

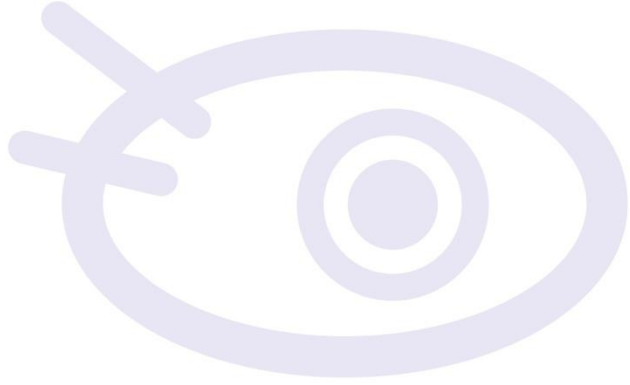
# Abstracts Book



TURKISH STATISTICAL ASSOCIATION

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

15-17 October | Ankara



**IDSSC 2024 ABSTRACTS BOOK**

**PUBLISHER:**

Turkish Statistical Association

**EDITOR:**

Prof. Dr. Turhan Menteş

**E-BOOK COVER and LAYOUT:**

Fatih Cengiz Öz [foztasarim.com]

**ISBN: 978-625-95613-1-8**

# 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

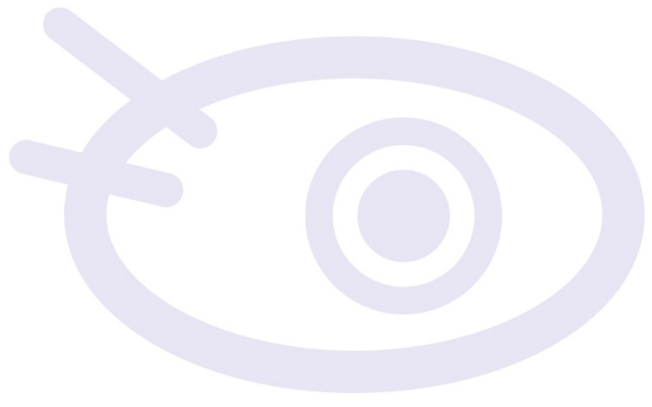


YAPAY ZEKA FABRİKASI

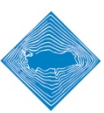


T.C. İÇİŞLERİ BAKANLIĞI

**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

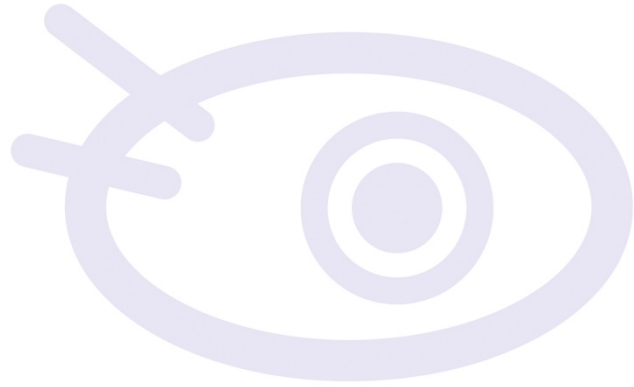
## CONGRESS SECRETARIAT

### Valör Kongre Organizasyonları

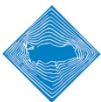
Turan Güneş Bulvarı Galip Erdem Caddesi 576. Sokak No.28 06550 Oran / Ankara  
T. +90 312 491 88 88 F. +90 312 491 99 89 M. vebist@valor.com.tr



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 teknolojiye 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 anlamamıza, 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.*”

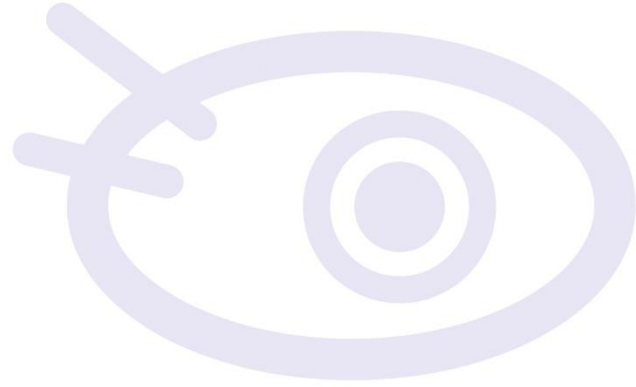
#### KAYNAKÇA

- 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](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  
2024



**Abstracts**



TURKISH STATISTICAL ASSOCIATION

## CONTENTS

Comparison of Methods Used in Predicting National Electricity Energy Consumption Amounts and an Application with United Kingdom Data.....	5
Fuzzy Generalized Additive Regression Functions for Time Series Forecasting.....	6
Feature Importance Analysis of Machine Learning Classifiers on Single-Cell RNA Sequencing Data.....	7
Yerel Yönetimlerde Uygulanmış Bir Araştırma Modeli: Ankara İli Örneği.....	8
Intuitionistic Fuzzy Cascade Forward Neural Network Regression Functions for Time Series Prediction.....	9
Computer Vision-Based Detection of Table Occupancy in Restaurants .....	10
Predicting Drought Levels in the Continental US Using Hybrid and Ensemble Machine Learning and Deep Learning Algorithms .....	11
İlaç Sektöründe Gelişmiş Talep Tahmini: Entegre Makine Öğrenmesi ve Derin Öğrenme Yaklaşımları .....	12
Discovering the Synergy Between Basketball Players Using Association Rule Learning .....	13
On Improving Computational Performance of Least Squares Multiple Birth SVM with Clustering and Genetic Algorithms .....	14
Comparative Analysis of LLM-Based and Conventional NLP-Based Methods for Multi-Class Text Classification in Turkish.....	15
Derin Öğrenme Yöntemleriyle Akarsu Debisi Tahmini .....	16
Türkiye'deki Boşanma Vakalarının Makine Öğrenimi ile Değerlendirilmesi: Bir Hayatta Kalım Analizi .....	17
An Overall Equipment Efficiency Predictive Analysis of a Hydraulic Press System by Time Series Forecasting with Topological Features.....	18
Sigortacılıkta Araç Sigortası Talep Sınıflandırmasında Gelişmiş Makine ve Derin Öğrenme Yöntemleri: Yenilikçi Yaklaşımlar ve Performans Değerlendirmeleri .....	19
IMDb Score Estimation using Movie Dialogues: A Text Mining and Machine Learning Hybrid Approach .....	20
Kuantum Hesaplama ve Bulanık Mantık ile İleri Seviye Makine Öğrenimi Yöntemleri .....	21
Artificial Neural Networks in Marketing: A Comprehensive Review of Consumer Behavior Applications .....	22

Comparative Analysis of Monthly Average Rainfall in Amasya Province with Exponential Smoothing Methods .....	24
Evaluating the Impact of Client Selection Algorithms on Performance and Training Speed in Heterogeneous Resource Allocation and Non-IID Data.....	25
SARIMA-YSA Hibrit Model Yaklaşımı: Türkiye’de Turizm Üzerine Bir Uygulama .....	26
Yapay Zeka Araçları ile Veri Analizine Genel Bir Bakış.....	27
Selecting Differentially Expressed Genes in Genomic Data Using the Length of ROC (L-ROC) Method.....	28
Analyzing and Improving the Impact of Software Changes on IT Operations: A Process Mining and Forecasting Approach .....	29
Türkiye’nin Ulusal Yapay Zeka Stratejisinin Kelime Bulutu (Wordclouds) Aracıyla İncelenmesi.....	30
High Dimensional Classification for Binary Features based on Decision Trees: An Empirical Study on ESG Reporting Companies in Türkiye.....	31
Zaman Serisi Tahmin Kombinasyonları İçin Çok Kriterli Karar Verme Yöntemine Dayalı Yeni Bir Yaklaşım .....	32
Hybrid Forecasting Approaches for Air Travel and Energy Demand: A Comparative Study Using ATA Method and Machine Learning Algorithms.....	33
Nesne Tanıma ve Sınıflandırmada Kullanılan Derin Öğrenme Yöntemleri .....	35
Bilimin İstatistiği: Bilimetri ve Türkiye'deki Bilimetri Çalışmalarının Bilimetrik Analizi .....	36
Classification of Brand Perception Using Random Forest: Brand Preference, Brand Loyalty, and Brand Trust.....	37
Yapay Sinir Ağları ile İstanbul Elektrik Talep Tahmini.....	38
Economic and Structural Changes in Cancer Drug Sales: Analysis of Turkey between 2010-2018.....	39
Eğitim İstatistiklerinde R ile Veri Kontrolünün Otomasyonu .....	40
Zaman Serisi Görüntüleri ve Yapay Zekâ Teknikleri İle Tarım Ürün Türü Sınıflaması.....	41
Elektronik Veri Araştırma Merkezi: Türkiye'de Veriye Dayalı Araştırmaların Güvenli ve Erişilebilir Platformu .....	42
Bölgesel Alt Kırılımlarda İnsani Gelişme Endeksi ve Bileşenleri, Metodolojik Yaklaşımlar .....	43

Veri Demokrasisi: Etkin Yönetişim ile Güçlenen Organizasyonlar .....	44
Integrating RFM Analysis with Power BI for Enhanced Customer Segmentation and CLV Insights.....	45
Avrupa İnsan Hakları Mahkemesi Yargı Kararlarının Metin Madenciliği Yöntemleriyle Tahmin Edilmesi.....	46
Forecasting nonlinear models: A simulation study.....	47
Sokak Hayvanlarına Yönelik Nefretin Duygu Analiziyle Tespiti .....	48
Rastgele Ormanlar ve Tekrarlayan Sinir Ağları Yöntemlerinin Teknoloji Endeksi Üzerinden Karşılaştırılmalı Analizi. ....	49
Birliktelik Kuralı Analizi ile Ürün Kod Atama Uygulaması .....	50
MEDOS: Kısa Dönemli İstatistiklerin Mevsimsellikten Arındırılması için Bir R-Shiny Uygulaması. ....	51
Airline Business Model Evolution: A K-Means Clustering Analysis.....	52
Yapay Zeka Destekli Veri Sınıflandırma ve Yönetim Modeli . ....	53
Covid-19 ve Deprem Süreçlerinde Uzaktan Eğitim: Öğrenci Beklentileri ve Deneyimlerinin Makine Öğrenimi ile Karşılaştırılması. ....	54
Enhancing Time Series Decomposition with Artificial Intelligence: Improving Pattern Identification and Interpretation in Economic Data .....	55
Estimation based on Ranked Set Sampling for the Kumaraswamy-Normal Distribution.....	56
On Estimating the Parameters of Logistic Distribution using Particle Swarm Optimization.....	57
Borsa İstanbul Bist100 Endeksi için Kurulan Modellerin Karşılaştırılması .....	58
Yapay Zekanın Kurumsal Dönüşümdeki Rolü. ....	59
Açık Bankacılık Platformu. ....	60
Panel Data Regression Analysis and Its Application to Real Data.....	61
İş İlanlarından Ötesi: Meslek ve Beceri Etiketleme ile Geleceğin İşgücünü Şekillendirmek .....	62
TCMB Açık Veri Platformu: EVDS Geçmişi ve Geleceği. ....	63
Mega Trendler, Mikro Analizler: Büyük Veri Yaklaşımları.....	64

## Comparison of Methods Used in Predicting National Electricity Energy Consumption Amounts and an Application with United Kingdom Data

Mehmet Akansel<sup>1</sup>, Nefise Ertoy<sup>1\*</sup>

<sup>1</sup>Bursa Uludag University, Department of Engineering, Faculty of Engineering, Bursa, Turkey

### Abstract

Electricity demand forecasting is vital for the safe, economical, and environmentally sustainable operation of energy systems. Recently, studies on electricity demand forecasting methods have focused on both improving classical methods and integrating new technologies. In this study, the aim is to forecast electricity demand using the electricity demand dataset of the United Kingdom. For this purpose, various forecasting models such as Seasonal Autoregressive Integrated Moving Average (SARIMA), Prophet, Long Short-Term Memory (LSTM), and XGBoost, Linear Trees were implemented using the Python programming language in the Colab environment. The performances of the models were compared using accuracy measurement methods such as Root Mean Square Error (RMSE) and Mean Absolute Percentage Error (MAPE). The results show the superiority of the LSTM method over other methods. Additionally, the results indicate that the Prophet and XGBoost models also perform well.

