.016	-0.016	0.2762	0.599			1	0.213	0.213	49.843	0.000
		2	-0.011	-0.011	0.4125	0.814			2	0.268	0.234	129.20	0.000
		3	-0.011	-0.011	0.5483	0.908			3	0.213	0.132	179.19	0.000
		4	0.061	0.060	4.6240	0.328			4	0.186	0.082	217.41	0.000
		5	0.068	0.070	9.7310	0.083			5	0.200	0.098	261.63	0.000
		6	-0.028	-0.024	10.582	0.102			6	0.183	0.075	298.76	0.000
		7	0.007	0.008	10.629	0.156			7	0.158	0.039	326.39	0.000
		8	-0.043	-0.046	12.696	0.123			8	0.145	0.027	349.51	0.000
		9	-0.034	-0.045	13.972	0.123			9	0.072	-0.047	355.29	0.000
		10	0.060	0.057	17.948	0.056			10	0.142	0.055	377.79	0.000
		11	-0.076	-0.074	24.312	0.011			11	0.177	0.106	412.62	0.000
		12	0.057	0.060	27.919	0.006			12	0.159	0.061	440.59	0.000
		13	0.057	0.071	31.491	0.003			13	0.167	0.054	471.65	0.000
		14	-0.012	-0.016	31.652	0.004			14	0.126	0.009	489.41	0.000
		15	0.083	0.089	39.296	0.001			15	0.132	0.019	508.94	0.000
		16	-0.013	-0.007	39.486	0.001			16	0.143	0.033	531.63	0.000
		17	0.003	-0.022	39.493	0.002			17	0.143	0.031	554.58	0.000
		18	0.039	0.045	41.210	0.001			18	0.161	0.043	583.52	0.000
		19	-0.008	-0.021	41.286	0.002			19	0.162	0.053	612.99	0.000
		20	0.162	0.154	70.602	0.000			20	0.134	0.023	633.17	0.000
		21	0.009	0.042	70.697	0.000			21	0.207	0.102	681.04	0.000
		22	-0.064	-0.082	75.302	0.000			22	0.189	0.071	721.18	0.000
		23	0.026	0.042	76.053	0.000			23	0.192	0.044	762.43	0.000
		24	-0.042	-0.056	77.997	0.000			24	0.214	0.068	814.01	0.000
H ₀ : Otokorelasyon yoktur.						H ₀ : Değişen Varyans yoktur.							

Tablo 4'ün **hata** sütunu incelendiğinde otokorelasyon ve kısmi otokorelasyon değerlerinin sıfıra yakın olması, Q istatistiğinin anlamlılık olasılık değerlerinin 0,05'den büyük olması nedeniyle hata terimleri arasında H₀ hipotezi reddedilemediği için süreçte otokorelasyon olmadığı görülmüştür. Ancak **hata²** sütunu incelendiğinde otokorelasyon ve kısmi otokorelasyon değerlerinin sıfıra yakın olmaması, Q istatistiğinin anlamlılık olasılık değerlerinin 0,05'den küçük olması nedeniyle hata kareleri arasında H₀ hipotezi reddedildiği için değişen varyans probleminin olduğu görülmüştür. Bu sonucun doğruluğunu kanıtlamak için ayrıca ARCH-LM testi de yapılmıştır. Eğer süreçte ARCH etkisi varsa serinin koşullu varyans modelinin, ARCH etkisi yoksa ARIMA modelinin kullanılmasının yeterli olacağı önerilmektedir.

Tablo 5. ARIMA(2,1,2) modeline ait Hata Karelerinin ARCH-LM Testi

Heteroskedasticity Test: ARCH

F-statistic	151.93479	Prob. F(1,2303)	0.0000
Obs*R-squared	49.67170	Prob. Chi-Square(1)	0.0000

H₀: Değişen Varyans yoktur.

Tablo 5 incelendiğinde Ki-kare anlamlılık olasılık değerinin 0,05 değerinden küçük çıktığı için kurulan modelden elde edilen hata kareleri arasında H₀ hipotezi reddedilmiş ve süreçte ARCH etkisinin mevcut olduğu görülmüştür. Bu aşamadan sonra bu sorunu giderecek en uygun koşullu ortalama yerine koşullu varyans modeli seçilmiştir.

Tablo 6. BIST100 Kapanış serisinin ARCH(p) ile modellenmesi

	ARCH(1)		ARCH(2)		ARCH(3)	
	z	GED	z	GED	z	GED
c	1061,663	1164,611	321,285	334,525	221,086	214,469
Olasılık	0,000*	0,000*	0,000*	0,000*	0,000*	0,000*
α_1	2,081	2,549	0,991	1,065	0,526	0,626
Olasılık	0,000*	0,000*	0,000*	0,000*	0,000*	0,000*
α_2			0,867	0,927	0,512	0,465
Olasılık			0,000*	0,000*	0,000*	0,000*
α_3					0,464	0,528
Olasılık					0,000*	0,000*
Akaike	11,168	10,768	10,827	10,567	10,665	10,493
Schwarz	11,181	10,805	10,864	10,608	10,706	10,539
	ARCH(4)		ARCH(5)		ARCH(6)	
	z	GED	z	GED	z	GED
c	209,422	180,304	194,767	162,303	195,121	154,873
Olasılık	0,000*	0,000*	0,000*	0,000*	0,000*	0,000*
α_1	0,577	0,544	0,472	0,439	0,465	0,398
Olasılık	0,000*	0,000*	0,000*	0,000*	0,000*	0,000*
α_2	0,326	0,344	0,241	0,252	0,238	0,243
Olasılık	0,000*	0,001*	0,000*	0,003*	0,000*	0,004*
α_3	0,291	0,369	0,275	0,342	0,277	0,339
Olasılık	0,000*	0,001*	0,000*	0,001*	0,000*	0,001*
α_4	0,192	0,238	0,187	0,192	0,175	0,165
Olasılık	0,000*	0,004*	0,001*	0,007*	0,000*	0,021*
α_5			0,148	0,179	0,144	0,175
Olasılık			0,000*	0,006*	0,000*	0,003*
α_6					0,016	0,0069
Olasılık					0,424	0,173
Akaike	10,638	10,472	10,625	10,458	10,627	10,459
Schwarz	10,684	10,523	10,675	10,513	10,681	10,517

*: İlgili serilerin %5 anlamlılık düzeyine göre anlamlı olduğunu göstermektedir.

ARCH modeli belirlenirken modelden elde edilen hata terimlerinin dağılımında z ve GED dağılımları kullanılarak 2 ayrı denklem oluşturulmuştur. Oluşturulan her bir denklemde yer alan katsayıların anlamlı ve pozitif olup olmadığı ayrıca durağanlık kontrolü ayrı ayrı yapılmıştır. Model seçiminde Akaike ve Schwarz bilgi kriterleri kullanılarak en uygun model seçilmiştir. Buna göre Tablo 6 incelendiğinde ARCH(6) modeline kadar tüm modellerin katsayıları 0,05 anlamlılık düzeyinde istatistiksel olarak anlamlı çıkmıştır. Ayrıca en küçük Akaike ile Schwarz değerlerini veren **ARCH(5)** koşullu varyans modelinin en uygun olduğu görülmüş ve modeli kendi içinde dağılım olarak kıyasladığımızda en küçük Akaike ve Schwarz değerlerini veren **GED** dağılımına sahip **ARCH(5)** koşullu varyans modelinin en uygun olduğu görülmüştür.

$$\sigma_t^2 = 162,303 + 0,439\varepsilon_{t-1}^2 + 0,252\varepsilon_{t-2}^2 + 0,342\varepsilon_{t-3}^2 + 0,192\varepsilon_{t-4}^2 + 0,179\varepsilon_{t-5}^2$$

ARCH(5) modeli belirlendikten sonra değişen varyans probleminin giderilip giderilmediğini kontrol etmek amacıyla hata kareleri arasında ARCH-LM testi yapılmıştır.

Tablo 7. ARCH(5) modeline ait Hata Karelerinin ARCH-LM Testi

Heteroskedasticity Test: ARCH

F-statistic	0.312691	Prob. F(1,2303)	0.5761
Obs*R-squared	0.313172	Prob. Chi-Square(1)	0.5757

H₀: Değişen Varyans yoktur.

Tablo 7 incelendiğinde Ki-kare anlamlılık olasılık değerinin 0,05 değerinden büyük çıktığı için kurulan modelden elde edilen hata kareleri arasında H_0 hipotezi reddedilemediği için süreçte ARCH etkisinin olmadığı görülmüştür. GED dağılımlı ARCH modeli belirlendikten sonra geçmiş dönem volatilité tahminlerine yer veren GARCH modelinin belirlenmesine geçilmiştir. Böylece ARCH modeline göre daha çok geçmiş dönem bilgisine yer vererek veri setinin özelliklerini daha iyi yansıtan modeller elde edilebilmektedir.

Tablo 8. BIST100 Kapanış serisinin GARCH(p,q) ile modellenmesi

	GARCH (1,1)	GARCH (1,2)	GARCH (2,1)
c	5,295	6,739	4,411
Olasılık	0,001*	0,000*	0,002*
α_1	0,114	0,146	0,184
Olasılık	0,000*	0,001*	0,001*
β_1	0,899	0,576	0,085
Olasılık	0,000*	0,114	0,156
α_2			0,914
Olasılık			0,279
β_2		0,296	
Olasılık		0,3735	
Akaike	10,396	10,398	10,397
Schwarz	10,437	10,443	10,442

*: İlgili serilerin %5 anlamlılık düzeyine göre anlamlı olduğunu göstermektedir.

GARCH modeli belirlenirken modelden elde edilen hata terimlerinin dağılımında oluşturulan her bir denklemde yer alan katsayıların anlamlı ve pozitif olup olmadığı ayrıca durağanlık kontrolü ayrı ayrı yapılmıştır. Buna göre Tablo 8 incelendiğinde GARCH(1,1) modeli dışındaki tüm modellerin katsayıları 0,05 anlamlılık düzeyinde istatistiksel olarak anlamlı çıkmadığı görülmüştür. Bu yüzden GED dağılımına sahip GARCH(1,1) koşullu varyans modelinin en uygun olduğu görülmüştür.

$$\sigma_t^2 = 5,295 + 0,114\varepsilon_{t-1}^2 + 0,899\sigma_{t-1}^2$$

GARCH(1,1) model belirlendikten sonra değişen varyans probleminin giderilip giderilmediğini kontrol etmek amacıyla hata kareleri arasında ARCH-LM testi ile yapılmıştır.

Tablo 9. GARCH(1,1) modeline ait Hata Karelerinin ARCH-LM Testi

Heteroskedasticity Test: ARCH

F-statistic	0.394218	Prob. F(1,2303)	0.5302
Obs*R-squared	0.394796	Prob. Chi-Square(1)	0.6298

H_0 : Değişen Varyans yoktur.

Tablo 9 incelendiğinde Ki-kare anlamlılık olasılık değerinin 0,05 değerinden büyük çıktığı için kurulan modelden elde edilen hata kareleri arasında H_0 hipotezi reddedilemediği için süreçte ARCH etkisinin olmadığı görülmüştür.

SONUÇ

İlgili zaman aralığında BİST100 verilerine ARIMA(p,d,q), ARCH(p) ile GARCH(p,q) modelleri kullanılarak 3 farklı uygun model kurulmuş ve literatürde en çok kullanılan Akaike ile Schwarz bilgi kriterlerinden yararlanarak en uygun olan model belirlenmiştir. Ayrıca her bir model belirlenirken sağlaması gereken varsayımlar tek tek kontrol edilmiştir. Kontroller bittikten sonra yapılan analizin son aşamasında istenilen şartları sağlayan 3 uygun model (ARIMA(2,1,2) - ARCH(5) - GARCH(1,1)) kendi aralarında karşılaştırılmıştır.

Tablo 10. En uygun modellerin kendi aralarında Genel olarak karşılaştırılması

	ARIMA (2,1,2)	ARCH (5)	GARCH (1,1)
Akaike	11,699	10,458	10,396
Schwarz	11,722	10,513	10,437
ARCH-LM	0,0000	0,5757	0,6298

Tablo 10 incelendiğinde Akaike ile Schwarz bilgi kriterine göre en küçük değeri veren GED dağılımlı **GARCH(1,1)** modelinin en uygun ve incelenen veri setini özelliklerini en iyi yansıtan model olduğu görülmüştür.

KAYNAKLAR

- Bollerslev, T. (1986). “Generalized Autoregressive Conditional Heteroskedasticity”, *Journal of Econometrics*, Vol. 37: 307-327.
- Bozkurt, K., Pekmezci, A. ve Tekin, H.A. (2022). Box-Jenkins Yöntemiyle Turizm Talebinin Tahmin Edilmesi: Türkiye Örneği. *Anatolia: Turizm Araştırmaları Dergisi*, 33(2): 77-86.
- Çuhadar, M. (2006). Turizm Sektöründe Talep Tahmini için Yapay Sinir Ağları Kullanımı ve Diğer Yöntemlerle Karşılaştırmalı Analizi (Antalya İlinin Dış Turizm Talebinde Uygulama). Süleyman Demirel Üniversitesi Sosyal Bilimler Enstitüsü Doktora Tezi, Isparta, 1-189.
- Engle, R. F. (1982). “Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of the United Kingdom Inflation”, *Econometrica*, 50(4): 987-1007.
- Göktaş, P., Pekmezci, A. ve Bozkurt, K. (2019). Ekonometrik serilerde uzun dönem eşbütünlük ve kısa dönem nedensellik-makroekonomik verilerle Eviews ve Stata uygulamaları. Gazi Kitabevi.
- Karakaş, E. (2019). Türkiye’nin Otomotiv İhracat Gelirinin ARIMA Modeli ile Tahmin Edilmesi. *Journal of Yaşar University*, 14(55): 318-328.
- Özkoç H.H., Pekmezci A. (2021), *Finansta Güncel Konular*, “Finansal Ekonometride Değişen Varyans Modellerinin Tahminine İlişkin bir Uygulama” Nobel Akademik Yayıncılık, Ed: Erkan POYRAZ, Türkiye, 59-86.

Yapay Zekanın Kurumsal Dönüşümdeki Rolü

Ali Osman BİLGİN¹, Kudret Elif BERKMAN¹, Ahmet Akın ATASOY¹

¹TÜİK - Veri Yönetişimi Daire Başkanlığı, Ankara, Türkiye

Özet

Türkiye İstatistik Kurumu (TÜİK), ülkemizin istatistiksel verilerinin üretilmesi ve yaygınlaştırılmasında kritik bir role sahiptir. Yapay zeka teknolojilerinin hızla gelişmesiyle birlikte, büyük hacimli veri kümeleriyle uğraşan TÜİK için de yeni bir dönüşüm dönemi başlamıştır. Bu bildiri, yapay zeka alanında TÜİK'in rolü, kurumsal dönüşümdeki potansiyel faydaları, karşılaşılabilecek zorluklar, şimdiye kadar yapılan çalışmalar ve yapay zeka dönüşümünün nasıl gerçekleştirilebileceği ele alınmaktadır.

Araştırmanın amacı, TÜİK'in yapay zeka teknolojileri ile alakalı süreçlerini incelemek ve bu süreçlerin kurumun verimliliğine ve etkinliğine olan etkilerini değerlendirmektir. Çalışmamızda, TÜİK'in mevcut yapay zeka projelerini, karşılaştığı zorlukları ve ileri dönemdeki projelerini analiz eden nitel bir yaklaşım kullanılmıştır.

Temel sonuçlar, yapay zekanın TÜİK'e sunduğu fırsatları ortaya koymaktadır. Yapay zeka teknolojilerinin kullanılmasıyla veri toplama ve işlemede etkinlik sağlanırken, aynı zamanda tahmin ve modelleme yetenekleri sunulmaktadır. Tahmin edileceği üzere karmaşık veri kümelerinin etkili görselleştirilmesi, istatistiksel bilgilerin geniş kitlelere ulaştırılmasını kolaylaştırmaktadır. Ayrıca, rutin görevlerin otomasyonu ile insan kaynaklarının daha stratejik alanlara yönlendirilmesi mümkün olacaktır. Bununla birlikte, veri güvenliği ve gizliliği endişeleri ve gerekli altyapı yatırımları gibi zorlukların da aşılması gerektiği tespit edilmiştir.

TÜİK'in yapay zeka ekibi tarafından Açık İş İlanları İçin Mesleklerin Sınıflandırılması, Açık İş İlanlarından Beceri İlgörülere Çıkarılması ve E-fatura Verilerine CPA Kodlarının Atanması, Medya Takibin Yapay Zeka İle Otomatikleştirilmesi, Sınıflama Tahmin Portalı, Arama Terimleri İlgörülere,

Çevrimdışı Konuşmadan Metne Çözümleyici, Çevrimdışı Dil Modeli Asistanı, Bilgi Getirme Destekli Asistan projeleri yapılmaktadır. İlgili projeler, kurumun yapay zeka alanındaki bilgi birikimi ve tecrübesini artırırken, veri analizi ve sınıflandırma başta olmak üzere iş süreçlerini iyileştirmektedir.

Bu bildiri, sürdürülebilir dijital dönüşüm için yapay zekanın önemi vurgulanmakta ve TÜİK'in yapay zeka dönüşümünün mevcut durumunu, karşılaşılan zorlukları ve gelecek vizyonunu ortaya koyarak, kurumun ve ülkemizin istatistiksel altyapısının güçlendirilmesine katkıda bulunmayı amaçlamaktadır.

Anahtar Kelimeler: Yapay zeka, Dijital dönüşüm, Metin Sınıflandırma, Metin Madenciliği

GİRİŞ

Bu çalışma, TÜİK'in yapay zeka alanındaki süreçlerini analiz ederek, bu teknolojilerin kurumsal dönüşümdeki rolünü açıklamaktır. Yapay zeka, büyük veri analizi, dijital dönüşüm (iş süreçlerinin otomatikleştirilmesi) gibi alanlarda kullanılarak kurumsal performansı ve verimliliği artırmaktadır. Özellikle TÜİK gibi büyük veri hacimlerine sahip bir kurum, yapay zeka çözümleri sayesinde hem veri toplama ve işleme süreçlerini hızlandırmakta hem de elde edilen verilerin analizi ve raporlanmasını daha etkili bir hale getirmektedir. Bu çalışmada yapay zekanın TÜİK'teki mevcut ve gelecekteki potansiyel uygulamaları incelenerek, yapay zeka tabanlı dönüşümün kurum üzerindeki etkileri değerlendirilmiştir.

Araştırmanın temel amacı, yapay zeka uygulamalarının kuruma katkılarını ve bu süreçlerin verimliliğe olan etkilerini göstermektir. Bu süreçte karşılaşılan zorluklar ve yapay zekanın getirdiği fırsatlar da değerlendirilmiştir. Böylelikle yapay zekanın veri odaklı bir kurum olan TÜİK'e nasıl katkı sağladığı ve bu teknolojinin daha geniş bir çerçevede kullanılması için öneriler getirilmiştir.

ARAÇ VE YÖNTEMLER

Malzeme

Bu çalışmada kullanılan materyaller, TÜİK'in mevcut veri altyapısı ve yapay zeka projelerinde kullanılan araçlardır. Açık kaynaklı Doğal Dil İşleme (Natural Language Process - NLP) kütüphaneleri, derin öğrenme algoritmaları ve büyük dil modelleri bu çalışmada kullanılan ana araçlar arasında yer almaktadır. TÜİK'in mevcut teknolojik altyapısı üzerine inşa edilen bu materyaller, verilerin toplanması, işlenmesi ve analiz edilmesinde kullanılmıştır.

Verilerin Toplanması

Veri toplama sürecinde TÜİK'in Merkezi Dağıtım Sistemi (MEDAS) (TÜİK, 2024a), web kazıma teknolojileri, web servisler ve Uygulama Programlama Arayüzü (Application Programming Interface - API) bulunmaktadır. Kurumun yapay zeka çalışmaları kapsamında kullandığı veri kaynakları yapısal ve yapısal olmayan veri kümelerinden elde edilen bilgiler üzerine kurulmuştur. Aşağıda her çalışmaya özel kullanılan veriler anlatılmaktadır.

- Açık İş İlanları İçin Mesleklerin Sınıflandırılması ve Açık İş İlanlarından Beceri İlgörürleri Çıkarılması Projeleri: Her iki proje de aynı veri kaynağından beslenmektedir. Bu veriler, TÜİK'in kendi işgücü istatistiklerinin bulunduğu MEDAS, iş ilanlarının bulunduğu kamu kurumları ve özel sektör verileridir. Veriler her bir kurum ya da firmanın alt yapısına göre web kazıma, web servis, API ile alınmaktadır.
- E-fatura Verilerine Faaliyetlere Göre Ürün Sınıflandırması (Classification of Products by Activity - CPA) Kodlarının Atanması Projesi: Gelir İdaresi Başkanlığı (GİB) ile yapılan protokol çerçevesinde, e-arşiv ve e-fatura verileri, web servisleri aracılığıyla elde edilmektedir. Bu çalışmada e-fatura verileri kullanılmıştır.
- Medya Takibin Yapay Zeka İle Otomatikleştirilmesi Projesi: Yazılı ve görsel medya haberlerinden elde edilen veri setleri bir özel sektör firması tarafından derlenip kurum veri tabanlarına web servis ile aktarılmaktadır.
- Sınıflama Tahmin Portalı Projesi: TÜİK Sınıflama Sunucusundaki ürün, faaliyet ve meslek sınıflaması verileri kullanılmıştır (TÜİK, 2024b).
- Arama Terimleri İlgörürleri Projesi: TÜİK web sitesinde ziyaretçiler tarafından yapılan aramalara ilişkin veriler kullanılmıştır.
- Çevrimdışı Konuşmadan Metne Çözümleyici: TÜİK bünyesinde gerçekleştirilen ulusal ya da uluslararası toplantıların ses kayıtları ve ALO 124 TÜİK tarafından cevaplayıcılarla yapılan ses kayıtları kullanılmıştır.

İstatistiksel Analizler ve Yapay Zeka Teknikleri

Örneklem Büyüklüğünün Belirlenmesi ve Örnekleme Seçimi

Çalışmada örneklem büyüklüğünün belirlenmesinde ve seçiminde Cochran tarafından önerilen yöntemler kullanılmıştır. Örneklem seçiminde Cochran'ın önerdiği Basit Rastgele Örnekleme (BRÖ) yöntemi kullanılmıştır. Bu yöntem, kitle büyüklüğü, güven düzeyi, kitle oranı ve kabul edilebilir hata payı faktörlerini dikkate alarak minimum örneklem büyüklüğünü hesaplamaktadır (Cochran, 1977).

$$n = \frac{N \cdot z^2 \cdot p \cdot (1-p)}{d^2 \cdot (N-1) + z^2 \cdot p \cdot (1-p)}$$

N: Kitlede bulunan toplam birey sayısı

z: Güven düzeyine karşılık gelen z-skoru

P: Kitlede incelenen özelliğe sahip bireylerin oranı

d: Hata payı veya tolerans sınırı

Metin İşleme

Metin verilerinin işlenmesinde, Terim Frekansı ve Ters Doküman Frekansı (Term Frequency-Inverse Document Frequency - TF-IDF) vektörleştirme tekniği kullanılmıştır (Salton ve Buckley, 1988). Bu teknik, metin sınıflandırma ve bilgi çıkarımı görevlerinde hızlı sonuç gerektiren projelerde özellik çıkarımı için etkili bir yöntem olarak tercih edilmiştir. Bu teknik aşağıdaki şekilde hesaplanır:

- Terim Frekansı (Term Frequency -TF): Bir kelimenin bir belgedeki tüm kelimeler içindeki oranını ifade eder.

$$tf(t,d) = (\text{sayısı } t \text{ kelimesinin belge } d \text{ 'de}) / (\text{belge } d \text{ 'deki tüm kelime sayısı})$$

- Ters Doküman Frekansı (Inverse Document Frequency - IDF): Bir kelimenin tüm belge koleksiyonundaki nadirliğini ifade eder.

$$\text{idf}(t) = \log(\text{toplam belge sayısı} / t \text{ kelimesini içeren belge sayısı})$$

- TF-IDF: TF ve IDF değerlerinin çarpımıdır.

$$\text{tfidf}(t,d) = \text{tf}(t,d) * \text{idf}(t)$$

Sonuç olarak bu teknik ile bir kelimenin bir dokümandaki önemini, o kelimenin ilgili dokümanda geçme frekansı ile bütün dokümanlarda geçme yaygınlığı arasındaki ilişkiyi kullanarak belirler.