**Keywords:** Electricity Demand Forecasting; SARIMA, Prophet, LSTM; XGBoost; Linear Regression Trees

## Fuzzy Generalized Additive Regression Functions for Time Series Forecasting

**Damla İtter Fakhouri<sup>1\*</sup>, Ozge Cagcag Yolcu<sup>2</sup>, Ufuk Yolcu<sup>2</sup>**

<sup>1</sup>Department of Statistics, Mimar Sinan Fine Arts University, İstanbul, Türkiye

<sup>2</sup>Department of Statistics, Marmara University, İstanbul, Türkiye

### Abstract

Fuzzy logic is an approach used to model uncertainty and imprecise information. Fuzzy approaches are commonly used in time series forecasting (TSF). In TSF, fuzzy logic can be effectively utilized, particularly to handle uncertainties and manage complexities within the data. This can be suitable for managing complexity and uncertainty in time series data, as time series are often subject to noise, variability, and unexpected situations. Fuzzy logic-based TSF models can handle such scenarios in a more flexible manner and help make forecasts more realistic. Some of them perform rule-based principles to generate the predictions. On the other hand, particularly, fuzzy regression functions (FRFs) yield more effective outcomes due to their non-rule-based nature. However, FRFs create models with a linear structure to obtain forecasts and can model linear relationships between input and output. However, most time series often contain nonlinear relationships. At this point, using model structures capable of modeling nonlinear relationships instead of linear models will produce more effective forecasting results. Unlike traditional linear models, generalized additive models (GAMs) allow for more complex relationships by incorporating smoothing functions of explanatories. This characteristic makes GAMs particularly effective in scenarios where relationships may not be strictly linear and can adapt to capture nonlinear patterns in the data more accurately. In this study, fuzzy generalized additive regression functions (F-GARFs) approach is proposed for time series forecasting. In the mechanism of the proposed F-GARFs, firstly, lagged and real-valued time series observations are clustered using the fuzzy c-means (FCM) algorithm to generate memberships. Certain nonlinear transformations are then created for each membership value. Memberships and the corresponding transformations are individually constructed for each cluster alongside lagged real-valued time series, thereby constituting input sets for F-GARFs. The proposed F-GARFs have been applied to various real-world time series data. Forecasts generated by each F-GARF are weighted by corresponding membership values at the respective time points to obtain the final forecasts' results. Findings indicate that the proposed F-GARFs yield effective and reliable forecasting outcomes.

**Keywords:** Fuzzy Regression Functions, Generalized Additive Model, Fuzzy C - Means, Time Series, Forecasting

# Feature Importance Analysis of Machine Learning Classifiers on Single-Cell RNA Sequencing Data

**Bilge Özlüer Başer<sup>1\*</sup>**

<sup>1</sup>Department of Statistics, Mimar Sinan Fine Arts University, Istanbul, Turkey

## Abstract

Single-cell RNA sequencing (scRNA-seq) has emerged as a pivotal technology for profiling gene expression at single-cell resolution, offering unprecedented insights into cellular heterogeneity and complex biological systems. However, working with scRNA-seq data presents several significant challenges: high dimensionality, sparsity, noise and technical variability, batch effects, cell type identification, data integration, computational complexity, and scalability. scRNA-seq holds excellent potential in biological research but requires continuous innovation and improvement to overcome these challenges. This study systematically evaluates the performance and feature importance of various machine learning classifiers applied to scRNA-seq data, specifically aiming to distinguish between peripheral blood mononuclear cells (PBMCs) from a healthy donor stained with TotalSeq™-B antibodies and extranodal marginal zone B-cell tumor (MALT: Mucosa-Associated Lymphoid Tissue) cells stained with a panel of TotalSeq™-B antibodies. Both datasets consist of single-cell gene expression data processed by Cell Ranger 3.0.0 and sequenced on Illumina NovaSeq, using 10x Genomics technology. The goal is to gain insights into the underlying biological mechanisms and identify potential biomarkers for further biological and clinical research. The preprocessing pipeline involved quality control, filtering, normalization, logarithmic transformation, and dimensionality reduction, resulting in a compact, informative feature set for model training. Several classification algorithms were applied to the data, yielding high-performance results. Hyperparameter optimization was conducted with cross-validation to ensure robust model performance. The results demonstrate that, despite the complexity of the data, the classifiers consistently identified several critical genes, underscoring their relevance in differentiating between healthy and malignant cell populations. The feature importance analysis revealed a high degree of concordance across different models, highlighting the robustness of these machine learning approaches in biomarker discovery. This study emphasizes the efficacy of combining advanced preprocessing techniques with robust machine learning classifiers to analyze scRNA-seq data. The comprehensive feature importance analysis enhances our understanding of the underlying biological mechanisms and paves the way for more accurate and interpretable models in single-cell genomics.

**Keywords:** Single-Cell RNA Sequencing, Gene Expression, Machine Learning, Classification Algorithms, Biomarker Discovery.

## Yerel Yönetimlerde Uygulanmış Bir Araştırma Modeli: Ankara İli Örneği

**Dr. Filiz Aydın**<sup>1\*</sup>

<sup>1</sup>Kamu Yönetimi/ Hacettepe Üniversitesi, Ankara, Türkiye

### Özet

Bu bildiriye sosyal bilimlerin alt bilim dalı olan yerel yönetimler alanında nitel ve nicel 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 nelere dikkat edilmesi gerektiği ve alan çalışması yöntemlerine ilişkin genel tespitlerde bulunulacaktır. Tüm bu tespitler büyük ölçüde söz konusu doktora tezinde kullanılan alan araştırması yöntemi üzerinden ele alınacak ve benzer çalışmalarda kullanılan yöntemler de değerlendirilecektir. Bildiriye konu olan ve örnek alınacak olan doktora tezinin alan araştırması 2022 yılının Temmuz ayında gerçekleştirilmiş olup Ocak 2024 tarihinde kabul edilmiştir. Söz konusu alan araştırmasında nicel ve nitel olmak üzere iki farklı veri toplama tekniği kullanılmıştır. İlçeler 2012 yılı öncesinde kırsal alana sahip olanlar arasından; *ilçe nüfusu, kırsal alan nüfusu, ilçelerin alanı (km<sup>2</sup>), ilçelerde km<sup>2</sup> ye düşen nüfus, Ankara şehir merkezine mesafe (km), 2022 sosyo- ekonomik gelişmişlik sıralaması* skoruna (SEGE) göre 3 gruba ayrılmış ve bu 3 gruptan benzer ilçeler içinden rastgele 9 tanesi seçilmiştir. Ankara ilinin 25 ilçesinin 8’i (Altındağ, Çankaya, Etimesgut, Keçiören, Mamak, Pursaklar, Sincan ve Yenimahalle) 6360 sayılı Kanunun yasalaştığı 2012 tarihinde kırsal alanın bulunmaması nedeni ile kapsam dışında bırakılmıştır. Kalan 17 ilçe ise üç gruba ayrılarak içlerinde 9’u seçilmiştir. Ankara İlının 9 ilçesinde (Akyurt, Kazan, Polatlı, Ayaş, Çubuk, Nallıhan, Kalecik, Çamlıdere, Bala) gerçekleştirilen nicel araştırma da 380 kişi ile yüz yüze anket, nitel araştırma kısmında ise 19 mülakat yapılmıştır.

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



## Intuitionistic Fuzzy Cascade Forward Neural Network Regression Functions for Time Series Prediction

Asiye Zuhalt Baltaci<sup>1\*</sup>, Ozge Cagcag Yolcu<sup>2</sup>, Ufuk Yolcu<sup>2</sup>

<sup>1</sup>Big Data Coordination Office, Yalova University, Yalova, Türkiye

<sup>2</sup>Department of Statistics, Marmara University, Istanbul, Türkiye

### Abstract

Fuzzy set-based time series prediction models can be considered into two categories: rule-based and computation-based. While computational approaches like fuzzy regression functions outperform rule-based methods such as ANFIS, they have significant issues, including the limitation to model only linear relationships and the rigid assumptions inherent in the multiple linear regression models they employ for determining these relationships. It is a fact that models that determine relationships with neural networks with non-linear structures can only model non-linear relationships. Moreover, one of the main inadequacies of all these models is that they are insufficient to comprehensively address internal uncertainty as they ignore both the degrees of neutrality (hesitation) and non-memberships. At this point, although prediction models based on intuitionistic fuzzy sets have been proposed that take both of these phenomena into account, they can either model only linear or non-linear relationships. However, time series data sets, which are considered in almost all fields, contain both linear and non-linear relationships between inputs and outputs. This study introduces a new intuitionistic fuzzy time series prediction model. The proposed model uses cascade forward neural networks (CFNNs), which use memberships and non-memberships as separate input batches for each intuitionistic fuzzy set, as a regression function. Due to the unique architecture of the CFNN, each regression function is capable of modeling both linear and non-linear relationships simultaneously. The inputs for each intuitionistic fuzzy cascade forward neural network regression function (IF-CFNN-RF) are created via fuzzification of the real-valued time series with the intuitionistic fuzzy clustering algorithm and generating membership and non-membership values. Both membership and non-membership values and their certain non-linear transformations are combined separately with the lagged variables of the real-valued time series to create input batches for each IF-CFNN-RF. The predictive performance of the proposed IF-CFNN-RFs was comparatively demonstrated over various implementations in which widely used time series in the relevant literature. The findings confirm that the proposed IF-CFNN-RFs produced both effective and consistent/reliable prediction results.

**Keywords:** Intuitionistic Fuzzy Set, Cascade Forward Neural Network, Prediction, Time Series, Linear & Non-Linear Relationships, Intuitionistic Fuzzy Clustering.

## Computer Vision-Based Detection of Table Occupancy in Restaurants

Ali Kerem Güler<sup>1\*</sup>, Ali Musa<sup>1</sup>, Mehmet Göktürk<sup>2</sup>

<sup>1</sup>Protel A.Ş.

<sup>2</sup>Department of Computer Engineering, Gebze Technical University, Kocaeli, Türkiye

### Abstract

This study presents a new approach to monitoring and analyzing table occupancy in a restaurant setting using object recognition and human detection algorithms. The method involves creating a custom image dataset of tables of different colors, shapes, and sizes, and training a model on this dataset using the YOLO (You Look Only Once) algorithm. The system is designed to identify tables and calculate occupancy measurements based on the number of people detected in the relevant area around each table. In addition, information including table occupancy is recorded via logging in a time series dataset format to facilitate future operational planning and time-based analysis. Preliminary tests indicate that the methodology successfully detects and monitors table occupancy in real time and provides valuable information for restaurant management. By recording and analyzing this data, businesses can make informed operational decisions and improve their service quality to increase customer satisfaction.

**Keywords:** Object Detection, Image Processing, Table Occupancy Detection, Human Detection

## Predicting Drought Levels in the Continental US Using Hybrid and Ensemble Machine Learning and Deep Learning Algorithms

İrem Tanrıverdi<sup>1\*</sup>, İnci Batmaz<sup>2</sup>

<sup>1</sup>Department of Statistics, Middle East Technical University, Ankara, Turkey

<sup>2</sup>Department of Statistics, Middle East Technical University, Ankara, Turkey

### Abstract

The importance of drought prediction as a major environmental and socioeconomic challenge cannot be overstated. This research undertakes a comprehensive study of drought phenomena by combining two complex datasets detailing weather data and soil properties across several counties in the United States (US). We analyze weather data such as temperature, wind speed, humidity, precipitation and soil properties such as nutrient levels, rooting characteristics, oxygen content, salinity, toxicity levels, workability, slope across the US from 2000 to 2020, comparing algorithms that take into account spatio-temporal structures. Specifically, we evaluate Gradient Boosting Machines (GBM) techniques such as XGBoost, LightGBM, and CatBoost. We also develop deep learning models including Long-Short Term Movements (LSTM), Convolutional Neural Networks (CNN), and Transformer architectures. Hybrid models such as CNN-LSTM and Attention-LSTM are also explored. For ensemble methods, we implement stacking and voting classifiers to increase robustness and performance. When assessing drought in the US, it is essential to consider the varying characteristics and drought patterns across different regions. Each region has distinct climate and soil properties, which can significantly influence the performance of predictive models. Therefore, an effective approach is to first divide the data into clusters (representing different regions) and then perform separate modeling for each cluster. This strategy allows for the identification of the most suitable models for each region and the determination of which features are most significant for those models. By tailoring the models to the specific conditions of each region, we can enhance the accuracy and relevance of drought predictions and gain deeper insights into regional drought dynamics. To the best of our knowledge and based on a comprehensive review of existing literature, our study is pioneering in utilizing these advanced methodologies for an in-depth analysis of drought score data in the U.S. Our findings indicate model accuracies ranging from 0.5938 to 0.9783. The highest performance is achieved by the Attention-LSTM Hybrid Model, while the XGBoost exhibits the lowest performance.

**Key words:** Deep Learning, Machine Learning, Hybrid models, Ensemble methods, Clustering

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

Recep Yılıkcı<sup>1\*</sup>, Ece Özbilen<sup>1</sup>

<sup>1</sup> 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

## Discovering the Synergy Between Basketball Players Using Association Rule Learning

Erva Karayürek<sup>1\*</sup>, Emre Arslan<sup>2</sup>

<sup>1</sup>Metric Software and Consulting, İstanbul Turkey

<sup>2</sup>Metric Software and Consulting İstanbul, Turkey

### Abstract

Relationships between players can significantly impact the overall performance and game results. Analysis of the harmony between players is one of the most important key team performance indicators. Therefore, it is crucial to mine the compatibility between players for basketball decisionmakers to be able to identify the best-fitting players, thereby maximizing the performance of the team and players. In this study, Apriori algorithm is employed to discover the synergy between basketball players in terms of contribution to the points-assists and scores. The data used in this study is comprised of scorers and assists of each point of an American team in a season in the NBA. After performing analysis on the item sets generated by Apriori algorithm, associations are observed between specific players taking into account the confidence, ranging between 20% and 35%. Lift and conviction metrics yielded scores over 1, indicating positive associations and high reliability of the relationships, respectively. These findings provide insights for coaches, enabling them to select players based on the identified synergies to optimize team performance.

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

## On Improving Computational Performance of Least Squares Multiple Birth SVM with Clustering and Genetic Algorithms

Güvenç Arslan<sup>1\*</sup>

Department of Statistics, Kırıkkale University, Kırıkkale, Türkiye

### Abstract

Multiple Birth Support Vector Machines (MB-SVM) were introduced as a powerful extension of SVMs. Although the basic idea is similar as in SVMs, in this approach optimal non-parallel hyperplanes are used for each class category. Different implementations of this approach can be found in the literature. One of these is the Least Squares MB-SVM. An appealing property of this implementation is that an analytical solution is obtained, which is then used for classification. On the other hand, this solution involves matrix computations of sizes depending on the number of attributes and the size of the data set. In this study, we propose to use clustering and genetic algorithms before applying the Least Squares MB-SVM algorithm to improve the computational performance of MB-SVM.

**Keywords:** Support Vector Machine, Multiple Birth, Clustering, Genetic Algorithm

## Comparative Analysis of LLM-Based and Conventional NLP-Based Methods for Multi-Class Text Classification in Turkish

**Furkan Ayık<sup>1\*</sup>, Mehmet Yalçın<sup>2</sup>, Celal Akçelik<sup>3</sup> ve Serap Mungan Tanhan<sup>4</sup>**

<sup>1</sup>Turkish Technology, Istanbul, Türkiye

<sup>2</sup>Department of Statistics, Yıldız Technical University, Istanbul, Türkiye

<sup>3</sup>RDC Partner, Istanbul, Türkiye

<sup>4</sup>Turkish Technology, Istanbul, Türkiye

### Abstract

The growing volume of data sources and collected data has made extracting information from text data increasingly important, with text classification being a key focus. Natural Language Processing (NLP) models enable high-accuracy classification with less time and cost, facilitating quicker decision-making. The importance of these models has also risen in the Turkish language, but due to the morphological complexity of Turkish, it has been less explored compared to English. In this study, multi-class text classification models were developed using Turkish-language cabin report data from an airline company. The performance of Machine Learning - Boosting models was compared with Large Language Models (Llama, BERT, ConvBERT and XLM Roberta). Key text representation techniques (Fast-Text, Word2Vec, TF-IDF) were utilized for the Machine Learning models, while fine-tuning was applied to the LLM-based models. The BERT Base Uncased - Turkish model emerged as the most successful, achieving a 93% accuracy rate, demonstrating the superior performance of LLM-based methods in capturing Turkish nuances. Catboost, one of the Machine Learning models, achieved 86% accuracy. The study highlights the potential of LLMs in Turkish text classification while confirming the effectiveness of machine learning-based NLP methods, especially in resource-limited scenarios.

**Keywords:** Natural language processing, Large language model, Text classification, Machine learning, Bert

## Derin Öğrenme Yöntemleriyle Akarsu Debisi Tahmini

Mustafa Erden<sup>1</sup> ve Mehmet Seren Korkmaz<sup>1\*</sup>

<sup>1</sup>Meteoroloji Mühendisliği Bölümü, Samsun Üniversitesi, Samsun, Türkiye

### Özet

Tatlı su kaynakları, nüfus artışı ve iklim değişikliğinin etkisi ile birlikte hem insan hem de diğer canlılar için daha önemli hale gelmiştir. Bu kaynakların etkin bir şekilde değerlendirilmesi; içme suyu temini, tarım ve enerji üretimi için kritik rol oynamaktadır. Bu çalışmada; Samsun Terme Çayı'nın günlük debi tahmini için (Q1t), akarsu gözlem istasyonundan bir önceki güne ait debi değeri(Q1t-1), menbasındaki akarsu gözlem istasyonunda önceki güne ait gözlenen debi değeri(Q2t-1) ve havzada bir önceki güne ait gözlenen yağış değeri verileri (Pt-1) kullanılarak, derin öğrenme yöntemleri olan Tekrarlayan Sinir Ağları (RNN), Uzun Kısa Süreli Bellek (LSTM) ve Uyarlamalı Ağ Tabanlı Bulanık Çıkarım Sistemi (ANFIS) yaklaşımlarının performansları sınanmıştır. Her bir modeli değerlendirmek için RMSE, R<sup>2</sup>, NSE ve KGE performans metrikleri kullanılmıştır. RNN modeli düşük ve orta debi değerlerinde başarılı tahminler yaparken, yüksek debi değerlerinde tahmin hatalarının arttığı görülmüş, LSTM modeli; RNN'e göre daha iyi bir performans sergilemiş, ancak yüksek debi değerlerinde benzer sapmalar görülmüştür. En iyi performansı gösteren ANFIS modeli; düşük, orta ve yüksek debi değerlerinde tutarlı ve doğru tahminler yapmıştır. Korelasyon analizi, önceki dönem debi değerlerinin mevcut dönem debi değerlerini önemli ölçüde etkilediğini ortaya koymuştur. Ayrıca, Uzaktan Algılama tabanlı yağış veri setlerinin derin öğrenme yöntemlerinde kullanılabilmesi tespit edilmiştir.

**Anahtar Kelimeler:** Derin öğrenme, Debi, RNN, LSTM, ANFIS



## Türkiye’deki Boşanma Vakalarının Makine Öğrenimi ile Değerlendirilmesi: Bir Hayatta Kalım Analizi

**Gülçin Kendirkıran<sup>1\*</sup>, Şahika Gökmen<sup>2</sup>, Anıl Eralp<sup>3</sup>**

<sup>1,2</sup> Ekonometri Bölümü, Ankara Hacı Bayram Veli Üniversitesi, Ankara, Türkiye

<sup>3</sup> Ekonometri Bölümü, Bolu Abant İzzet Baysal Üniversitesi, Bolu, Türkiye

### Özet

Bu çalışma, Türkiye’deki boşanma üzerinde etkisi olan nitelikleri, makine öğrenme alanında kullanılan hayatta kalım analiziyle değerlendirmektedir. Hayatta kalım yöntemlerinde ikili sınıf etiketine sahip veri kümeleri bir süreç bilgisinin de varlığı ile derlenmekte ve bu sayede ikili sınıfın ortaya çıkış olasılıkları değil, ikili sınıfın ortaya çıkmasındaki süreç üzerinde etkisi olan faktörler analiz edilebilmektedir. Bu yöntem sağlık bilimleri ile ortaya çıkmakla birlikte sosyal bilimlerde de sıklıkla uygulanmaktadır. Yöntemin günümüz yapay zeka algoritmaları içerisinde ele alınışı özellikle ağaç tabanlı yöntemlerle yapılmaktadır. Bu yöntemlerin uygulanması aşamasında, gerçek verilerin ele alındığı örneklerde, ikili hedef değişken sınıflarındaki gözlem sıklıkları dengelenmiş veya dengelenmemiş olarak karşımıza çıkabilmektedir. Özellikle dengelenmemiş sınıfların yer aldığı veri kümelerinde, sınıf oranları arasındaki dengesizliğin görece fazla olduğu durumlarda model performansları zayıflamakta ve bu durum uygulamadan elde edilen sonuçlar için bir sorun oluşturabilmektedir. Makine öğrenimi yöntemleri, Türkiye Aile Yapısı Araştırması gibi büyük verilerin değerlendirilmesinin yanı sıra sözü edilen ikili sınıflardaki veri dengesizlikleriyle baş edebilme özelliğine de sahiptir. Bu aşamada, dengelenmemiş sınıf sorununa çözüm olarak aşırı örnekleme ve yetersiz örnekleme teknikleri güncel çalışmalarda önerilmektedir. Dolayısıyla söz konusu çalışmanın öncelikli amacı, Türkiye İstatistik Kurumu’nun hazırladığı 2021 Aile Yapısı Araştırması (AYA) verilerini kullanarak boşanma sürecinin hayatta kalım makine öğrenmesi algoritmasıyla incelenmesi ve bu sayede yapay zekayla gelişen ileri tekniklerin sosyal bilimlerdeki yerinin boşanma gibi önemli bir konuyla değerlendirilmesidir. Bu doğrultuda çalışmada ikincil amaç olarak, aşırı örnekleme ve yetersiz örnekleme tekniklerinin ele alınarak uygun hiper-parametrelerin de belirlenmesiyle elde edilen farklı Rassal Hayatta Kalım Ağaçları’nın performanslarının incelenmesini kapsamaktadır. Ardından, karşılaştırılan performanslara bağlı olarak boşanma üzerinde etkili olan niteliklerin belirlenmesindeki değişkenlerin önem sıralamaları boşanma konusu özelinde değerlendirilmektedir. Bu bağlamda, çalışmadan elde edilen temel bulgulara göre aşırı örnekleme tekniği kullanıldığında model performansının yaklaşık olarak %90’a ulaşarak, diğer yaklaşımlardan daha iyi performans sergilediği gözlenmiş ve söz konusu model referans alınarak model seçim sonuçları tartışılmıştır.

**Anahtar Kelimeler:** Hayatta Kalım Analizi, Time-to-event, Değişken Önem Sıralaması, Model Seçimi.

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

**Korkut Anapa<sup>1\*</sup>, İsmail Güzel<sup>2</sup>, Ceylan Talu Yozgatlıgil<sup>3</sup>**

<sup>1,2</sup>Institute of Applied Mathematics, Middle East Technical University, Ankara, Turkey

<sup>2</sup>Network Technologies Department, TÜBİTAK ULAKBİM, Ankara, Turkey

<sup>3</sup>Department 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

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

Volkan OBAN<sup>1\*</sup>, Bulut Karadağ<sup>1</sup>, Gökhan Görmüş<sup>1</sup>, Ashlan Çandır<sup>1</sup>

<sup>1</sup>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.