Derin Öğrenme Tabanlı Metin İşleme

Karmaşık metin işleme görevleri için, dönüştürücü (transformer) tabanlı doğal dil işleme tekniklerinden yararlanılmıştır. Cümleler İçin Dönüştürücülerden Gelen İki Yönlü Kodlayıcı Temsilleri (SentenceBidirectional Encoder Representations from Transformers - SBERT), yöntemi kullanılmıştır. Bu yöntem cümlelerin anlamsal ve bağlamsal özelliklerini yakalamak için dikkat (attention) mekanizmasını kullanan bir derin öğrenme modelidir (Reimers ve Gurevych, 2019). İlgili yöntemde iki farklı kelime, derin öğrenme yöntemi ile vektörize edildikten sonra kosinüs benzerliği ile uzaklıklarına bakılır (Singhal, 2001). Kosinüs uzaklıkları yakın olan vektörler anlamca birbirlerine daha yakın olur. Bu yaklaşım, metinler arasındaki semantik benzerliklerin etkili bir biçimde ölçer.

Sınıflandırma Analizleri

Sınıflandırma problemlerinin çözümünde literatürde sıklıkla tercih edilen dört yöntem olan Lojistik Regresyon (LR), Naive Bayes (NB), Rastgele Orman (RO) ve Destek Vektör Makinası (DVM) algoritmaları kullanılmıştır. LR, ikili sınıflandırma problemlerinde sonuç vermektedir. Kurumuzdaki bütün problemler ikiden fazla sınıfa sahip olduğu için "bire karşı hepsi" yöntemi ile genişletilerek çoklu sınıflandırma problemine uyarlanmıştır. NB, özelliklerin birbirinden bağımsız olduğu varsayımına dayanarak hızlı ve etkili bir sınıflandırma sağlamıştır. RO, birden fazla karar ağacını kullanan ve metin verilerinde başarılı olduğu düşünüldüğü için kullanılmıştır. DVM ise verileri en iyi ayıran hiper düzlemi bularak sınıflandırma yapmaktadır. Kompleks bir yöntem olduğu için kullanılmıştır. Bu yöntemlerin her birinin farklı avantaj ve dezavantajları bulunmaktadır. Sınıflandırma problemlerinde ilgili yöntemlerin hepsi kullanılmış başarı ve performansa göre uygun yöntem seçilmiştir (Han vd., 2011).

Eğitim Sonrası Niceleme (Post Training Quantization - PTQ)

Eğitim Sonrası Niceleme, bir modelin bellekte daha az yer kaplaması amacıyla sayısal temsilin daha düşük bitlerle yapılması işlemidir. PTQ, Büyük Dil Modellerinin (Large Language Models - LLM) maliyetini azaltmak için umut vadeden bir teknik olarak ortaya çıkmıştır. Özellikle, PTQ, LLM'lerde bellek tüketimini etkili bir şekilde azaltabilir ve hesaplama yükünü azaltabilir (Li vd., 2024).

Model Performans Değerlendirme

Model performanslarının değerlendirilmesinde, Doğruluk (Accuracy) ve F1-skoru metrikleri kullanılmıştır (Han vd., 2011). Bazı problemlerde kural bazlı yöntemler de uygulanmıştır. Kural bazlı yöntemlerde sadece doğruluk parametresi kullanılmıştır. Bu metrikler, farklı modellerin karşılaştırılmasında ve en uygun modelin seçilmesinde etkili olmuştur.

SONUÇLAR

Açık İş İlanları İçin Mesleklerin Sınıflandırılması

İş gücü piyasası analizlerinde açık iş ilanlarından elde edilen verilerin kullanımı, güncel beceri ihtiyaçlarını ve mesleki trendleri anlamak için giderek daha önemli hale gelmektedir (Colombo vd., 2019). TÜİK'in yürüttüğü "Açık İş İlanları İçin Mesleklerin Sınıflandırılması" projesi, bu alanda önemli bir adım teşkil etmektedir. Proje, ISCO-08 (Uluslararası Standart Meslek Sınıflaması) sınıflamasını kullanarak meslekleri sınıflandırmayı amaçlamaktadır.

ISCO-08, mesleklerin uluslararası düzeyde karşılaştırılabilir olmasını sağlayan ve işgücü piyasası araştırmalarında yaygın olarak kullanılan uluslararası bir standarttır (International Labour Organization, 2012).

Metin işlemek görevi için SBERT yöntemi kullanılmıştır. Ayrıca ISCO-08 sınıflaması oldukça geniş bir sınıflama olduğu için kural setleri de oluşturulmuştur. Sonuç olarak derin öğrenme ve kural setlerinin olduğu hibrit bir sistem geliştirilmiştir.

Etiketli veri bulunmamasından ötürü başarı oranını ölçmek için Cochran tarafından önerilen yöntem ile bir örneklem seçilmiştir (Cochran, 1977). Örneklemin çekildiği aşamadaki kitle büyüklüğü 42.225'dir. %95,00 güven düzeyinde, %50,00 kitle oranı ve % 0,05 hata payı ile örnek sayısı 381 olarak hesaplanmıştır. BRÖ ile seçilen 381 örnek için kurumdaki uzman personeller arasından 2 farklı kontrol grubu belirlenmiştir. Modelin sonuçları bu iki farklı grubun manuel yaptığı atama ile karşılaştırılarak %86,08 ve %86,82'lik başarı oranları elde edilmiştir.

Açık İş İlanlarından Beceri İçgörülerini Çıkarılması

Bu proje, Avrupa Beceriler, Yetkinlikler, Nitelikler ve Meslekler (European Skills/Competences, Qualifications and Occupations - ESCO) sınıflamasını kullanarak iş ilanlarından beceri kategorilerini çıkarmayı hedeflemektedir. Bu proje, işgücü piyasasındaki beceri taleplerini anlamak ve eğitim politikaları açısından büyük öneme sahiptir. Dijital, kodlama, yeşil, bilişsel, teknik, fiziksel ve sosyal becerilerin yanı sıra sertifikalar, dil becerileri ve deneyim taleplerinin analizi, işgücü piyasasının çok boyutlu bir resmini sunmaktadır.

Metin işlemek görevi için SBERT yöntemi kullanılmıştır. Ayrıca ESCO sınıflaması henüz geliştirilmekte olan bir sınıflamadır. Bu sebepten metin madenciliği teknikleri kullanılarak ilan açıklama alanlarından anlamlı bilgiler çıkarılmıştır. Sonuç olarak derin öğrenme ve metin madenciliği yöntemlerinin olduğu hibrit bir sistem kullanılmıştır.

Projede beceri çıkarma işlemi henüz tamamlanmadığı için başarı oranı ölçülmemiştir. Beceri çıkarma işlemi tamamlandıktan sonra ideal sayıda örnekler çekilerek konunun uzmanları tarafından yapılacak atamalar neticesinde başarı oranı ortaya çıkacaktır.

E-fatura Verilerine CPA Kodlarının Atanması

Bu çalışmada, Gelir İdaresi Başkanlığı tarafından dijital olarak tutulan e-fatura verilerine, CPA kodlarının atanması hedeflenmiştir. CPA kodları, ürünlerin ortak özelliklerine göre sınıflandırılmasını sağlar ve bu proje kapsamında E-fatura verilerine 6'lı düzeyde CPA kodları atanarak daha kapsamlı istatistiksel analizler yapılması amaçlanmıştır. CPA sınıflaması, toplamda 3.218 farklı sınıflama içerdiği için oldukça karmaşık çoklu sınıflandırma (multiclassification) problemidir. Bu denli geniş bir sınıflama yelpazesinde doğru eşleştirme sağlamak çok zordur.

Modelin girdi verisi olarak uzmanlar tarafından CPA kodları atanmış ürün açıklamalarının yer aldığı bir kaynak kullanılmıştır. Ancak, bu veri seti E-fatura verileriyle doğrudan ilişkilendirilmediği için temsiliyet açısından yetersiz kalmıştır. Temizleme işlemleri sonucunda, hedeflenen CPA kodlarının yalnızca bir kısmına ulaşılabilmiş ve ürün açıklamaları ile kodlar arasında beklenen eşleşme sağlanamamıştır.

Doğruluk oranı için uzmanlar tarafından yapılan CPA kod atamaları bile %22,00 başarı (accuracy) seviyesinde kalmıştır. Bunun en temel nedeni ürün açıklama verisinin CPA ataması için yeterli olmaması yani kapsam sorunudur. Kurulan modellerdeki en yüksek doğruluk oranı (accuracy) ise %5,60 olmuştur. Sonuç olarak hem kullanılan veri setinin yeterince temsil edici olmaması hem de karmaşık çoklu sınıflandırma problemi nedeniyle proje başarılı olamamıştır. Bu proje ile yapay zeka projelerinin hepsinde başarılı olunamayacağı kurumsal olarak idrak edilmiştir.

Medya Takibin Yapay Zeka İle Otomatikleştirilmesi

Bu proje, TÜİK'in bir firma ile ortaklaşa yürüttüğü medya takip sürecini yapay zeka kullanarak otomatikleştirmeyi amaçlamaktadır. Projenin amacı TÜİK hakkındaki tüm yazılı ve görsel medya haberlerinin daha verimli bir şekilde takip edilmesi ve analizlerin otomatikleştirilmesidir.

Projenin hedefi, haberlerin konusunu otomatik olarak tespit etmek, haber bültenini belirlemek, yazarını tanımlamak ve haberlerin olumlu veya olumsuz niteliklerini analiz etmektir. Bu hedefler doğrultusunda, manuel medya takibinin

zaman alıcı ve hata yapmaya açık yapısını iyileştirmek, insan kaynaklı hataları azaltmak ve genel verimliliği artırmak amaçlanmıştır.

Proje kapsamında, konu atama işlemi için TF-IDF ile birlikte NB, LR, DVM ve RO yöntemlerinin kullanıldığı ensemble yaklaşımı ile %83,74 doğruluk ve %73,58 F1-skoru elde edilirken, haber bülteni atamasında kural bazlı yöntemle %62,77 doğruluk sağlanmış, yazar tanıma ve duygu analizi bölümleri ise tamamlanamamıştır.

Sınıflama Tahmin Portalı

Ülkelerin ekonomik ve sosyal yapılarının incelenmesi için istatistiki bilgilere gereksinim duyulur.

Ancak bu bilgilerin gerek kullanıcılar arasında, gerekse uluslararası alanda anlamlı olabilmesi, belli standartlara dayandırılarak düzenlenmiş olmalarına bağlıdır. Bu bağlamda TÜİK sınıflama sunucu yıllardır ulusal ve uluslararası kullanıcılara değerli bir hizmet sunmaktadır (TÜİK, 2024b).

TÜİK'in mevcut sınıflama sunucusunda ilgili ürün, faaliyet ve meslek araması için geleneksel yöntem ile arama yapmak gerekmektedir. TÜİK tarafından geliştirilmekte olan "Sınıflama Tahmin Portalı" projesi ile yapay zeka destekli metin arama sistemi ile istenilen ürünün sınıflamasını bulma konusunda kullanıcılara daha etkin sonuçlar sunmayı hedeflemektedir.

Portal, metin verilerini SBERT yöntemi ile karşılaştırmaktadır. Proje kapsamında meslek sınıfı, faaliyet sınıfı ve amaca göre tüketim sınıfı tahminleme işlemleri gerçekleştirilmektedir. Henüz geliştirme aşamasında olan bu portalın geliştirilme ve test aşamalarının tamamlanması ile beraber ulusal anlamda önemli bir katma değer üretmesi düşünülmektedir.

Arama Terimleri İçgörülere

Bu proje kapsamında, TÜİK web sayfasında aranan kelimelerin ve kelime gruplarının çeşitli metin sınıflandırma yöntemleri kullanılarak sınıflandırılması amaçlanmaktadır. Elde edilecek analizler neticesinde kurumdaki mevcut arama motorunun geliştirilmesi için gerekli bilgilere sahip olunacaktır.

Ayrıca, TÜİK internet sitesi kullanıcılarının arama alışkanlıkları ve davranışları detaylı olarak incelenecektir. Bu sayede, kullanıcıların aradığı kelime ve kelime grupları yapay zeka destekli olarak kategorize edilerek, ana sayfa üzerinde daha kişiselleştirilmiş ve etkili bir hizmet sunulması hedeflenmektedir. Proje kapsamında, arama motorunun performansı artırılarak, kullanıcıların ihtiyaç duydukları bilgilere daha hızlı ve verimli şekilde ulaşması sağlanacaktır. Ayrıca kurum internet sitesinde aranılan kelime ve kelime gruplarının kategorize edilmesiyle beraber son kullanıcılara daha etkili hizmet verilmesinin sağlanması düşünülmektedir.

Proje, metinleri sayısallaştırma etmek için TF-IDF yöntemini kullanmıştır. Sınıflandırma algoritması için NB, LR, DVM, KA yöntemleri kullanılmıştır. Gerek başarı oranı gerekse performans kriterleri dikkate alındığında LR yöntemi ile en uygun model geliştirilmiştir. Yapılan analizler sonucunda doğruluk oranı %89,55 ve f1-skoru %89,21 olarak tespit edilmiştir.

Çevrimdışı Konuşmadan Metne Çözümleyici

Bu proje, çeşitli formatlardaki ses dosyalarını metne dönüştürmeyi amaçlayan bir uygulamadır. Kurum içindeki ses kayıtlarındaki veri gizliliği endişeleri için çevrimdışı bir çözüm olarak tasarlanmıştır. Kurum sunucularında çalışmaktadır. Bu sayede, ses kayıtları ve toplantılar yapay zeka destekli bir uygulama ile güvenli bir şekilde metin formatına dönüştürülmektedir.

Uygulama python programlama dili ile yapay zeka uygulamaları oluşturulmasını sağlayan Gradio kütüphanesi kullanılarak hazırlanmıştır (Abid, 2024). Uygulamanın arka planında OpenAI tarafından gerçekleştirilen açık kaynak kodlu Whisper kütüphanesi kullanılmaktadır (Radford vd., 2022). Ayrıca, bu projede, kurumda ilk kez bir model Grafik İşlemci Birimi (Graphics Processing Unit – GPU) üzerinde çalıştırılmıştır. Kullanılan GPU, RTX 4080 (16GB)'dir ve bu GPU üzerinde modelin hızlı ve oldukça iyi çalıştığı görülmüştür.

İlgili uygulama test aşamasında olduğu için henüz başarı oranları değerlendirilememiştir. İlerleyen süreçte kurumdaki toplantı ve telefon kayıtları için ideal örneklerle seçilerek başarı oranları belirlenecektir.

Çevrimdışı Dil Modeli Asistanı

TÜİK, veri gizliliği ve güvenliği konusundaki hassasiyetleri göz önünde bulundurarak, kendi sunucularında çalışacak bir "Çevrimdışı Dil Modeli Asistanı" geliştirmeyi hedeflemektedir. Bu projenin amacı, açık kaynaklı, çevrimdışı çalışabilen dil modeli kullanarak, kurumun ihtiyaçlarına özel bir yapay zeka asistanı oluşturmaktır. Çünkü mevcut durumda bireysel olarak kullanılan Gemini, Chat GPT, Claude.ai gibi platformlar kurumdaki veri gizliliği sebebiyle projeler içerisinde kullanılamamaktadır.

Bu projede, Meta'nın 8 milyar parametrelili Llama-3 modeli ile Google'ın 2 milyar parametrelili Gemma-

2 modeli üzerinde çeşitli deneyler yapılmıştır. Ancak, kullanılan GPU'ların (RTX 4080) bellek kapasitelerinin yetersiz kalması, modellerin tam performansla çalıştırılmamasına neden olmuştur. Bu sorunu çözmek amacıyla, başarıdaki azalma pahasına 4-bit ve 8-bit niceleme (quantization) yöntemleri uygulanmış, fakat bu yöntemlerle de beklenen performans elde edilememiştir. Sonuç olarak, modeller, Merkezi İşlem Birimi (Central Processing Unit – CPU) tabanlı bir ortamda (96 GB bellek, 2x48 GB) çalıştırılmak zorunda kalmıştır.

Konsept doğrulama (Proof Of Concept - PoC) işlemleri, mevcut altyapı kısıtlamaları nedeniyle tam kapasiteyle uygulanamamıştır. Buna rağmen, yapılan ön çalışmalar, projenin potansiyelini ve uygulanabilirliğini ortaya koymuştur. İlerleyen süreçte, kurumun bilişim altyapısının yapay zeka alanında güçlendirilmesi ve özelleştirilmiş dil modellerinin eğitimi için gerekli kaynakların sağlanması durumunda modelin tamamlanması düşünülmektedir. Bu sayede kurumun özel ihtiyaçlarına cevap verebilen, güvenli bir çevrimdışı dil modeli asistanının geliştirilmesi hedeflenmektedir.

Bilgi Getirme Destekli Asistan

Bilgi Getirme Destekli Asistan (Retrieval Augmented Generation - RAG) projesi, gelişmiş doğal dil işleme ve makine öğrenimi tekniklerini kullanarak, dinamik bilgi erişimi ve üretimi alanında yenilikçi bir çözüm sunmayı amaçlamaktadır. Bu sistem, büyük ölçekli veri tabanlarından semantik olarak ilgili bilgileri çıkarma ve bu bilgileri bağlamsal olarak uygun yanıtlar üretmek için kullanma yeteneğine sahip olması planlanmaktadır. RAG'ın temel amacı, geleneksel dil modellerinin statik bilgi sınırlamalarını aşarak, sürekli güncellenen ve genişleyen bir bilgi havuzundan yararlanmaktır.

TARTIŞMA VE SONUÇ

Yapay zekanın TÜİK'teki kurumsal dönüşümde oynadığı rol, hem operasyonel süreçleri iyileştirme hem de destekleyici sistemler açısından büyük önem taşımaktadır. Yapay zeka teknolojilerinin kurumdaki projelere entegrasyonu, özellikle sınıflandırma alanında hızlı ve doğru sonuçlar elde etmeyi mümkün kılmaktadır. Yapay zekanın kullanımının artması ile kurumdaki rutin yapılan işlerin azalacağı, böylece insan kaynaklarının daha stratejik işlere odaklanabileceği gösterilmiştir.

Yapay zekanın kullanılmasının artırılması sürecinde bazı zorluklarla da karşılaşılmaktadır. Yapay zeka uygulamalarının yaygınlaştırılması için nitelikli insan kaynağı ihtiyacı, güçlü altyapı, ulusal ve uluslararası standartlara uyum ve yüksek maliyetli yatırımlar gerekebilecektir. Veri güvenliği ve gizlilik konuları da diğer bir önemli meseldir. Özellikle yüksek hacimli veri işleme süreçlerinde, verilerin korunması ve izinsiz erişimlerin önlenmesi için şifreleme yöntemleri ve güvenlik protokollerine ihtiyaç duyulabilecektir. Ayrıca, yapay zeka uygulamalarının sürekli iyileştirilmesi ve doğru sonuçlar vermesi için geri bildirim mekanizmalarının kurulması gerektiği görülmüştür. Çünkü yapay zeka uygulamaları tek seferlik yapılan uygulamalar değil sürekli yaşayan canlı uygulamalardır.

Sonuç olarak, yapay zeka teknolojileri TÜİK'in kurumsal dönüşümünü hızlandırarak karar almada veri toplama, işleme ve analiz süreçlerini daha verimli hale getirmektedir. Yatırımlar ve güvenlik politikaları gibi zorlukların

üstesinden gelinmesine binaen elde edilen kazanımlar, yapay zekanın gelecekte TÜİK’te daha fazla kullanılacağını göstermektedir.

Referanslar

- Abid A, 2024. Gradio: Build and share delightful machine learning apps, all in Python. GitHub. Access address: <https://github.com/gradio-app/gradio>; Date of access: 02.09.2024.
- Cochran WG, 1977. Sampling techniques. 3rd ed., John Wiley & Sons, New York.
- Colombo E, Mercurio F, Mezzanzanica M, 2019. AI meets labor market: Exploring the link between automation and skills. Information Economics and Policy, 47: 27-37.
- Han J, Kamber M, Pei J, 2011. Data mining: concepts and techniques. Morgan Kaufmann Publishers Inc.
- International Labour Organization, 2012. International Standard Classification of Occupations: ISCO08. ILO, Geneva.
- Li S, Ning X, Wang L, Liu T, Shi X, Yan S, Dai G, Yang H, Wang Y, 2024. Evaluating Quantized Large Language Models. arXiv. Access address: <https://arxiv.org/pdf/2402.18158>; Date of access: 29.08.2024.
- Reimers N, Gurevych I, 2019. Sentence-BERT: Sentence embeddings using Siamese BERT-networks. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing (EMNLP), 3982-3992.
- Radford A, Wu J, Child R, Luan D, 2022. Whisper: A General-Purpose Speech Recognition Model. OpenAI.
- Salton G, Buckley C, 1988. Term-weighting approaches in automatic text retrieval. Information Processing & Management, 24(5): 513-523.
- Singhal A, 2001. Modern Information Retrieval: A Brief Overview. IEEE Data Engineering Bulletin, 24(4): 35-43.
- Türkiye İstatistik Kurumu (TÜİK), 2024a. Merkezi Dağıtım Sistemi (MEDAS). Access address: <https://biruni.tuik.gov.tr/medas/timeout.zul>; Erişim tarihi: 19.09.2024.
- Türkiye İstatistik Kurumu (TÜİK), 2024b. Sınıflama Sunucusu. Access address: <https://biruni.tuik.gov.tr/DIESS/>; Erişim tarihi: 20.09.2024.

Teşekkür

Kurum Başkanımız Dr. Erhan ÇETİNKAYA, Başkan Yardımcımız Tuğba DEĞİRMENCİ, Daire Başkanımız İsmail AKGÜL ve Grup Başkanımız Bilal KURBAN olmak üzere çalışmaya ait projelerin verilerinin ve altyapılarının sağlanmasında görev alan bütün TÜİK çalışanlarına ve çalışmalarımıza akademik olarak katkıda bulunan Dr. Ahmet KURNAZ’a teşekkür ederiz.

Çıkar çatışması

Yazarlar çıkar çatışması olmadığını beyan etmişlerdir.

Yazar Katkıları

Ali Osman BİLGİN, Açık İş İlanları İçin Mesleklerin Sınıflandırılması, Açık İş İlanlarından Beceri İçgörülerini Çıkarılması, Medya Takibin Yapay Zeka İle Otomatikleştirilmesi, Sınıflama Tahmin Portalı projelerinde görev olarak araç ve yöntemler ile sonuçlar kısmındaki ilgili bölümleri oluşturdu. Ayrıca giriş ile tartışma ve sonuç kısımlarını oluşturdu.