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

Yunus EROĞLU<sup>1\*</sup>

<sup>1</sup>Industrial Engineering, Engineering Faculty, Gaziantep University, Gaziantep, Turkiye

### 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.

**Keywords:** IMDb score, Text mining, Machine learning, Estimation

## Kuantum Hesaplama ve Bulanık Mantık ile İleri Seviye Makine Öğrenimi Yöntemleri

Cumali Yaşar<sup>1\*</sup>, İhsan Yılmaz<sup>1</sup>

Eğitim Fakültesi, Çanakkale Onsekiz Mart Üniversitesi, Çanakkale, Türkiye  
Bilgisayar Mühendisliği Bölümü, Çanakkale Onsekiz Mart Üniversitesi, Çanakkale, Türkiye

### Özet

Makine öğrenimi, büyük veri analizi ve yapay zeka alanlarında hızla gelişmekte olan bir disiplindir. Ancak veri setlerinin karmaşıklığı ve belirsizlikler, mevcut makine öğrenimi tekniklerinin etkinliğini sınırlandırabilir. Bu nedenle, kuantum hesaplama ve bulanık mantık gibi yeni teknolojiler, makine öğrenimi süreçlerinde daha güçlü ve esnek çözümler sunmak için önemli fırsatlar yaratmaktadır. Bu çalışmada, kuantum hesaplama ve bulanık mantığın bir araya getirilerek ileri seviye makine öğrenimi yöntemlerinin nasıl geliştirilebileceği incelenmiştir. Kuantum hesaplama, klasik bilgisayarların sınırlarını aşarak kubitler yardımıyla çok daha hızlı ve karmaşık hesaplamalar yapabilme kapasitesine sahiptir. Paralel hesaplama gücü sayesinde, kuantum bilgisayarlar çok boyutlu veri setlerinde işlem süresini büyük ölçüde azaltarak makine öğrenimi algoritmalarının verimliliğini artırabilir. Kuantum algoritmaları özellikle optimizasyon ve veri sınıflandırma süreçlerinde dikkate değer performans artışları sağlayabilir. Bulanık mantık ise, verilerin kesin kategorilere ayrılması yerine belirsizliklerle başa çıkabilme yeteneği sunmaktadır. Geleneksel makine öğrenimi algoritmaları çoğunlukla kesin sınırlar gerektirirken, bulanık mantık, her veri noktasının bir kategoriye belirli bir derecede ait olmasına olanak tanır. Bu esneklik, özellikle belirsiz ve karmaşık veri kümelerinde daha anlamlı sonuçlar elde edilmesini sağlar. Bu çalışmada, bulanık mantık ile kuantum hesaplamanın entegre edilmesiyle oluşturulan hibrit bir makine öğrenimi modeli sunulmaktadır. Önerilen modelin performansı, geleneksel makine öğrenimi yöntemleri ile karşılaştırıldığında hem hız hem de doğruluk açısından üstünlük sağlamaktadır. Kuantum hesaplamanın paralel işlem gücü, bulanık mantığın belirsizliklerle başa çıkma kapasitesi ile birleştiğinde, makine öğrenimi algoritmaları daha hızlı ve güvenilir sonuçlar elde etmektedir. Bu yaklaşım, özellikle büyük veri setleri üzerinde çalışan sistemlerde önemli bir avantaja sahiptir. Çalışmanın sonunda, hibrit modelin sağlık, finans ve mühendislik gibi belirsizliklerle sıkça karşılaşılan alanlardaki potansiyel uygulamaları üzerinde durulmuştur. Ayrıca, veri güvenliği ve büyük veri yönetimi gibi konularda da bu modelin sağlayabileceği avantajlar tartışılmıştır. Gelecekte, kuantum hesaplama ve bulanık mantığın daha fazla entegre edilmesi, makine öğrenimi algoritmalarının sınırlarını daha da ileri taşıyacak ve çok daha karmaşık problemlerin çözülmesine olanak sağlayacak şekilde uygulaması geliştirilmiştir.

**Anahtar Kelimeler:** Kuantum Hesaplama, Bulanık Mantık, Makine Öğrenimi, Veri Bilimi, Büyük Veri Analizi, Yapay Zeka, Belirsizlik Yönetimi

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

**İbrahim Halil Efendioğlu<sup>1\*</sup>**

<sup>1</sup>Business Administration, Gaziantep University, Gaziantep, Turkiye

## Abstract

This paper aims to systematically examine how artificial neural networks (ANNs) are used to predict consumer behavior in the marketing field and to evaluate the existing literature in this area. By systematically reviewing the literature between 2020 and 2024, it explores the advantages, applications, and challenges of ANNs in predicting consumer behavior. The review highlights ANNs' potential to improve marketing strategies through more accurate predictions, dynamic customer segmentation, and optimization of personalized marketing efforts. Despite their effectiveness, challenges such as the "black box" nature of ANNs and data requirements are identified as areas needing further exploration. Artificial neural networks (ANNs), as part of artificial intelligence (AI), have emerged as potent tools for analyzing and predicting consumer behavior in marketing. ANNs offer substantial benefits by processing complex and large-scale datasets, leading to more precise marketing strategies. They enable businesses to predict consumer preferences, personalize recommendations, and optimize marketing campaigns. Despite increasing applications of ANNs in consumer behavior, there still needs to be a gap in synthesizing the literature to understand their capabilities and limitations better. This paper aims to review recent studies on ANNs and consumer behavior, offering insights into their application and providing recommendations for future research. The literature on ANNs shows rapid growth in their application to consumer behavior, primarily in consumer segmentation, purchase prediction, and marketing strategy optimization. ANNs are uniquely positioned to model nonlinear and complex relationships, making them more effective than traditional statistical methods like regression analysis. They have been applied across various marketing functions, such as customer segmentation, sales forecasting, and personalized marketing, by analyzing past consumer behavior and predicting future trends. Studies reveal that ANN-based approaches outperform conventional methods in accuracy, particularly in consumer behavior prediction. They allow companies to optimize marketing campaigns by targeting distinct consumer segments with tailored messages, leading to improved customer engagement and higher conversion rates. This research utilized a systematic literature review (SLR) to analyze publications from 2020 to 2024. Articles were selected from the Scopus database using keywords like "consumer behavior" and "artificial neural networks." The selection criteria focused on peer-reviewed journal articles in marketing-related fields. Following the PRISMA guidelines, the SLR method ensured a structured approach, filtering articles based on their relevance to marketing and consumer behavior. The review identified 25 critical studies focusing on ANN applications in marketing, showing a steady increase in publications from 2020 to 2024. The studies predominantly applied empirical methodologies, most utilizing ANN for predictive modeling. Key trends include the rise of dynamic customer segmentation, personalized recommendations, and demand forecasting. Several studies showed that ANNs could model complex and nonlinear relationships, offering higher accuracy in predicting consumer behavior than traditional techniques. However, challenges such as the need for large datasets and computational power and a lack of transparency in ANN decision-making ("black box problem") were noted. This review's theoretical contributions emphasize ANNs' role in enhancing marketing strategies by improving the accuracy of consumer behavior predictions. However, methodological challenges, particularly the interpretability and explainability of ANN models, are significant barriers to widespread adoption in marketing. Practically, ANNs provide marketers with tools

to personalize marketing strategies and optimize resource allocation through dynamic customer segmentation and demand forecasting. The findings suggest that future research should improve ANN transparency, explore ethical considerations, and address data quality issues. ANNs represent a powerful tool for predicting consumer behavior and optimizing marketing strategies. Their ability to model complex, nonlinear relationships offers marketers significant advantages in customer segmentation and personalized marketing efforts. However, to realize their full potential in marketing, the challenges of explainability and data requirements must be addressed. Future research should develop more transparent and explainable ANN models and explore ethical and data privacy issues related to ANN applications in consumer behavior. This study focused on a specific time frame (2020-2024), and further research should examine the long-term impacts of ANNs on marketing. Additionally, more studies are required to explore ANN applications across different cultural and demographic contexts and improve the transparency of these models. Future research should also focus on integrating ANN models with real-time data to enhance their responsiveness to consumer behavior changes.

**Keywords:** Artificial Neural Networks, Consumer Behavior, Marketing

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

**Keziban Kılıç Topal<sup>1\*</sup>, Vedide Rezan Uslu<sup>2</sup>**

<sup>1</sup>Department of Educational Sciences, Amasya University, Amasya, Türkiye

<sup>2</sup> 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.

**Keywords:** Rainfall, Rainfall Forecasting, Genetic Algorithm, Time Series, Double Seasonal Holt Winters.



## Evaluating the Impact of Client Selection Algorithms on Performance and Training Speed in Heterogeneous Resource Allocation and Non-IID Data

Övgü Özdemir<sup>1\*</sup>, Ayşe Dilek<sup>1</sup>, Pınar Karagöz<sup>2</sup>

<sup>1</sup>Department of R&D, Proven Technologies, Ankara, Türkiye

<sup>2</sup>Department of Computer Engineering, Middle East Technical University, Ankara, Türkiye

### Abstract

Federated learning has emerged as a powerful paradigm for training machine learning models across distributed devices while preserving data privacy. The effectiveness of federated learning heavily depends on the client selection process, as it directly influences both the convergence speed and overall performance of the model. The challenges of non-IID data distributions and heterogeneous resource allocation among clients make client selection a critical factor in optimizing training outcomes. This study investigates the effects of various client selection algorithms on performance and training speed in federated learning environments with non-IID data and heterogeneous resource allocation. We evaluate loss-based selection algorithms, such as FedCorr [1] and Power-of-Choice [2], alongside random selection. Additionally, we propose a straightforward client selection strategy that leverages the processor type of clients to enhance training speed. The proposed approach is compared with existing methods under different data distribution and resource allocation scenarios. The algorithms were tested using the MNIST and CIFAR-10 datasets, with resource heterogeneity simulated through a setup consisting of an NVIDIA Jetson device, three Raspberry Pi 4 units, and four computers with Intel CPUs. We developed a web-based, no-code platform for federated learning and conducted all experiments using this platform. The results demonstrate the potential of utilizing client selection methods based on local loss and processor type to improve both performance and training speed in federated learning with non-IID data and heterogeneous resources.

**Keywords:** Federated Learning, Client Selection, Non-IID, Heterogeneous Resource

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

**Kadriye Nurdanay Öztürk<sup>1\*</sup>, Serpil Türkyılmaz<sup>2</sup>**

<sup>1</sup>İstatistik ve Bilgisayar Bilimleri Bölümü, Bilecik Şeyh Edebali Üniversitesi, Bilecik, Türkiye

<sup>2</sup>İstatistik ve Bilgisayar Bilimleri Bölümü, Bilecik Şeyh Edebali Üniversitesi, Bilecik, Türkiye

### Özet

Bu çalışmada, Mevsimsel Ototregresif Bütünleşik Hareketli Ortalama (SARIMA) modeli ve Yapay Sinir Ağları (YSA) modeli birleştirilerek hibrit bir SARIMA-ANN 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-ANN hibrit modelinin öngörü performansları Ortalama Karesel 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

## Yapay Zeka Araçları ile Veri Analizine Genel Bir Bakış

### Özge Taş<sup>1\*</sup>

Kapadokya Meslek Yüksekokulu, Kapadokya Üniversitesi, Nevşehir, Türkiye

### Özet

Yapay zeka (YZ) teknolojilerinin hızlı gelişimi ve büyük veri (big data) kavramının ortaya çıkışı, veri analizi alanında devrim niteliğinde değişikliklere yol açmıştır. Bu çalışma, YZ araçlarının veri analizi süreçlerindeki rolünü, uygulamalarını ve etkilerini kapsamlı bir şekilde incelemektedir. Veri biliminde verilerin derinlemesine incelenmesi analiz edilmesi ve yorumlanması hem yapay zeka algoritmalarında karşımıza çıkarken yapay zeka uygulamalarıyla da veriyi işlemek şeklinde roller ortaya çıkmıştır. YZ tabanlı veri analizi araçları, verinin analiz edilmesi ihtiyacına cevap vererek, geleneksel analiz yöntemlerinin ötesine geçen yetenekler sunmaktadır. Yapay zeka, insan zekasını taklit eden ve belirli görevleri yerine getirmek için öğrenme, problem çözme ve karar verme yeteneklerine sahip bilgisayar sistemlerini ifade eder. Veri analizi bağlamında YZ, özellikle makine öğrenimi (machine learning) ve derin öğrenme (deep learning) alt dalları aracılığıyla uygulanmaktadır. Bu teknolojiler, büyük veri setlerinden anlamlı örüntüler çıkarma, tahminler yapma ve otomatik karar verme süreçlerini mümkün kılmaktadır. Makine öğrenimi algoritmaları, verilerden öğrenerek performanslarını zaman içinde artırmaktadır. Denetimli öğrenme (supervised learning), denetimsiz öğrenme (unsupervised learning) ve pekiştirmeli öğrenme (reinforcement learning) gibi farklı yaklaşımlar, çeşitli veri analizlerinde uygulanabilmektedir. Derin öğrenme ise, özellikle büyük ve karmaşık veri setlerinde etkili olan, çok katmanlı yapay sinir ağlarını kullanan bir makine öğrenimi alt kümesidir. Teknolojinin hız kazanmasıyla verilerin işlenmesi ve analizi konuları çalışmalarda hız kazanmaktadır. Bu durumda büyük verinin işlenmesi anlamlı bilgilere dönüştürülmesi ciddi bir zorluk teşkil etmektedir. Yapay zeka bu noktada devreye girerek verilerin işlenmesi, analiz edilmesi ve yorumlanması süreçlerinde kritik bir rol oynar. Bu süreçte yapay zeka (YZ) araçları, özellikle makine öğrenimi ve derin öğrenme teknikleri, geleneksel veri analiz yöntemlerini geride bırakarak daha etkin ve verimli sonuçlar sunmaktadır. Bu makalede, YZ araçlarının veri analizinde nasıl kullanıldığını hangi tekniklerin öne çıktığını ve bu tekniklerin sağladığı avantajlar detaylandırılacaktır.