Kudret Elif BERKMAN, Çevrimdışı Konuşmadan Metne Çözümleyici, Çevrimdışı Dil Modeli Asistanı, Bilgi Getirme Destekli Asistan, E-fatura Verilerine CPA Kodlarının Atanması projelerinde görev olarak araç ve yöntemler ile sonuçlar kısmındaki ilgili bölümleri oluşturdu.

Ahmet Akın ATASOY, Arama Terimleri İçgörülerini projesinde görev olarak araç ve yöntemler ile sonuçlar kısmındaki ilgili bölümü oluşturdu.

Airline Business Model Evolution: A K-Means Clustering Analysis

Cem Ersöz^{1*}, Çağlar Üçler¹

¹ Özyeğin University, Faculty of Aviation and Aeronautical Sciences, İstanbul, Türkiye

Abstract

Traditional business models of airlines are evolving, driven by the dynamics of the aviation market. Full-service carriers started to apply economizing strategies, moving towards low-cost carriers. This paper claims that airlines apply hybrid business models to remain competitive, and statistical analyses were made here as proof. K-Means Clustering was applied to 10 major US airlines operating based on the MIT Airline Data Project to assess the 2010-2019 period for optimal cluster numbers determined by the elbow method. Associated variables were selected from the same database using the Delphi Method. Correlations were examined successfully. Nevertheless, Southwest was positioned in the same cluster as the full-service network carriers, which contradicts its business model. Similarly, Hawaiian and Alaska are in the low-cost carrier cluster but have business-class offerings or large aircraft operations. Consequent analysis with 3 clusters yielded further dynamic shifts among these airlines. In general, the analysis indicated evidence for hybrid approaches.

Key words: K-means Clustering, Elbow Method, Airline Business Model.

INTRODUCTION

Airline business models are evolving. Airlines must alter their services to match the new imperative. On the one hand, low-cost carriers (LCCs) compete beyond cost and price (Lawton and Solomko, 2005), trying to enhance quality through complementary to-be-sold services. On the other hand, full-service network carriers (FSNCs) leverage economizing initiatives. Sometimes, even an airline-within-airline strategy is used to achieve LCC advantages in FSNCs (Homsombat et al., 2014; Pearson and Merkert, 2014).

This implies that airlines adopt hybrid business models (Lohman and Koo, 2013; Pereira and Caetano, 2017). Nevertheless, the related literature is of a qualitative nature, and to the best knowledge of the authors there is no numerical explicit evidence for hybrid airlines. Consequently, this study aims to provide evidence for diminishing boundaries between low-cost and full-service carriers.

The structure of this paper is as follows: First, a short literature review is made for converging/ hybrid LLC and FSNC business models. Then, the sample, K-Means clustering leveraging Delphi, is explained as the method followed by the analysis and discussion of the data

Literature Review

Alamdari and Fagan (2005) demonstrated that contemporary LCCs diverged from the traditional LCC archetype. They identified Ryanair and Easyjet as the airlines that conformed most closely to the original LCC model, but even those two maintained only 79% conformance (Alamdari and Fagan, 2005, p. 384).

Tsoukalas et al. (2008) and Belobaba et al. (2009) further demonstrated that FSNCs diverge from the traditional FSNC archetype. Their studies also demonstrated a narrowing of unit costs between US-based LCCs and FSNCs between 1995 and 2006. This continued divergence from both strategic propositions is now collectively known as 'airline business model convergence' (Daft and Albers, 2013, 2015).

Traditionally, business model convergence represents a weakening in an airline's strategic position; however, convergence also has positive effects if it reflects the diffusion of efficient processes and practices among firms' (Daft and Albers, 2013, p. 47). Airline business model convergence has raised questions regarding the possible existence of an integrated cost leadership/differentiation strategy leading toward hybrid approaches (Moir and Lohmann, 2018).

Moreover, there is a “lack of a consistent and standardized approach to analyzing airline business models” (Mason and Morisson, 2008). Consequently, related variables, i.e., metrics, must be defined to compare airlines with each other.

MATERIAL AND METHODS

Material

MIT Airline Data Project database (MIT, 2024) was used as the data source. Data for 2010-2019 was utilized to prevent irregularities during the COVID-19 pandemic. The following 10 operational major airlines in the US market were investigated, as shown in Table 1, where the numbers associated with the company names have been used as identifiers in the statistical analyses and associated charts.

Table 1. Airlines analyzed

Airline Name	Business Model
American	Full-service carrier
Delta	Full-service carrier
United	Full-service carrier
Southwest	Low-cost carrier
jetBlue	Low-cost carrier
Alaska	Full-service carrier
Hawaiian	Full-service carrier
Allegiant	Low-cost carrier
Frontier	Low-cost carrier
Spirit	Low-cost carrier

Methods

The Collection of the Data

The following variables, which can reveal the business models of the airlines, were identified with the Delphi Method by 1 aeronautical engineer with 28 years of experience, 1 statistician with commercial air cargo and aviation management background of 16 years, 1 aviation management academician with 14 years of experience and leveraging the scientific work of Mason and Morisson (2008) and Lohman and Koo (2013):

- System Total Operating Revenue (\$ Billions)
- Total System Available Seat Miles (Millions)
- Total System Revenue Passenger Miles (millions)
- System Total Operating Expense (ex-Fuel and ex-Transport Related) (\$ Million)
- System Total Labor and Related Expenses (\$ Billions)
- Total Fuel Expense (\$ Million)
- Fuel Expense per ASM (\$US Dollars per Available Seat Mile)
- Total Other Outsourcing Expense (\$ Million)
- System Passenger Yield (Cents per Revenue Passenger Mile)
- Total System Load Factor (percentage)
- Average Daily Block Hour Utilization of Total Operating Fleet
- Average Daily Airborne Hours of Total Operating Fleet
- Average Stage Length Flown of Total Operating Fleet
- Average Seats per Departure (Total Operating Fleet)
- System Total Enplaned Passengers (thousands)
- Departed Seats (thousands)
- Revenue per passenger

Statistical Analysis

The elbow method was used to determine the optimal number of clusters. K – means clustering algorithm was employed to identify clusters. Major principal components (PC) were determined and plotted over each other for each year to examine the changes. Correlations of principal components with the variables for performance indication were computed and plotted.

RESULTS AND DISCUSSION

The elbow plot (Figure 1) shows a significant reduction in the within-cluster sum of squares (WSS) between 1 and 2 clusters, suggesting that adding a second cluster dramatically improves the clustering fit. After 2 clusters, the WSS reduction becomes less significant, so adding more clusters only yields minor improvements to the fit. Based on this pattern, 2 appears to be the optimal number of clusters.

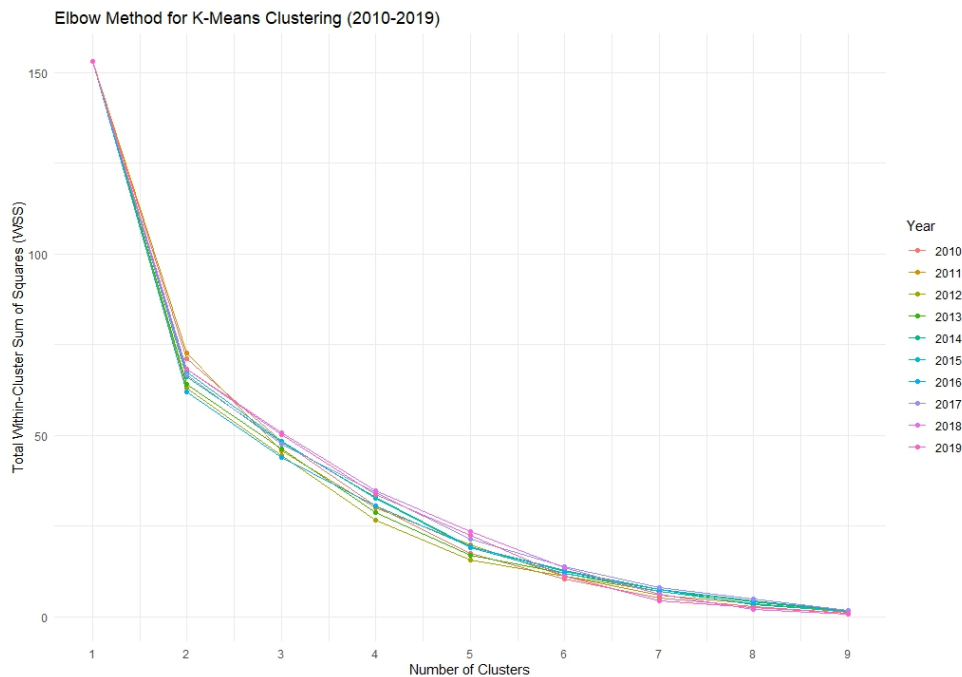


Figure 1. The result of the Elbow Method

Clustering the data into 2 yielded the first two principal components, as shown in Figure 2. K-means resulting from the data yielded similar results grouped in the same clusters. American, Delta, United, and Southwest formed cluster 1, i.e., full-service network carriers; JetBlue, Alaska, Hawaiian, Allegiant, Frontier, and Spirit formed cluster 2, i.e., low-cost carriers. This is an expected result, except Southwest, which is known as a low-cost carrier and Alaska and Hawaiian, which are known as full-service carriers. This can be based on the fact that Southwest's fleet size is very large and applies a mixed business strategy involving differentiated services. This embodies the proof of the initial claim that airlines are evolving their strategies towards a hybrid business model, diminishing boundaries between strategies.

Correlations of the variables with PC1 remain stable over the years in general, as shown in Figure 3. Only load-factor (LF) seems to experience large fluctuations, especially in 2015, which fuel prices and the related ticket price change can explain. PC1 captures the most significant variance in the data (60%-64%). PC1 is inversely related to airline financial metrics like Revenue, Expenses and capacity metrics like Available Seat Miles (ASM), Revenue Passenger Miles (RPM). Airlines with high cash flow and traffic contribute negatively to PC1, which suggests that the negative side of PC1 represents a scale or size factor that is strongly related to larger, full-service carriers. PC1 positively correlates with Load Factor, suggesting that airlines focusing on maximizing seat utilization and block hours (likely low-cost carriers) dominate the positive side of PC1.