**Anahtar Kelimeler:** Yapay zeka, Veri analizi, Büyük veri

## Selecting Differentially Expressed Genes in Genomic Data Using the Length of ROC (L-ROC) Method

**Ebru Ozturk<sup>1\*</sup>, Merve Basol Goksuluk<sup>2</sup>, Dincer Goksuluk<sup>2</sup>**

<sup>1</sup>Department of Biostatistics, School of Medicine, Hacettepe University, Ankara, Turkiye

<sup>2</sup>Department of Biostatistics, School of Medicine, Erciyes University, Kayseri, Turkiye

### Abstract

Microarray and RNA sequencing (RNA-Seq) are vital tools in genetic data analysis, enabling precise molecular diagnoses and the identification of disease-associated genes. The inherent nature of genomic data is characterized by its high dimensionality, encompassing numerous features, such as genes or gene regions, numbering in the thousands. Consequently, effective feature selection is imperative to mitigate model complexity, enhance modeling accuracy, and yield unbiased estimations aligned with clinical outcomes. This study evaluates the feature selection performance of the length of ROC (L-ROC) method in RNA-Seq studies and compares it with other established methods. RNA-Seq data underwent preprocessing steps including low count filtering, normalization, and transformation. Various methodologies were applied to identify differentially expressed (DE) features, including L-ROC [1], area under the ROC curve (AUC), generalized AUC (gAUC) [2], and negative binomial generalized linear models (GLMs) as implemented in DESeq2 [3]. A simulation study varied parameters such as sample sizes, rates of differential expression, and levels of overdispersion to comprehensively evaluate method efficacy. The L-ROC method demonstrates comparable or superior performance in identifying DE features under conditions of pronounced overdispersion, a common occurrence in RNA-Seq datasets. However, its efficacy diminishes with slight overdispersion. True detection rates increase with higher differential expression rates or sample sizes as expected. L-ROC emerges as a preferable approach for selecting DE features, particularly under conditions of high overdispersion. While GLMs perform well under the negative binomial model assumption, their efficacy may diminish otherwise. L-ROC's distributional independence renders it robust for RNA-Seq data analysis, highlighting its practical significance primarily in RNA-Seq data.

**Keywords:** RNA Sequencing, differential expression, genomic data, feature selection.

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

Ayşegül Kaçar<sup>1\*</sup> ve Erkan Işıklı<sup>2</sup>

<sup>1</sup>Department of Industrial Engineering, Graduate School of Education, Istanbul Technical University, Istanbul, Turkey

<sup>2</sup> Department of Industrial Engineering, Faculty of Business Administration, Istanbul Technical University, Istanbul, Turkey

### Abstract

This study introduces a new methodology that combines process mining with predictive analytics to analyze and enhance the impact of software changes on IT Operations and Service Desk processes. The methodology leverages process mining to uncover the relationships between incidents triggered by software changes and to identify critical points within these processes. Insights from this analysis are then used as inputs for forecasting models to more accurately predict operational metrics such as incident frequency and resolution times. The forecasting results provide decision-makers with reliable, data-driven information that aids in optimizing process improvements and resource allocation. Enabling the early identification of potential risks associated with software changes, this methodology supports proactive intervention planning, thereby enhancing operational efficiency and effectiveness while preserving customer satisfaction and service quality. The present study offers a unique framework for operational risk management and efficiency improvement in IT processes through large-scale data analytics, contributing to the advancement of industry standards. Furthermore, the flexible and data-driven nature of this methodology demonstrates broad applicability for process optimization across various sectors, including banking, healthcare, and manufacturing.

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

## Türkiye'nin Ulusal Yapay Zeka Stratejisinin Kelime Bulutu (Wordclouds) Aracıyla İncelenmesi

Gamze Sinem Özer<sup>1\*</sup>

<sup>1</sup>Siyaset Bilimi ve Kamu Yönetimi, Sosyal Bilimler Enstitüsü, Akdeniz Üniversitesi, Antalya, Türkiye

### Özet

Bu çalışmada, Türkiye'nin Dijital Dönüşüm Ofisi tarafından hazırlanan Ulusal Yapay Zeka Stratejisi'nin içeriği, kelime bulutu (wordclouds) aracıyla incelenerek dokümanda en sık vurgulanan konuların ortaya çıkarılması amaçlanmıştır. Yapay zeka, son yıllarda küresel ölçekte hızla gelişen ve pek çok sektörde dönüşüme yol açan yenilikçi bir teknoloji alanıdır. Türkiye de bu alanda stratejik bir yol haritası belirleyerek, ulusal düzeyde yapay zeka teknolojilerinin geliştirilmesi ve uygulanması için kapsamlı bir strateji belgesi oluşturmuştur. Bu stratejinin ana hatlarını ve öncelikli alanlarını anlamak amacıyla, strateji dokümanı metin madenciliği tekniklerinden bir olan kelime bulutu analizine tabi tutulmuş ve metinde en sık kullanılan kelimeler görselleştirilmiştir. Dolayısıyla kelime bulutu analizi ile strateji dokümanının içeriği derinlemesine incelenmiş ve dokümandaki temel kavramlar, öncelikli alanlar ve odak noktaları belirlenmiştir. Elde edilen sonuçlar, Türkiye'nin yapay zeka alanındaki hedefleri, öncelikleri ve stratejik yaklaşımlarını daha iyi anlamaya olanak tanımaktadır. Analiz, özellikle “yapay zeka”, “veri”, “inovasyon”, “güvenlik” ve “etik” gibi anahtar kelimelerin öne çıktığını göstermiştir. Bu bulgular, Türkiye'nin ulusal yapay zeka stratejisinin hem teknolojik gelişmelerin hem de etik ve toplumsal etkilerin göz önünde bulundurularak hazırlandığını ortaya koymaktadır. Çalışma, aynı zamanda bu stratejinin, ulusal ve uluslararası düzeyde nasıl konumlandığını ve hangi alanlarda derinleşmeye ihtiyaç duyduğunu belirlemek açısından da önemli ipuçları sunmaktadır.

**Anahtar Kelimeler:** Ulusal yapay zeka stratejisi, Türkiye, Kelime bulutu, Doküman analizi.

## High Dimensional Classification for Binary Features based on Decision Trees: An Empirical Study on ESG Reporting Companies in Türkiye

**Dr. Önder DORAK**<sup>1\*</sup>

<sup>1</sup>İşletme Bölümü/Sayısal Yöntemler ABD, Anadolu Üniversitesi, Eskişehir, Türkiye

### Abstract

This study aims to explain the impact of ESG (environmental, social, governance) binary (Boolean) metrics on ESG rating and to determine the most significant binary metrics – those that have only True/False values- for each main ESG categories. The study specifically aims to highlight the key binary metrics that are more accessible and interpretable for green investors, as these binary indicators are often simpler to comprehend and analyze compared to their numerical metrics used in ESG methodology. In parallel with this purpose, 230 binary metrics were selected, after numerical and sector specific metrics were disregarded, from a larger dataset of 730 ESG metrics. The study includes BIST (Istanbul Stock Exchange) companies publishing ESG sustainability reports between 2008-2023. Data is collected from LSEG Eikon (London Stock Exchange Group). The methodology used to classify the observations is the recursive partitioning algorithm CART (Classification and Regression Trees) chosen for its ability to handle binary data effectively and to identify hierarchical structures among the metrics. By using this approach, the study delves into how binary indicators can predict ESG performance and the relative importance of each metric within the context of environmental, social, and governance criteria. The study conducts four separate classifications, using the ESG combined score, environmental pillar score, social pillar score and governance pillar score as the target variables. The well-known standard three-way holdout method is followed for machine learning procedure by splitting 20% of data for testing and remaining 80% for training and validation. The results could be organized into three main aspects. Firstly, key binary metrics are identified that are most significantly impacts the overall ESG classification thanks to feature importance metrics and the hierarchy of the metrics in the decision trees after pruning. Secondly, the study also shows how each environmental, social and governance categories are shaped by the binary indicators. The most significant metrics for environmental, social and governance pillars are “Climate Change Commercial Risk Opportunities”, “Policy Forced Labor”, and “External Consultants”, respectively. Lastly, the study evaluates the performance of decision trees for classification that is solely based on binary data. The model for overall classification performs well with an average accuracy of 84% with balanced performance across the four classes. However, the relatively low recall (68%) and precision (68.3%) metrics indicate potential challenges in correctly identifying positive instances. Important features for overall classification also align with the important features determined in the main category (ESG) level classifications which points out the consistency of the classification algorithm. Findings suggest that binary metrics have significant impact on ESG scoring and classification.

**Keywords:** Machine learning, Decision trees, Environmental, social, and governance (ESG) investing, Feature selection, Binary data classification

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

Tuğba Yasemin Karagöz<sup>1\*</sup>, Erkan Işıklı<sup>2</sup>

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

<sup>2</sup>Endüstri 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 tekli 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 seti üstünde çalışılmıştır. Önerilen modelin göreceli performansı; geleneksel tekli 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 tekli tahmin modellerinden daha iyi sonuç verdiğini, ele alınan kombinasyon modellerinden de çoğunlukla daha başarılı olduğunu ortaya koymuştur.

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



## Hybrid Forecasting Approaches for Air Travel and Energy Demand: A Comparative Study Using ATA Method and Machine Learning Algorithms

Tuğçe Ekiz Yılmaz<sup>1\*</sup>, Güçkan Yapar<sup>1</sup>

Department of Statistics, Dokuz Eylül University, Izmir, Türkiye

### Abstract

This study aims to explore the application of advanced time series forecasting techniques on two critical datasets: the AirPassenger dataset and the Electricity Demand dataset for New South Wales (NSW) in Australia. By focusing on two distinct domains—air travel and energy consumption—this research seeks to assess the effectiveness of both traditional statistical methods and modern machine learning approaches in accurately capturing patterns, trends, and anomalies in these datasets. The AirPassenger dataset, a widely used benchmark in time series forecasting, contains monthly international airline passenger numbers from 1949 to 1960. This dataset is ideal for analyzing long-term trends and pronounced seasonal effects, offering insight into the cyclical nature of air travel demand. Conversely, the Electricity Demand dataset for NSW provides daily electricity consumption data, reflecting short-term fluctuations driven by various factors such as weather, industrial activity, and socio-economic conditions. The combination of these datasets presents a unique challenge in balancing long-term seasonal forecasting with short-term demand spikes. For the modeling approach, we employ traditional time series methods, focusing primarily on Exponential Smoothing and the ATA Method. The ATA method is an innovative and highly efficient forecasting technique, offering significant improvements over traditional methods like Exponential Smoothing. It not only provides faster computation and greater simplicity, but also delivers superior forecasting accuracy by effectively addressing issues like initial value selection and adapting to various demand patterns. Unlike other methods, ATA addresses the issue of selecting initial values, making it a more efficient option for handling time series data. Its strength lies in its ability to capture both stable and irregular demand patterns, making it particularly suitable for forecasting in domains such as air travel and energy consumption. In addition to traditional methods, we implement advanced machine learning models, including XGBoost, Long Short-Term Memory (LSTM) networks, and LightGBM. These machine learning techniques, known for their ability to capture complex, non-linear relationships in data, are further tested in hybrid and ensemble configurations to explore the best model combinations. Our approach aims to leverage the strengths of each method, combining ATA's efficiency and simplicity with machine learning's capacity to obtain accurate forecasts in the data. To evaluate model performance, we utilize standard error metrics such as Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE). These metrics will enable us to compare the accuracy, robustness, and reliability of each model across both datasets. Additionally, we investigate how the combination of ATA with machine learning models can further improve predictive accuracy, particularly in capturing the unique characteristics of air travel and energy consumption. Our findings are expected to provide valuable insights into the strengths and limitations of different forecasting models when applied to varying types of time series data. This research contributes to the fields of transportation and energy forecasting by highlighting the effectiveness of hybrid approaches and offering practical applications for decision-making in both sectors.

**Keywords:** Time Series Forecasting, ATA Method, Exponential Smoothing, AirPassenger Dataset, Electricity Demand Forecasting, Machine Learning, XGBoost, LSTM Networks, LightGBM, Hybrid Forecasting Models

## Nesne Tanıma ve Sınıflandırmada Kullanılan Derin Öğrenme Yöntemleri

**Erdem Uysal<sup>1</sup>, Hamza Erol<sup>1</sup>**

<sup>1</sup>Mersin Üniversitesi, 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

**Bilimin İstatistiği:  
Bilimetre ve Türkiye'deki Bilimetre Çalışmalarının Bilimetre Analizi**

**Yunus Emre Karaman<sup>1\*</sup>**

Yönetim Bilişim Sistemleri, İktisadi İdari ve Sosyal Bilimler Fakültesi, Kapadokya Üniversitesi,  
Nevşehir, Türkiye

**Özet**

Bu çalışma, bilimin nicel analizi olarak bilinen bilimetre yöntemlerinin temel ilkelerini ve Türkiye'de bu alanda yapılan çalışmaların bilimetre analizini sunmayı amaçlamaktadır. Bilimetre, bilimsel yayınların sayısal analizine dayalı olarak bilimsel üretkenliği, etki gücünü ve bilimsel iş birliğini ölçmeyi hedefleyen bir disiplindir. Çalışmada, Türkiye'deki bilimetre çalışmalarının yaygınlık düzeyi, kullanılan yöntemler, konu başlıkları ve araştırmacılar arasındaki iş birliği ağları ele alınmıştır. Web of Science, Scopus ve Google Scholar gibi uluslararası veri tabanlarından elde edilen verilere dayanarak, Türkiye'de bilimetre alanında yapılan çalışmaların sayısı, yıllara göre dağılımı, en çok atıf alan çalışmalar ve öne çıkan araştırma alanları ortaya konmuştur. Ayrıca, Türkiye'nin bu alandaki küresel bilimsel katkısı ve diğer ülkelerle iş birliği düzeyi de incelenmiştir. Bu çalışma, Türkiye'nin bilimsel performansının küresel ölçekte daha iyi anlaşılmasına ve gelecekteki politika geliştirmelerine katkı sunmayı hedeflemektedir.

**Anahtar Kelimeler:** Bilimetre, Nicel analiz, Bilimin istatistiği

## Classification of Brand Perception Using Random Forest: Brand Preference, Brand Loyalty, and Brand Trust

Muhammed İkbal Yılmaz<sup>1\*</sup>, Özge Yıldız<sup>2</sup> and Sahika Gökmen<sup>3,4</sup>

<sup>1,2</sup>Undergraduate Student, Department of Econometrics, Faculty of Economics and Administrative Sciences, Hacı Bayram Veli University, 06500, Ankara, Turkey

<sup>3</sup>Associate Professor, Department of Econometrics, Faculty of Economics and Administrative Sciences, Hacı Bayram Veli University, 06500, Ankara, Turkey

<sup>4</sup>Researcher, Statistics Department, Uppsala University, Uppsala, Sweden.

### Abstract

Brands play a crucial role in today's world, helping brand executives develop strategies, gain a competitive advantage, and connect with their target audience. Brand Preference (BP), Brand Loyalty (BL), and Brand Trust (BT) are key concepts embraced by fashion experts and capital producers, and they are further explored through university students in this study. This research addresses a gap in the literature by analyzing students' BP, BL, and BT levels using machine learning and the Random Forest algorithm. Although real-time data from online shopping platforms can assess personal preferences, identifying the most influential factors affecting brand attitudes with machine learning is still uncommon. By applying decision tree classifiers on various dataset subsamples and using a meta-predictor to reduce overfitting, this study aims to improve prediction accuracy. The research is based on data from undergraduate students at Ankara Hacı Bayram Veli University's Econometrics and Economics departments during the 2023/2024 spring semester. The 'Brand Perception Scale' data is analyzed with the Random Forest algorithm, known for its accuracy in classification. The dependent variables BP, BL, and BT are predicted based on twenty-two independent variables. The main results show that brand reputation, recognizability, fashion, and quality are the most influential factors for all three variables, while demographic factors such as birthplace, gender, and department have minimal impact.

**Keywords:** Machine Learning Classification, Brand Trust, Brand Loyalty, Brand Preference, Random Forest, Decision Trees

## Yapay Sinir Ağları ile İstanbul Elektrik Talep Tahmini

**Hayriye Yasak Özkal<sup>1\*</sup>, Buse Kaylan<sup>1</sup>, Meltem Sipahi<sup>1</sup>, Muammer Altan Çakır<sup>2</sup>**

<sup>1</sup>Veri Mühendisliği ve İş Analitiği Anabilim Dalı, İstanbul Teknik Üniversitesi, İstanbul, Türkiye

<sup>2</sup>Fizik 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, EPIAŞ 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))

## Economic and Structural Changes in Cancer Drug Sales: Analysis of Turkey between 2010-2018

Mustafa Asım Özalp<sup>1\*</sup>, Osman Tolga Kaskatı<sup>2</sup>

<sup>1</sup>Department of Actuarial Science, Hacettepe University, Ankara, Türkiye  
<sup>2</sup>BYS Grup AŞ.

### Abstract

In this study, a time series analysis on the economic and patient parameters affecting cancer drug sales in Turkey is conducted. Using data between 2010-2018, factors such as TL exchange rate, inflation rates and annual patient numbers are taken into consideration. With the help of the Quandt-Andrews breakpoint test, possible structural changes over the period are identified and their effects on drug access are analysed. The findings reveal that the depreciation of the TL and the changes in regulations have a negative impact on access to medicines for cancer patients. The results emphasise that economic instability and regulatory changes can have serious consequences on health services. These findings suggest that health policies and drug access strategies should be reconsidered.

**Keywords:** Time series analysis, Drug access, Inflation, Quandt-Andrews breakpoint test, Drug access, Economic fluctuations

## Eđitim İstatistiklerinde R ile Veri Kontrolünün Otomasyonu

**Fethi Şaban Özbek**<sup>1\*</sup>

<sup>1</sup>Sosyal İstatistikler Daire Başkanlığı/Türkiye İstatistik Kurumu

### Özet

Resmi istatistik üretim sürecinde veri kontrolünün otomasyonunun sağlanması veya geliştirilmesi, yalnızca her yıl kontrol aşamasında harcanan zaman ve emeğin azaltılmasını sağlamakla kalmaz, aynı zamanda verilerin doğruluğunu ve güvenilirliğini de artırır. Otomasyon sayesinde, insan hatası riski en aza indirgenir ve veri kontrol süreçlerinde tutarlılık sağlanır. Bu da karşılaşılabilecek sorunlara ve anormalliklere daha hızlı müdahale edilmesine olanak tanır. Ayrıca, tekrarlayan ve zaman alıcı görevlerin otomatik hale getirilmesi, uzmanların daha stratejik ve analitik görevler üzerinde çalışmasına imkân tanır, böylece istatistik üretim sürecinin genel verimliliği ve kalitesi artar. Kazanılan zaman ve emek, yeni yöntemlerin geliştirilmesi, veri analizinin derinleştirilmesi ve karar alma süreçlerinin iyileştirilmesi gibi daha katma değerli alanlara yönlendirilebilir, bu da istatistiksel çıktının genel olarak daha etkili ve amaca uygun olmasını sağlar. Bu ve benzeri faydalar doğrultusunda, eğitim istatistiklerinde veri kontrolünün otomasyonunun geliştirilmesine yönelik çalışmalar yürütülmektedir. Bu çalışma ile Türkiye İstatistik Kurumu'nda Resmi İstatistik Programı kapsamında üretilen eğitim istatistikleri üretim sürecinde R programlama dili ile veri kontrolünün otomasyonunun sağlanarak her yıl Excel gibi programlar aracılığıyla düzenli olarak yapılan kontrol işlemlerinin R komut dosyaları yardımı ile otomatik hale getirilmesi çalışmaları hakkında bilgi verilmektedir. Bu çalışmalar kapsamında, Merkezi Dağıtım Sistemine yüklenen yaklaşık 540 bin satırdan oluşan ulusal eğitim istatistikleri verilerinin içsel kontrolünü sağlayan, okullaşma oranları hesaplamaları için oluşturulan tabloların tutarlılığını (İstatistik Bölge Birimleri Sınıflandırması (İBBS) 1, 2 ve 3 değişken toplamalarının kontrolü) ve bir önceki yıla göre değişim oranlarını kontrol eden R komut dosyaları geliştirilmiş ve bu adımlara ait veri kontrolünün otomasyonu sağlanmıştır. Bu ürünlere ait kontrol adımlarının geliştirilmesinin yanı sıra eğitim istatistikleri kapsamında üretilen diğer istatistik ürünlerine ait veri kontrolü otomasyonunun sağlanması/geliştirilmesi planlanmaktadır.

**Anahtar kelimeler:** Eğitim istatistikleri, İstatistik üretim süreci, R, Veri otomasyonu



## Zaman Serisi Görüntüleri ve Yapay Zekâ Teknikleri İle Tarım Ürün Türü Sınıflaması

**Mert Kubilay Uludağ<sup>1\*</sup>**

<sup>1</sup>CBS Grup Başkanlığı/TÜİK, Ankara, Türkiye

### Özet

Yapay zeka (YZ), son yıllarda birçok alanda devrim yaratarak, karmaşık veri analizlerinde ve karar verme süreçlerinde insan gücünü tamamlayıcı bir rol üstlenmiştir. YZ teknikleri kullanılarak uydu görüntülerinin sınıflandırılmasında yüksek doğruluk oranlarına ulaşılmaktadır. Bu sayede, tarımsal alanların haritalanması, rekolte tahmini ve ürün sağlığı takibi gibi konularda daha etkin ve güvenilir sonuçlar elde edilmektedir. Bu çalışmada, ilk fazda Türkiye'nin Kırklareli ve Ankara illerinde, buğday, mısır ve ayçiçeği ürünlerinin haritalanması, ikinci fazda ise Kırklareli ilinde buğday, ayçiçeği, mısır, arpa, yulaf, kanola, tritikale, pirinç, yonca, çayır otu, macar fiği, yem bezelyesi, kavaklık ve üzüm olmak üzere toplamda 15 tarımsal ürünün haritalanması hedeflenmiştir. Ayrıca ikinci fazda farklı algoritmaların doğruluğa etkisinin de araştırılması amaçlanmıştır. Çalışmada, ilk fazda Sentinel-1 ve Sentinel-2 ikinci fazda ise sadece Sentinel-2 uydu görüntülerinden elde edilen veriler kullanılarak, ilk fazda rastgele orman sınıflayıcısı ikinci fazda ise hem rastgele orman hem de LightGBM algoritması aracılığıyla tarım ürün türü haritaları oluşturulmuştur. Ayrıca, NDVI ve BSI gibi görüntü indekslerinin de kullanılmasıyla sonuçların doğruluğunun artırılması amaçlanmıştır. Çalışmanın ikinci fazında Sentinel-2 görüntülerinin çözünürlüğünü arttırma ve 20 metrelik bantların 10 metreye getirilmesi üzerine de çalışılmıştır. Ankara ve Kırklareli illerinde aynı dönem ve ürünlerle yapılan çalışmaların karşılaştırılması sonucunda, farklı coğrafi bölgelerdeki doğruluk analizleri değerlendirilmiştir. Ayrıca ikinci fazda farklı algoritmaların sınıflamaya etkisi ve algoritmaların performans farkı gözlemlenmiştir. Sonuç olarak, bu çalışmada elde edilen verilerin, tarım istatistiklerinin hesaplanmasında ve gelecekteki tahminlerde kullanılabilirliği araştırılmıştır.