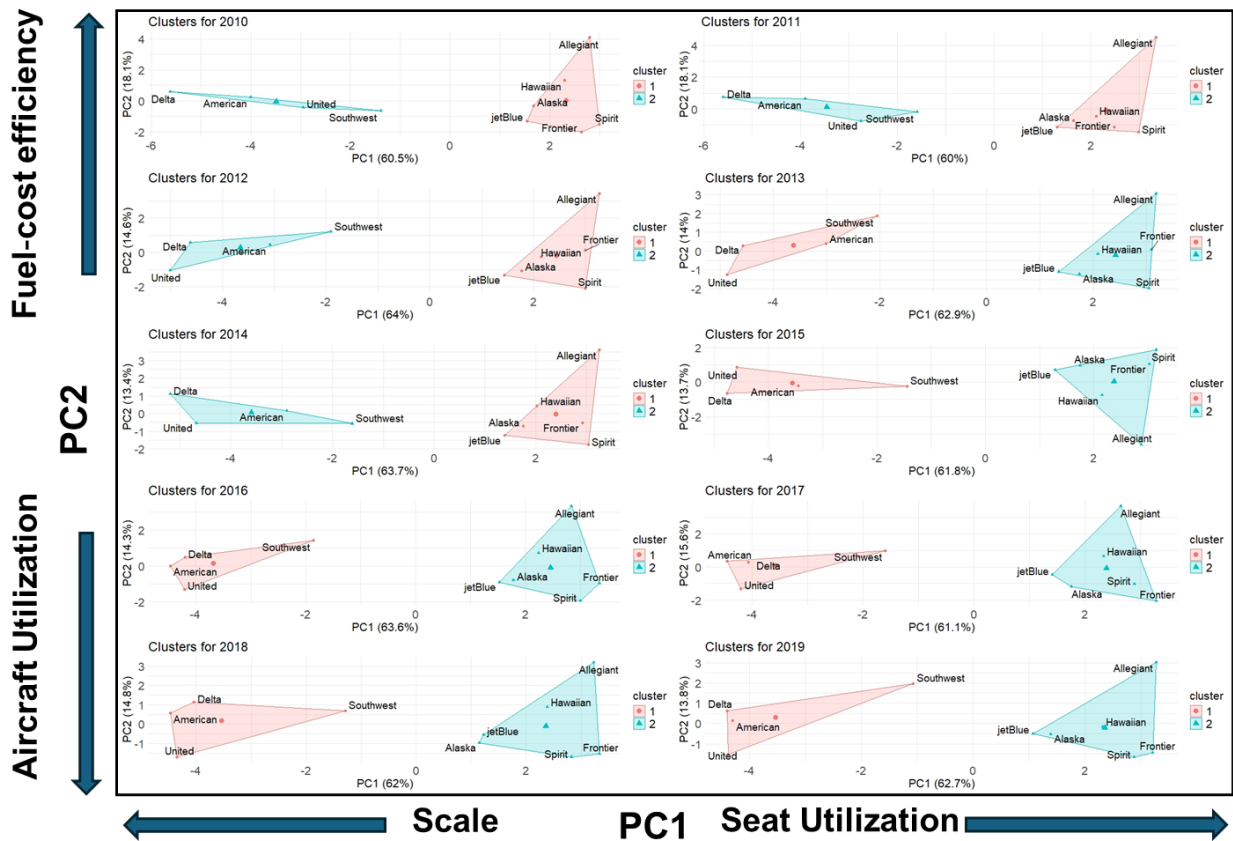


Figure 2. Results of the K-Means Clustering Method with 2 Clusters

Correlations of the variables with PC2 (Figure 4) indicate a major change in 2015. The correlations of aircraft and block hours changed polarity from negative to positive. Similarly, fuel expenditure per ASM, seats per departure, and load factor changed polarity from positive to negative. The next year, all correlations returned to their initial condition. This is again based on substantial decreases in fuel prices during 2015. PC2 primarily captures aircraft utilization time (flight and block hours) and fuel cost-efficiency (fuel expense per available seat mile). Airlines that are fuel-cost-efficient (i.e. spend less on unit fuel) and produce smaller daily flight times generally contribute positively to PC2 in most years. Fuel prices dropped significantly in 2015, affecting the relationship between operational cost-efficiency and airline performance. Longer flight times contributed positively to PC2 in 2015, while fuel efficiency became less important. PC2 reflects a balance between aircraft utilization (lower flight/block hours) and cost-efficiency, with 2015 marking a significant shift due to external factors (e.g., fuel prices).

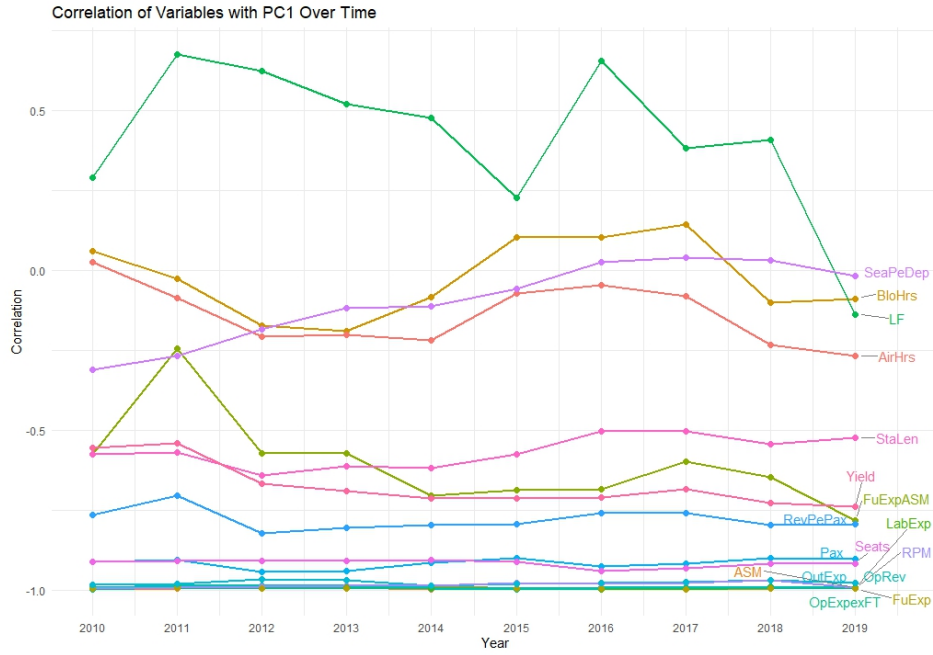


Figure 3. Correlations of variables with PC1 over years

To underline the hybrid strategy of some airlines, the same analyses were repeated with 3 clusters as shown in Figure 5. The motivation was whether further airlines are forming together with Southwest a new cluster for hybrid business strategy airlines. A third cluster might identify the airlines with emerging hybrid business models.

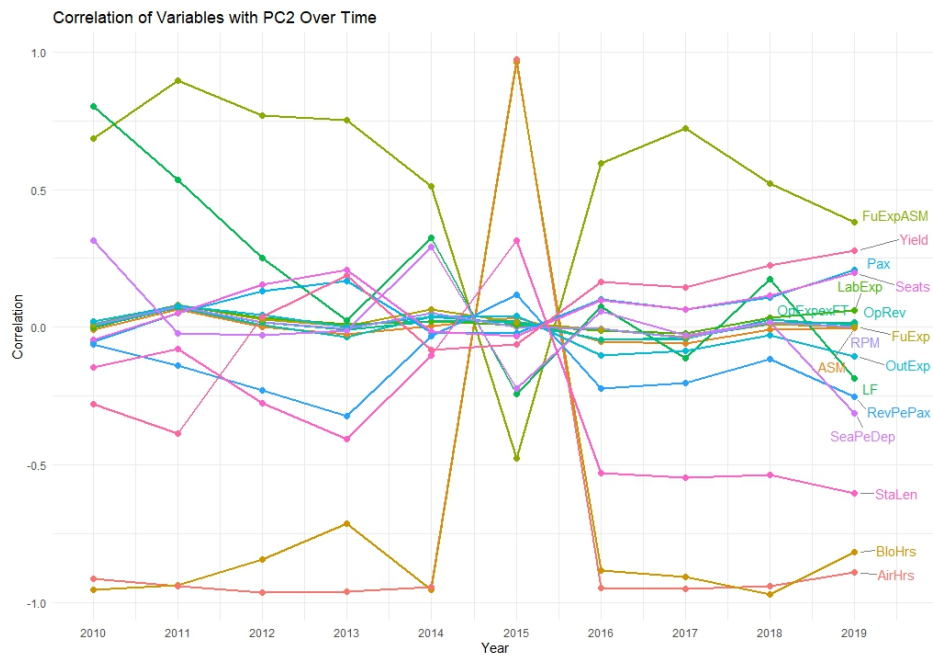


Figure 4. Correlations of variables with PC2 over years

Figure 5 identifies three distinct groups of airlines. Low-cost carriers such as Frontier, Spirit, and Allegiant consistently form one cluster, indicating their stable, cost-focused business model. Full-service carriers like American, Delta, and United predominantly cluster together, but airlines such as Alaska and Hawaiian fluctuate between clusters, suggesting they adopt hybrid strategies by incorporating cost-saving elements typically seen in low-cost models. In 2019, Southwest forms the third cluster in the middle of the other 2, which is evidence of a perfect blend of a FSC and a LCC.

The analysis suggests a degree of business model convergence over the years, particularly with some full-service carriers adopting hybrid strategies to remain competitive. While the clusters maintain a general distinction between low-cost and full-service models, the blurred lines seen in some airlines, such as jetBlue and Alaska, and Southwest indicate the pressure to combine cost-efficiency with service differentiation. This trend highlights the evolving nature of the airline industry, where carriers adapt their strategies to capture broader market segments while optimizing operational performance. This indicates the new paradigm shift for the hybrid model with shift of strategies (see Fig. 6).

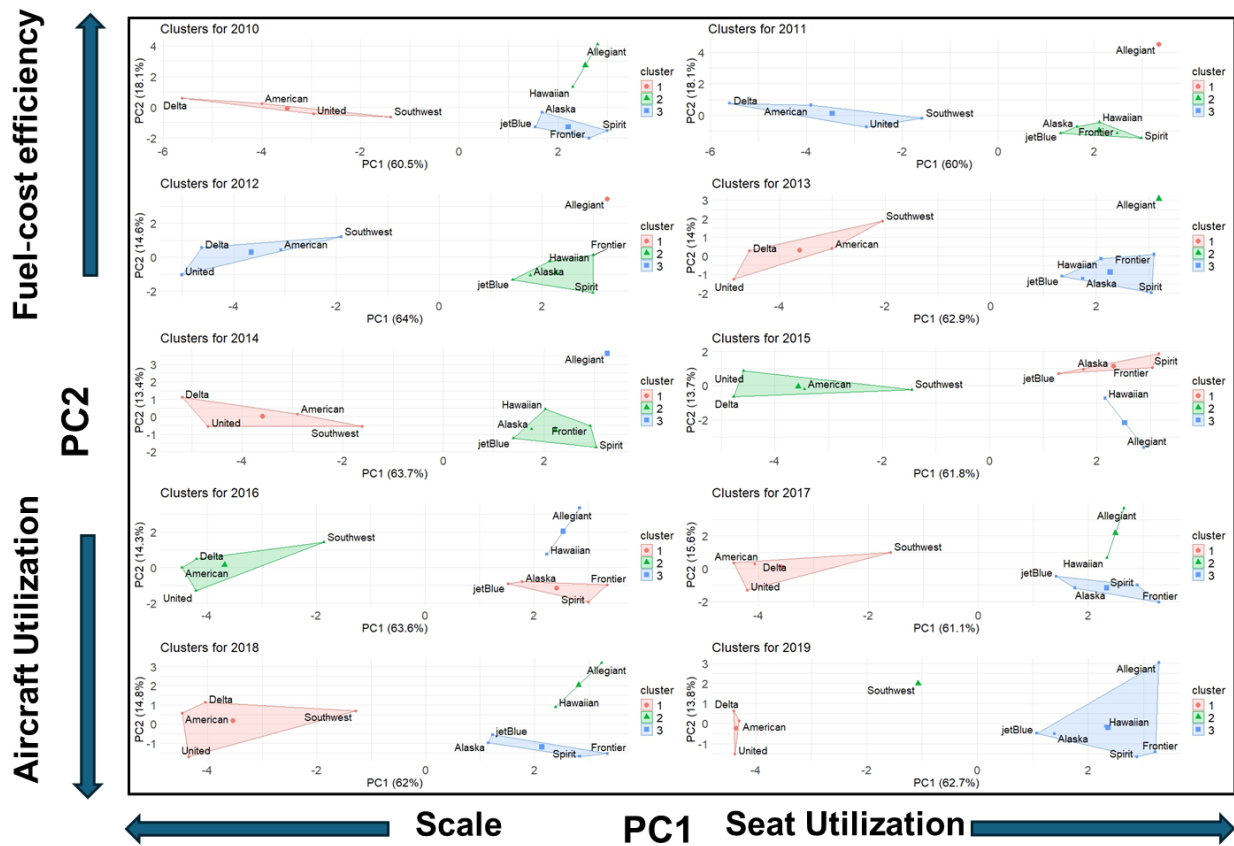


Figure 5. Results of the K-Means Clustering Method with 3 Clusters

The new imperative in the hybrid model implies that distinct differentiating strategies of LCCs are being adopted by FSCs and vice versa. A very distinctive change is e.g. the adaptation of low-cost subsidiaries by FSCs, which is evident in the sector. Then, as the numerical analysis indicates, the high seat utilization rate of LCCs is sought after by FSCs while LCCs started to offer on board entertainment and shifted towards business focus.

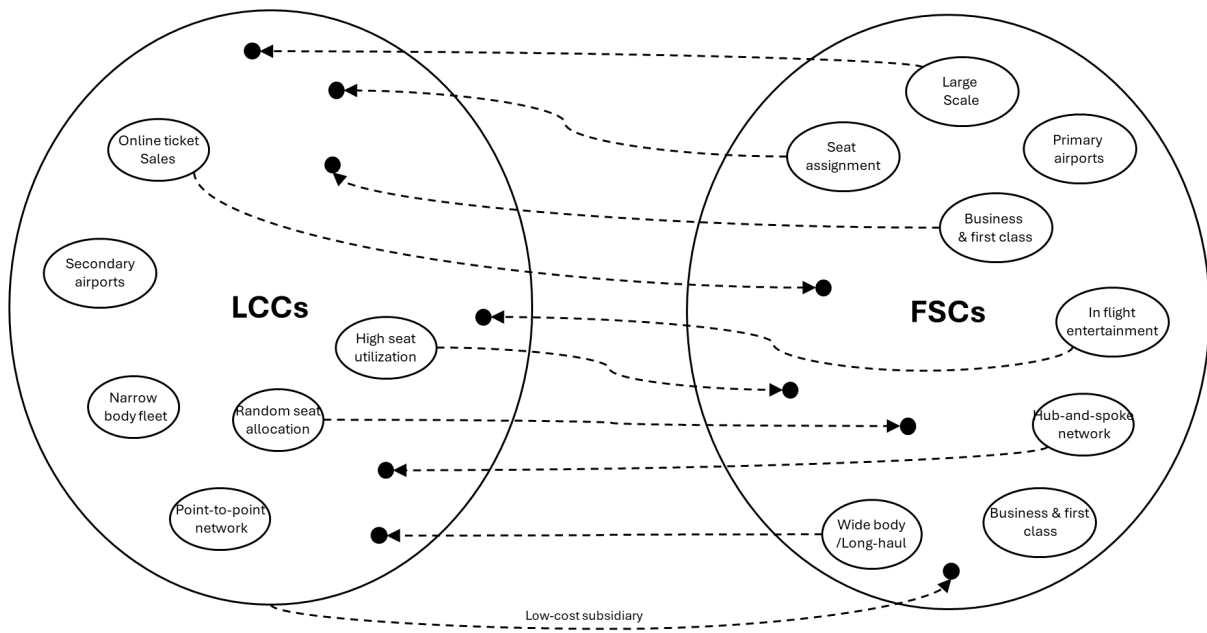


Figure 6. Hybrid business model with shift of strategies

CONCLUSION

K-means clustering was applied to US airlines between 2010 and 2019, which delivers evidence of a hybrid business model approach toward low-cost strategy.

The scientific contribution of this study is the application of the k-means method for the assessment of business strategies. It testifies the paradigm shift involving strategy adaptation towards hybrid business models. The managerial implication is that this work proves the need for the use of mixed strategies in full-service carriers based on historical evidence.

The study only covers the US airline market, which is a limitation. A future study on the worldwide market must be done.

References

- Abacı Alamdari F., Fagan S., 2005. Impact of the adherence to the original low-cost model on the profitability of low-cost airlines. *Transport Reviews*, 25 (3): 377-392. <https://doi.org/10.1080/0144164042000335798>.
- Belobaba P., Odoni A., Barnhart C., 2009. *The global airline industry*. John Wiley & Sons, USA.
- Daft J., Albers S., 2013. A conceptual framework for measuring airline business model convergence. *Journal of Air Transport Management*, 28: 47-54. <https://doi.org/10.1016/j.jairtraman.2012.12.011>.
- Daft J., Albers S., 2015. An empirical analysis of airline business model convergence. *Journal of Air Transport Management*, 46: 3-11. <https://doi.org/10.1016/j.jairtraman.2015.03.007>.
- Homsombat W., Lei Z., Fu X., 2014. Competitive effects of the airlines-within-airlines strategy - Pricing and route entry patterns. *Transportation Research Part E: Logistics and Transportation Review*, 63 (C): 1-16.
- Lohmann G., Koo T.T.R., 2013. The airline business model spectrum. *Journal of Air Transport Management*, 31: 7-9. <https://doi.org/10.1016/j.jairtraman.2012.10.005>.
- Lawton T.C., Solomko S., 2005. When being the lowest cost is not enough: Building a successful low-fare airline business model in Asia. *Journal of Air Transport Management*, 11 (6): 355-362.

Mason K.J., Morrison W.G., 2008. Towards a means of consistently comparing airline business models with an application to the 'low cost' airline sector. *Research in Transportation Economics*, 24 (1): 75-84.

MIT Airline Data Project, 2024. MIT Airline Data Project. Access address: <http://web.mit.edu/airlinedata/www/default.html>; Date of access: 12.05.2024.

Moir L., Lohmann G., 2018. A quantitative means of comparing competitive advantage among airlines with heterogeneous business models: Analysis of U.S. airlines. *Journal of Air Transport Management*, 69: 72-82. <https://doi.org/10.1016/j.jairtraman.2018.01.007>.

Pearson J., Merkert R., 2014. Airlines-within-airlines: A business model moving East. *Journal of Air Transport Management*, 38: 21-26.

Pereira B.A., Caetano M., 2017. Business model innovation in airlines. *International Journal of Innovation: IJI Journal*, 5 (2): 184-198.

Tsoukalas G., Belobaba P., Swelbar W., 2008. Cost convergence in the US airline industry: An analysis of unit costs 1995-2006. *Journal of Air Transport Management*, 14 (4): 179-187. <https://doi.org/10.1016/j.jairtraman.2008.04.003>.

TCMB Açık Veri Platformu: EVDS Geçmişi ve Geleceği

Avni Arıkan¹, Atakan Eren Ölmez^{1*}

¹TCMB/ Veri Yönetişimi ve İstatistik Genel Müdürlüğü, Ankara, Türkiye

Özet

Veri kütüphaneciliği günümüz dünyasında giderek artan kullanım hacmine sahiptir. Veriyi doğru bir şekilde toplamak, analiz etmek ve kullanmak, bireyler, kurumlar ve ülkeler için büyük önem arz etmektedir. Dünyada da ilk veri dağıtım örneklerinden olan Elektronik Veri Dağıtım Sistemi - EVDS , TCMB iç kaynaklarıyla oluşturulan yaklaşık 25.000 zaman serisi içermektedir. Hem finansal alanındaki profesyoneller tarafından , hem kamuoyunda hem de dünya çapında güvenilir finansal verilere ulaşmak için sıkça kullanılan bir kaynaktır. Bu çerçevede, bir açık veri platformu olan EVDS'nin doğru kullanımı, veriye dayalı karar alma süreçlerinde etkin rol oynamaktadır. TCMB olarak amacımız, geniş kullanıcı kitlelerine erişerek kaliteli, zamanlı, doğru ve güvenilir veri kullanım davranışlarına katkıda bulunmaktır.

Anahtar Kelimeler: Açık Veri, Veri Portalı, Veri Dağıtım, EVDS, Büyük Veri, İstatistik, Veri Analizi, Veri Madenciliği, Mobil Cihazlar

VERİ

Veri (İngilizce ve Latince datum, çoğulu data) İstatistik biliminde, yorumlanmak ve sunulmak amacı ile toplanmış, çözümlenmiş ve özetlenmiş gerçekleri ifade eder. Gözlem ve deneye dayalı araştırmanın sonuçlarıdır veri.

Bilgi ile veri 21. Yüzyılın petrolü olmuştur. İstatistiksel veriler, en önemli bilgi kaynakları arasında yer alır. Ülkeler için, istatistiki veriler güçlü ve zayıf yönetim noktalarını ortaya koyan toplum anatomisi sağlar. Böylece politika ve karar vericiler için önemli bilgi olur.

İstatistiksel gözlemleri veya ölçümleri temsil eden veriler, bilimsel teorileri doğrulamak veya çürütmek için de kullanılır. İstatistiksel veriler ürün satışları, piyasa gelişmeleri veya ekonomik göstergeler hakkında bilgi vererek, yönetimin stratejik kararlar almasını sağlar.

Devlet kurumlarının kendi aralarında ve kamuoyu ile veri paylaşımı her zaman önemi olmuş, yönetimde şeffaflığın göstergesi olarak kabul edilmiştir. Verinin paylaşılması kadar veriyi paylaşan kurumun statüsü de verinin güvenilirliğini artıran unsur olmuştur. Basılı raporlar, bültenler ile başlayan veri paylaşımı, bilişim teknolojilerinin gelişmesine paralel olarak paylaşım yöntemi değişikliklerine uğramıştır. Verinin biriktirilmesini sağlayan veri tabanı uygulamalarında gelişimin yanı sıra kişisel bilgisayarların birbirleriyle konuşmasını sağlayan yapılarda özellikle İnternet dünyasındaki gelişim, basılı veri paylaşımının, elektronik ortamda veri paylaşımına dönüşmesini sağlamıştır. Bunun doğal sonucu olarak verinin kullanıcılarına ulaşma zamanı nano saliseler seviyesine kadar inmiştir. Verinin sunum olanakları değişmiş, kullanıcının çok daha az sürede neredeyse hiçbir işlem yapmadan analitik, stratejik kararlar alabilmesini sağlayan veri paylaşım yöntemleri gelişmiştir.

Verinin bilgisayarlar aracılığı ile paylaşılması, sunma ve edinme maliyetlerini düşürmüştür. Süre inanılmaz boyutlarda kısalmıştır.

Verinin ilk elde edilmiş haline ham veri diyoruz. Birçok veri ham olarak bir anlam ifade etmez. Elde edilen verinin çeşitli yöntemlerden biri veya bazıları ile işlenmesiyle birlikte veri anlamlandırılabilir.

Ancak, bu verilerden anlamlı sonuçlar çıkarmak ve bu bulguları kararlar için girdi olarak kullanabilmek için istatistik bilimi hakkında bilgi sahibi olmak da gerekiyor.

İSTATİSTİĞİN AMAÇLARI VE KAYNAKLARI

İstatistik, geçmiş ve şimdiki durumla ilgili toplanmış verileri özetleyip yorumlayarak analiz etmemizi ve böylece gelecek hakkında karar vermemizi kolaylaştıran bir bilim dalıdır. Kısaca, geçmişin verilerinden faydalanarak geleceğe ilişkin belirsizliği ve karar almada riski azaltma aracıdır. İstatistik bilgileri önemli olmakla birlikte iyi

anlaşılmadığında yanlış yorumlanmaya da açıktır. İstatistikle yalan da söylenir. “Üç çeşit yalan vardır: Yalanlar, kuyruklu yalanlar ve istatistik – Benjamin Disraeli (Birleşik Krallık eski başbakanı)”

Veriden Bilgiye, Bilgiden Karara

Eldeki verilerin anlaşılabilir olmasını sağlamak ilk amaçtır.

Anlaşılabilir hale gelmiş veriler ile geleceğe yönelik karar alınmasını sağlamak üzere modelleme araçlarını kullanır. İkinci amaç karar vericilere belirsizliği azaltılmış bilgiler sunmaktır.

Yönetim kararları sadece tecrübeye ve sezgiye değil, aynı zamanda verilere de dayandırılmalıdır. Özellikle bilgi teknolojilerindeki gelişmeler, yönetimlerin piyasa hakkında birçok veriyi daha kolay toplayabilmesini ve kullanabilmesini sağlıyor. İstatistik biliminde en önemli yanlışlar veriler toplanırken yapılan örnekleme hatalarından kaynaklanır. Hatalı elde edilen veri yanlış bilgiler sunulmasına yol açar. Bunun da yanlış kararlara neden olacağı açıktır. Bu nedenle İstatistik biliminde veri toplanırken ve analiz edilirken, çeşitli doğrulama yöntemleri kullanılır.