**Anahtar Kelimeler:** Makine öğrenmesi, Uzaktan algılama, Ürün türü sınıflaması

**Elektronik Veri Araştırma Merkezi:  
Türkiye'de Veriye Dayalı Araştırmaların Güvenli ve Erişilebilir Platformu**

**Serhat Atakul<sup>1</sup>, Ömer Faruk Ünal<sup>1\*</sup>**

<sup>1</sup>Türkiye İstatistik Kurumu, Ankara, Türkiye

**Özet**

Türkiye İstatistik Kurumu (TÜİK), anketler, idari kayıtlar ve diğer kaynaklardan elde ettiği mikro verileri, mevzuat çerçevesinde araştırmacıların istifadesine sunmaktadır. Mikro veriler birey veya girişim düzeyindeki kayıtların gizliliğini ve güvenliğini sağlamak amacıyla, isim, unvan ve adres gibi tanımlayıcı bilgiler çıkarılmakta ve vergi kimlik numarası gibi hassas veriler değiştirilerek anonimleştirilmektedir. Bu mikro veri setleri, hali hazırda yalnızca 14 ilde bulunan ve internet erişimi olmayan, veri transferine izin verilmeyen, kapalı bir ağ ortamında çalışan Veri Araştırma Merkezlerinde erişime açıktır. Türkiye'nin "Türkiye Yüzyılı" vizyonu doğrultusunda TÜİK tarafından geliştirilen Elektronik Veri Araştırma Merkezi (E-VAM), veri erişimine yeni bir soluk getirmiştir. E-VAM sayesinde araştırmacılar, mekânsal sınırlamalardan bağımsız olarak, istedikleri zaman ve mekânda, güvenli bir ortamda mikro verilere erişim sağlayabilmekte ve analizler gerçekleştirebilmektedir. E-VAM, veri gizliliği ve güvenliğine azami derecede riayet etmektedir. Fare ve klavye hareketlerini kayıt altına alma, dosya indirme ve gönderimini engelleme gibi önlemler sayesinde, mikro veriler yetkisiz erişim ve kullanıma karşı korunmaktadır. Araştırmacılar, E-VAM portalı üzerinden mikro veri setlerini inceleyebilir, ihtiyaç duydukları veri setlerini sepetlerine ekleyebilir ve e-Devlet aracılığıyla başvurularını tamamlayabilirler. Bu kullanıcı dostu arayüz, kurumsal süreçlerin de optimize edilmesine katkı sağlamaktadır. E-VAM'ın pilot çalışması dört üniversite ile başarıyla tamamlanmış olup, hem insan kaynağı hem de enerji ve işgücünden tasarruf sağladığı gözlemlenmiştir. E-VAM, üniversiteler ve kamu kurumları için merkezi bir veri erişim ve paylaşım ekosistemi oluşturmayı hedeflemektedir. TÜİK'in öz kaynakları ile geliştirilen bu platform, Ulusal Yapay Zekâ Stratejisi'nin amaçlarından biri olan Kamu Veri Alanı projesinde de altyapı olarak kullanılma onayı almıştır. E-VAM, verinin çoklanmasını ve mükerrer veri tutulmasını engelleyerek dijital dönüşüme ivme kazandırmaktadır. Büyük hacimli mikro verilerin kullanımını kolaylaştıran ve araştırmacıların kodlarını paylaşabileceği bir ortam sunan platform, büyük veri analitiği ve yapay zekâ teknolojilerinin entegrasyonuna da imkân tanımaktadır. Bu sayede, farklı konulardaki mikro veri setleri birleştirilerek model oluşturma ve istatistiksel tahminler yapılabilecek, böylece bilgiye dayalı karar alma süreçleri desteklenecektir. E-VAM, Türkiye'de veriye dayalı araştırmaların ve inovasyonun önünü açan mühim bir adımdır. Bu platform sayesinde, araştırmacılar daha hızlı, güvenli ve verimli bir şekilde çalışabilecek, Türkiye'nin bilgi toplumu olma yolundaki yürüyüşüne katkıda bulunabilecektir.

**Anahtar Kelimeler:** Mikro veri, uzaktan erişim.

## Bölgesel Alt Kırılımlarda İnsani Gelişme Endeksi ve Bileşenleri, Metodolojik Yaklaşımlar

Derya Tuncer<sup>1\*</sup>

<sup>1</sup>Sosyal İstatistikler Daire Başkanlığı, Türkiye İstatistik Kurumu, Ankara, Türkiye

### Özet

Bir ülkenin gelişimini değerlendirmede göz önüne alınabilecek kriterin ekonomik büyümenin yanı sıra insanların iyi bir yaşam sürmeleri ve refah düzeyi ile de ilişkilendirilmesi gerektiğini vurgulayan İnsani Gelişme Endeksi (İGE); metodolojisi Dünya Bankası çalışmalarına dayanmakla birlikte UNDP tarafından tüm ülkeler için merkezi olarak hesaplanan ve insani gelişme raporları ile yayımlanmakta olan önemli bir sürdürülebilir kalkınma göstergesidir. UNDP'nin merkezi İGE metodolojisi, yıllar içerisinde bir takım değişimler geçirerek evrimleşmesine rağmen 3 temel bileşeni (uzun ve sağlıklı yaşam, bilgiye erişim ve makul, iyi bir yaşam standardına sahiplik) değişmemiş olup, bu üç bileşenin her birinde elde edilen endekslerin geometrik ortalaması ile hesaplanmaktadır. Bazı ülkeler, kendi metodolojilerini kullanarak veya UNDP'nin metodolojisini uyarlayarak kendi İGE hesaplamalarını yapabilmiş olmakla birlikte bu durum nadiren görülmektedir. Ülkelerin kendi metodolojik uyarlamaları ile ülke düzeyinde İGE hesaplaması bile nadir görülen bir olgu iken bölgesel alt kırılımlarda İGE ve bileşen endekslerinin hesaplanması adım adım ilerlenmesi gereken ve farklı yöntemler içeren hesaplamalar ile detaylı bir çalışma gerektirmektedir. Bu çalışmada; hesaplamaları devam eden ve Türkiye İstatistik Kurumu'nda Resmi İstatistik Programı kapsamında il bazında (İstatistik Bölge Birimleri Sınıflandırması (İBBS-3)) üretilmesi planlanan İGE ve bileşen endekslerinden genel hatlarıyla bahsedilmekle birlikte, temelde gelir endeksi hesaplamasında kullanılan ve il bazında henüz mevcut olmayan değişkenlerin SPSS Modeller aracılığıyla veri madenciliği yöntemleri ile model denemeleri yapılarak tahminine odaklanılmaktadır. Yapay sinir ağları, karar ağaçları, doğrusal regresyon ve K\_en yakın komşu algoritmaları ile oluşturulan regresyon modelleri ile gelir endeksi hesaplamasında kullanılabilir; IBBS\_2 düzeyinde (26 bölge) mevcut olmasına rağmen, IBBS\_3 düzeyinde (il bazında) henüz mevcut olmayan satın alma gücü paritesi değişkeninin il bazında tahminin yapılması planlanmaktadır. Alınan sonuçlara göre en iyi tahmin yapay sinir ağları ile modelleme sonucunda görülmüş olup, özellikle sinir ağları ile makine öğrenimi üzerinde durulacaktır.

**Anahtar kelimeler:** İnsani Gelişme Endeksi (İGE), Gelir Endeksi, Makine Öğrenmesi ve yapay sinir ağları, Regresyon, Veri Madenciliği

## Veri Demokrasisi: Etkin Yönetişim ile Güçlenen Organizasyonlar

**Bilal Kurban<sup>1</sup>, Dr. Gaye Mat Çelik<sup>1\*</sup>, Dr. Serhan Kökhan<sup>1</sup>**

<sup>1</sup>Veri Yönetişimi Daire Başkanlığı, Yapay Zekâ ve Veri Analizi Grup Başkanlığı, Türkiye İstatistik Kurumu, Ankara, Türkiye

### Özet

Dijital dönüşümün hızla geliştiği günümüzde, verinin stratejik bir kaynak olarak önemi her geçen gün artmaktadır. Verinin etkin kullanımı, özellikle kamu kurumlarında, karar alma süreçlerini iyileştirerek, kamu hizmetlerinin daha etkin ve verimli bir şekilde sunulmasını sağlamaktadır. Türkiye İstatistik Kurumu (TÜİK), bu bağlamda veri yönetişimi ve kamu veri alanı çalışmalarını merkeze alarak, veri demokrasisini tesis etme yolunda önemli adımlar atmaktadır. Veri yönetişimi, verinin yaşam döngüsü boyunca mevzuatlar, politikalar ve düzenlemeler çerçevesinde sahiplik, sorumluluk, erişim, metaveri, kalite, gizlilik, güvenlik, veri yönetimi ve kültürünün sağlanması, bu süreçte ulusal ve uluslararası iş birliği ile uyumlu bir altyapının oluşturulması disiplini. Kamu veri alanı, teknik, idari ve hukuki zeminde kamu kurumlarının veri sorumlusu oldukları hizmet verilerinin yetkilendirilmiş bir platformda işlemek üzere konsolide edilmesi aracılığıyla, kamuda ileri veri analitiği ve yapay zeka uygulamaları için elverişli bir ortam sağlayan bir ekosistemdir. Veri demokrasisi ise verinin tüm paydaşlar için erişilebilir ve kullanılabilir olmasını ifade etmektedir. Bu çalışmada, TÜİK'in geliştirdiği kapsamlı veri yönetişimi çerçevesi ve bu çerçevenin kamuda veri demokrasisini nasıl şekillendirdiği ele alınmaktadır. TÜİK, ulusal veri yönetim politikalarının geliştirilmesine katkıda bulunarak, Türkiye'nin veri devrimine uyum sağlamasını amaçlamaktadır. Bu doğrultuda, TÜİK'in vizyonu 2024-2028 Stratejik Planı'nda "veri yönetiminde öncü kurum olma" olarak ifade edilmektedir. Veri demokrasisi, veri temelli karar alma süreçlerinin şeffaflık, hesap verebilirlik ve katılımcılık ilkeleri ile entegre edilmesini gerektirmektedir. TÜİK, kamuya açık veri setlerini artırmakta ve bu verilerin kolayca erişilebilir olmasını sağlamak için çeşitli platformlar geliştirmektedir. TÜİK, bu anlayışla, veri yönetişimi süreçlerini güçlendirerek ve kamu veri alanını genişleterek, Türkiye'nin dijital dönüşümüne katkıda bulunmaktadır. Gelecekte, TÜİK'in bu alandaki çalışmalarının daha da ileriye taşınmasıyla veri demokrasisinin tüm paydaşlar için daha erişilebilir hale gelmesi öngörülmektedir.