Bilgi Oluşturma Kaynakları

Basit formüllerden karmaşık modellemelere kadar çok çeşitli yöntemler ile veri analiz edilebilir. Elde edilen bilgi tablo ve/veya grafik olarak kullanıcıya sunulur.

Bilginin kullanıcıya sunulması, başka bilimlerin katkıları ile sürekli gelişmektedir. Veriyi ve bilgiyi sunanlar zaman zaman aralarında yarışmaktadır. Sunumun bundan sonraki bölümünde bu konuyu ele alacağız.

AÇIK VERİ

“Bilginin herkes için ulaşılabilir olması gerekliliktir.”

Herhangi bir telif hakkı, patent ya da diğer kontrol mekanizmalarına tabi olmaksızın herkes tarafından ücretsiz ve özgürce kullanılabilen, tekrar kullanılabilen ve dağıtılabilen veriye açık veri denilmektedir.

- Açık verinin temel özellikleri: Kullanılabilirlik (erişim), tekrar kullanım (dağıtım), evrensel katılımın sağlanması.
- Veri hangi durumda açıktır: Teknik ve yasal olduğunda.
- Veri neden açılmalıdır: Şeffaflık, sosyal ve ticari değerlerin serbestleşmesi, katılımın sağlanması amaçlanıyorsa.

Her veri açık veri olma özelliğine sahip değildir. Devletin tüzel kişiliğinden veya kamu kurum ve kuruluşlarının hizmet sunum sırasında kendisinden kaynaklanan ya da üreterek sahip oldukları verilerin hassasiyetine bağlı olarak, kişisel verinin gizliliği ve mahremiyeti, kamu verisi güvenliği, ulusal güvenlik, ticari sır ve fikri mülkiyet konusu söz konusu olduğunda verinin açık olması mümkün değildir. Ülkeler bu konuda yasal düzenlemeler yapmaktadırlar.

Açık veri portalı nedir?

Açık Veri Projesi kapsamında hayata geçirilmesi hedeflenen Ulusal Açık Veri Portalı, vatandaşlarımıza ve bilim insanlarımıza anonimleştirilmiş ve mahremiyeti sağlanmış açık devlet verilerinin ortak bir açık veri portalı üzerinden sunulması ve bu bağlamda veriden değer üretmeyi hedefleyen veri paylaşım platformudur.

Açık veri platformlarının gelişmesinde internetin ucuzlayarak yaygınlaşması önemli yer tutar. Dünyanın hiçbir yeri artık ulaşamaz değildir. Bu durumun, küresel ekonominin temelini de oluşturduğu söylenebilir. Uydu iletişimi, sınırsız bilgi ağlarını birbirine bağlayan ve fiziksel mesafeleri sıfırlayan ana unsur haline geldi. İnternet, insanların her geçen gün gittikçe artan bilgi üretme, üretilen bilgiyi saklama/paylaşma ve ona kolayca ulaşma isteklerine ucuz ve pratik cevap geliştiren bir teknoloji oldu. İnternet yardımıyla pek çok alandaki bilgilere insanlar kolay, ucuz, hızlı ve güvenli bir şekilde erişebiliyorlar.

Veri dağıtım sistemleri, açık veri platformları ile kayıt ortamlarının, veri tabanlarının gelişim arasında doğrusal ilişki bulunmaktadır. Verinin dağıtılması kadar verinin toplanma hızının da elektronik teknolojiyle artmış olması

önemlidir. Çok değişik geniş bir yelpazeden toplanan t-2 frekansındaki verilerin kullanıma sunulması t-0,0 düzeyine indirilebilmiştir.

Veri Dağıtım İlkeleri

Avrupa Komisyonu İstatistik Program Komitesi (Statistical Programme Committee) Şubat 2005'te Avrupa İstatistikleri Uygulama Prensiplerini (Code of Practice) kabul ederek, bu prensiplere bağlı kalacağını beyan etmiştir. Uygulama prensipleri mevcut uluslararası standartlar temel alınarak oluşturulmuştur. 5429 sayılı Türkiye İstatistik Kanunu ile bu ilkelere uygun gerekli yasal düzenlemeler yapılmıştır. 5 yıllık dönemler için uygulanan Program da bu prensipler gözetilerek resmi istatistik üretim ve dağıtımını esas almaktadır.

TUİK nezdinde belirlenen temel prensipler şöyle sıralanabilir;

- Tarafsızlık ve şeffaflık: Resmi istatistik üreten kurum/kuruluşlar, verilerini, bilimsel bağımsızlık ilkesi çerçevesinde ve tüm kullanıcılara eşit mesafede, tarafsız, profesyonel ve şeffaf bir biçimde üretmeli ve dağıtmalıdır.
- Doğruluk ve Güvenilirlik: Resmi istatistikler konularına ve araştırma kapsamına göre, mevcut durumu ve değişimleri doğru ve güvenilir biçimde yansıtmalıdır.
- Tutarlılık ve Karşılaştırılabilirlik: Resmi istatistikler kendi içinde ve zaman içinde tutarlı, bölgeler ve ülkeler arasında karşılaştırılabilir olmalıdır.
- Resmi istatistikler, kurum ve kuruluşların web sayfalarından (html sayfaları, veri tabanları veya pdf dosyaları), basılı yayınlarla veya CD yoluyla yayımlanabilir. İstatistiklerin öncelikli yayımlama aracı web sayfalarıdır.

Merkez Bankası ve Açık Veri Platformu

TCMB'de 1990'lı yılların başına kadar, basılı bülten ve raporlar ile sürdürülen veri paylaşımı, kişisel bilgisayarların kullanımının artması, bilgisayarların birbirleriyle konuşması, teknolojinin gelişmesi ile birlikte kağıtlardan elektronik ortama taşınmıştır. 1993 yılında uygulamaya konulan ilk ürüne "Elektronik Veri Dağıtım Sistemi (EVDS)" adı verildi. EVDS, 1993 yılında karakter tabanlı bir uygulama olarak başladı. Bilgisayarlı Bilgi Sistemi (Bulletin Board System – BBS). Başlangıçta sadece 500 zaman serisinin sunumundan oluşuyordu. 1993 yılında Türkiye'de internet bağlantısının başlaması ile EVDS Tel-net uygulaması ile hizmet vermeye devam etti. Yine sunum karakter tabanlıydı. 1995 yılında internet kullanımının yaygınlaşması, grafik tabanlı işletim sistemlerinin kullanılmaya başlaması ile EVDS, yeni bir yüze kavuştu. 1988'e gelindiğinde, EVDS web tabanlı, grafik olanağı olan, yeni fonksiyonların eklendiği ve daha kullanıcı dostu bir sistem haline geldi. Zaman serisi sayısı ise 1800'e ulaştı. TCMB, dünyada kamu kurumları arasında, internet üzerinden açık veri platformu oluşturan ilkler arasındaydı. 2002 yılında, TÜSİAD tarafından, ekonomik göstergelerin tüm dünyaya ilk elden ve anında yayınlanması ve Türkiye ekonomisinin şeffaflığı açısından büyük önem taşıması nedeniyle EVDS'ye e-devlet ödülleri kapsamında jüri özel ödülü verildi.

2010'lu yıllar, veri dağıtımında, kullanıcıların da etkileşimli olarak raporlama sitemine katılmasını sağlayan uygulamaların geliştirildiği yıllar oldu. Veriyi kullananların, sunulan tablo ve grafikler üzerinde kendi ihtiyaçları doğrultusunda değişiklikler yapması ve sunumu kendilerine göre düzenlemesine olanak sağlayan uygulamalar geliştirilmeye başladı. 2016 yılında EVDS bu ihtiyaçları karşılamak üzere yeniden tasarlandı. Sunulan zaman serisi 28 000'i geçti.

Günümüzde EVDS, TCMB tarafından üretilen ve/veya diğer kuruluşlarca üretilip bünyesinde toplanan istatistiki verileri zaman serisi olarak, dinamik ve etkileşimli olarak sunan bir sistem olarak hizmet vermektedir. TCMB'nin dışarıya açılan yüzlerinden birisidir.

Ülke ekonomisinin izlenmesinde önemli bir kamu hizmeti görevi yürüten EVDS, birçok kesimden

kullanıcının ihtiyacını karşılıyor;

- Öğrenciler ödevleri için,
- Politika yapıcılar kararları için,

- Akademisyenler, arařtırmaları için,
- İř dnyası iř yařamında ihtiya duydukları veriler için,
- Ekonomi yazarları yazıları için,
- Yerli ve yabancı analistler ekonomimizi deęerlendirirken EVDS'den yararlanıyor.

Veri Daęıtımında, Gelecek

Son on yılda gerekleřen, mobil cihazlardaki geliřme, kablosuz baęlantılar, bnyk veri analizleri, ok boyutlu raporlar, birbiriyle yarıřan sinir aęları, yapay zekâyı herkes için mmykyn kılan bulut temelli hizmetler ve dięer geliřmelere bakıldıęında, bilim kurgu roman ve filmlerinin dnyasına hızlı bir gidiř olduęunu dnyřunmemek elde deęil. Masaüstü bilgisayar kullanımının yakın gelecekte olduka gdyk kalacaęı, mobil cihazların yanı sıra giyilebilir teknolojilerin olduka artacaęı dnyřunlmektedir. Veri daęıtımında kaliteli, gvenli ve hızlı olmanın yanında en bařarılı yorumların, gorsellięi bol mesajlar řeklinde iletileceęini dnyřunebiliriz. Alan, veri setleri olarak bilinen bnyk dijital koleksiyonları analiz etmek için veri tabanı ynyetimi ile istatistik ve yapay zekâ (sinir aęları ve makine oęrenme gibi) araların ok daha hızlı geliřeceęi slylenebilir. Donanım ve yazılımla ilgili geliřmeleri rnyn ve hizmetlerine en hızlı uyarlayan sektrylerden birisi bankacılıktır. Finans teknolojileri giderek onem kazanıyor. Fintech olarak adlandırılan bu alandaki geliřmeler, kurumlar ve bireylerin iř yapıř řekliyle yařamını daha derinden etkileyecektir.

EVDS'nin tm bu geliřmelerden etkilenmesi, yapay zeka ile geliřtirilmiř uygulamaları bnyyesine katması kaınılmaz olacaktır. Veri analitięi ile kalıpları ve oęrnyleri ortaya ıkararak yapay zekâ veri analitięi karar sryrelerine destek saęlayacaktır.

KAYNAKA

Ankara rnyversitesi, Kutyphane Dokymantasyon Daire Bařkanlıęı web sitesi (<http://kutuphane.ankara.edu.tr/>)

B.Beitler, A brief history of Microsoft Access

Bilgi Teknolojileri ve İletiřim Kurumu web sitesi (<http://www.mevzuat.gov.tr/Default.aspx>) History of Microsoft (<https://www.thoughtco.com/search?q=History+of+microsoft>) <https://en.wikipedia.org/>
<https://www.thoughtco.com/microsoft-history-of-a-computing-giant-1991140>

Koenig, G. The origin of FileMaker Pro

Martel, J. A personal history of dBase

Muهندس Beyinler web sitesi (<https://www.muهندسbeyinler.net/>)

The Linux Information Project. Database definition

Trkiye İstatistik Kurumu web sitesi (<http://www.tuik.gov.tr/>)

Wallerstein, Bildięimiz Dnyya Sonu, (ev: T. Birkan), Metis Yayınları, İstanbul.

Z.Beril Akıncı VURAL, Uęur BAKIR. "Distopyan Perspektiften Bilgi İletiřim Teknolojileri ve İnsanlıęın Geleceęi". Seluk rnyversitesi