**Anahtar Kelimeler:** Veri yönetişimi, kamu veri alanı, veri demokrasisi

## Integrating RFM Analysis with Power BI for Enhanced Customer Segmentation and CLV Insights

Duygu Demiray Akkaya<sup>1\*</sup>, Ergun Eray Akkaya<sup>2</sup>, Muhammet Mehmet Kır<sup>1</sup>

<sup>1</sup>Department of Computer Engineering, Beykoz University, Istanbul, Turkiye

<sup>2</sup>Department of Management of Information Systems, Istanbul Topkapı University, Istanbul, Türkiye

### 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.

## Avrupa İnsan Hakları Mahkemesi Yargı Kararlarının Metin Madenciliği Yöntemleriyle Tahmin Edilmesi

Nagihan Ünal<sup>1\*</sup>, Oktay Yıldız<sup>2</sup>

<sup>1</sup>Gazi Üniversitesi Bilişim Enstitüsü, Ankara, *Türkiye*

<sup>2</sup>Gazi 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 işleme, Metin sınıflandırma, Metin verisi artırma

## Forecasting nonlinear models: A simulation study

**PhD. Lorena Salaj**<sup>1\*</sup>

<sup>1</sup>Mediterranean University of Albania, Faculty of Informatics, Albania

### Abstract

Today, large quantities of data are being collected for analytical purposes. In most cases, they don't have similar data generating processes; thus, it might be difficult identifying a single forecasting model to apply to all possible scenarios. Considering a time series as an independent dataset, choosing whether to apply a linear or a nonlinear forecasting model separately, may lead to inaccurate results, which means inappropriate future interventions and decisions. This paper regards a simulation study for the quantitative evaluation of the dependence of the forecasting models' performance on the characteristics of the data. We have simulated several datasets with known time series characteristics, under controlled conditions, including linear data generating processes and nonlinear ones. Heterogeneous time series are obtained by combining several data generating processes into one time series, to identify the more appropriate forecasting model for real world datasets, while the length and the number of the series are varied in different scenarios to explore the effect of data availability on the performance of the forecasting models. We have performed forecasts using both linear and nonlinear forecasting models and compared their performance in terms of percentage errors.

**Keywords:** Time series, Forecasting, Nonlinear models, ARIMA

## Sokak Hayvanlarına Yönelik Nefretin Duygu Analiziyle Tespiti

**Halil Özgür Çelik<sup>1\*</sup>, Elif Çoker<sup>1</sup>**

<sup>1</sup>İstatistik Bölümü, Fen Edebiyat Fakültesi, Mimar Sinan Güzel Sanatlar Üniversitesi, İstanbul, Türkiye

### Özet

Günümüzde toplumsal yaşamın ayrılmaz bir parçası haline gelen sosyal medya platformları, bireylerin duygu ve düşüncelerini açıkça ifade edebildikleri alanlardır. Bireyler, bu platformlar aracılığıyla, kendilerini daha iyi ifade etme ve düşüncelerini daha etkili şekilde paylaşma olanağı bulmuştur. Bireyin haberleşmesini, eğlenmesini, öğrenmesini ve iş yapmasını kolaylaştıran bu platformların aynı zamanda bazı zararlı yanları da bulunmaktadır. Bu zararlı yanlar özellikle yanıltıcı bilgi, güvenlik tehlikesi ve hatta uzun süreli olarak etkisi bilinmeyen olumsuz içeriklerin oluşması olarak örneklendirilebilir. İçerik açısından sosyal medyanın en ciddi sorunlarından biri nefret söylemleridir. Toplumda ayrımcılık ve dışlanma duygularının artmasına neden olan nefret söylemleri sadece insanların değil, tüm canlıların yaşamını olumsuz yönde etkilemektedir. Sosyal medya platformlarında yer alan nefret söylemlerine maruz kalan kesimlerin başında ise sokak hayvanları gelmektedir. Sokak hayvanlarına dair, insanların sağlığını tehlikeye atmakta veya insanların can ve mal varlıklarına zarar vermekte olduğu şeklinde yanıltıcı içerikler, sokak hayvanlarına yönelik nefret söylemini körüklemektedir. Böylece sosyal medya aracılığı ile körüklenen nefret söylemi insan-insan, insan-hayvan ve insan-doğa ilişkisinde daha kutuplaştırıcı bir rol oynarken toplumsal yaşamı daha problemlili bir noktaya çekmektedir. Bu çalışmada sokak hayvanlarına yönelik nefret söylemi, yaygın sosyal medya platformu olan Twitter'dan (X olarak adı değişen platform çalışmanın yapıldığı süreçte Twitter olarak kullanılıyordu) toplanan veriler ile incelenmektedir. Yaklaşık iki yıllık bir süre içinde (2021-2023) Twitter platformunda sokak hayvanları ile ilgili atılan tüm tweetler ele alınarak bu tweetlerdeki metinler üzerinde bir duygu analizi (sentiment analysis) çalışması yapılmaktadır. Duygu analizi, bir metinde yer alan duygu veya duygusal tonun analiz edilmesidir. Twitter üzerinden gerçekleştirilen duygu analizi Twitter kullanıcılarının tweetlerinde yer alan belirli bir duyguya dair olan bilgilerin toplanması ve analiz edilmesidir. Duygu analizinde doğal dil işleme, yapay öğrenme ve derin öğrenme yöntemleri kullanılır. Geleneksel doğal dil işleme yöntemleri kelime köklerine odaklanarak bir duygu tespiti yapmaya çalışırken modern yöntemlerde ise metin örneklerinin bir kısmına dair duygu kutupları etiketlenerek yapay öğrenme ve derin öğrenme yöntemleriyle tahmin gerçekleştirilir. Bu çalışma kapsamında güncel modern yöntemlerden biri olan BERT (Bidirectional Encoder Representations from Transformers) modeli kullanılmıştır. Yakın geçmişte Google araştırmacıları tarafından ortaya çıkarılan bu model ile daha doğru ve anlamlı sonuçlar elde etmek mümkündür. Çalışma kapsamında 100.000 adet tweet içeriği, Python dilinde oluşturulmuş olan Snsrape kütüphanesi aracılığı ile elde edilmiştir. Bu tweetler yine Python dilindeki çeşitli kütüphaneler aracılığı veri temizliği ve veri ön işleme adımlarından geçirildikten sonra rastgele örnekleme yöntemi ile 500 tweet elle etiketlenerek duygu kutupları elde edilmiştir. BERT modeli bu örnekleme verisi ile çalıştırılarak tahmin etiketlerinin başarısı, doğruluk oranı (accuracy) metriği ile ölçülerek 0,75 düzeyinde bulunmuştur. Yeterli başarı düzeyi elde edildikten sonra kalan tüm tweetlerin duygu kutupları tahmin edilmiştir. Tahminler çeşitli veri analizi ve veri görselleştirme teknikleri ile ifade edilerek sokak hayvanlarına yönelik nefret söylemlerine dair çeşitli bilgilere ulaşılmıştır.

**Anahtar Kelimeler:** Duygu Analizi, BERT, Yapay Öğrenme, Derin Öğrenme, Doğal Dil İşleme



## Rastgele Ormanlar ve Tekrarlayan Sinir Ağları Yöntemlerinin Teknoloji Endeksi Üzerinden Karşılaştırılmalı Analizi

**Melike Yalçın Düzgün<sup>1\*</sup>, Uğur Karabey<sup>2</sup>**

<sup>1</sup>Aktüerya Bilimleri, Kırıkkale Üniversitesi, Kırıkkale, Türkiye

<sup>2</sup>Aktüerya Bilimleri, Hacettepe Üniversitesi, Ankara, Türkiye

### Özet

Günümüzde teknolojinin hızlı gelişimi, son dönemlerde finansal piyasalarda teknoloji sektörünü ön plana çıkarmıştır. Bu bağlamda, teknoloji sektörünün genel performansının gösteren Borsa İstanbul A.Ş.'de işlem gören teknoloji endekslerinin fiyat hareketlerinin değişim ve tahmin analizi önem kazanmıştır. Yapılan çalışmalarda, geleneksel yöntemlere kıyasla yapay zekâ yöntemlerinin finansal piyasaların analiz edilmesinde daha başarılı sonuçların elde edildiği saptanmıştır. Bu çalışmada, teknoloji endeksini içeren veri setine çeşitli teknik göstergeler dahil edilerek makine öğrenimi yöntemlerinden Rastgele Ormanlar (RF) ve derin öğrenme yöntemlerinden Tekrarlayan Sinir Ağları (RNN) kullanılmıştır. Fiyat hareketlerinin değişim ve tahmin analizi üzerinden bu yöntemlerin sonuçları karşılaştırılmıştır.

**Anahtar Kelimeler:** Rastgele ormanlar, Tekrarlayan sinir ağları, Teknoloji endeksi, Fiyat tahmini

## Birliktelik Kuralı Analizi ile Ürün Kod Atama Uygulaması

**Kadir Korhan Babadağ<sup>1\*</sup>, Semih Ergişi<sup>2</sup>**

<sup>1,2</sup>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 yapılp 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.

1. İfadeler üzerinde yapılan dilbilimsel işlemler,
2. Modelleme için eğitim veri setinin hazırlanması,
3. 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 İfadeler, Denetimsiz öğrenme, Birliktelik kuralı

## MEDOS: Kısa Dönemli İstatistiklerin Mevsimsellikten Arındırılması için Bir R-Shiny Uygulaması

**Muhammed Fatih Tüzen**<sup>1\*</sup>

<sup>1</sup>Yöntem 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+

## Airline Business Model Evolution: A K-Means Clustering Analysis

Cem Ersöz<sup>1\*</sup>, Çağlar Üçler<sup>2</sup>

<sup>1,2</sup> Havacılık ve Uzay Bilimleri Fakültesi, Özyeğin Üniversitesi, 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 carriers, which contradicts its business model. Consequent analysis with 3 clusters yielded further dynamic shifts among airlines. In general, Southwest and United indicated evidence for hybrid approaches.

**Keywords:** K-means clustering, elbow method, airline business model.

## Yapay Zeka Destekli Veri Sınıflandırma ve Yönetim Modeli

Tülin Demiralp<sup>1\*</sup>, Ahmet Künyeli<sup>1</sup>, Necati Kaya<sup>2</sup>, Maryat Coşkun<sup>1</sup>

<sup>1,2</sup> 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, öz geçmiş, sertifika, işe giriş, idari şartname ve teknik şartnameden oluşan 7 tipte toplam 2392 Türkçe kurumsal doküman barındıran bir veri seti hazırlanmıştır. Veri ön işleme, belirteçleme, kelime köklerinin bulunması, veri temizleme ve kelimelerin vektör değerlerinin bulunması işlemlerinin ardından eğitim setindeki dokümanlar kullanılarak BERTurk 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 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

## Covid-19 ve Deprem Süreçlerinde Uzaktan Eğitim: Öğrenci Beklentileri ve Deneyimlerinin Makine Öğrenimi ile Karşılaştırılması

**Döndü Soytürk<sup>1\*+</sup>, Osman Keçeçi<sup>1</sup>, Şahika Gokmen<sup>1,2</sup>**

<sup>1</sup>İktisadi ve İdari Bilimler Fakültesi, Ekonometri Bölümü, Hacı Bayram Veli Üniversitesi, Ankara, Türkiye.

<sup>2</sup>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 çıktılarını 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 anlaşılabilmesi 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

## Enhancing Time Series Decomposition with Artificial Intelligence: Improving Pattern Identification and Interpretation in Economic Data

**Dr. Afzal Mahmood<sup>1\*</sup>, Dr. Kanwal Zahra<sup>2</sup>, Dr. Badar Rizvi<sup>3</sup>**

<sup>1</sup>Economics, University of Education, Lahore, Pakistan

<sup>2</sup>Business School, University of Central Punjab, Lahore, Pakistan

<sup>3</sup>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.

### IJS Codings:

- **C22:** Time-Series Models
- **C32:** Modeling Time-Series Data
- **G17:** Financial Forecasting
- **O33:** Technological Change: Choices and Consequences; Diffusion Processes

## Estimation based on Ranked Set Sampling for the Kumaraswamy-Normal Distribution

**Özge Gürer<sup>1\*</sup>, Birdal Şenoğlu<sup>1</sup>**

<sup>1</sup>Department of Statistics, Ankara University, Ankara, Turkey

### Abstract

Although traditional statistical methods generally are based on the assumption of normality, non-normal distributions are more common in real life problems. In this study, location and scale parameters of the Kumaraswamy-Normal (KwN) distribution, which can be used as a reasonable alternative to the normal distribution, are estimated under the assumption of known shape parameters, see Cordeiro and De Castro (2011) for the family of Kumaraswamy-Generalized distributions. Since KwN distribution has very flexible tails, it can easily be used in modeling data sets that can be encountered in many fields of science. Another important point addressed in this study is the use of ranked set sampling (RSS) instead of simple random sampling (SRS) in the selection of sample units. RSS was firstly used by McIntyre (1952) and is based on the idea of ranking the observations in a data set based on some auxiliary variables or visual assessments, without conducting any actual measurements. In this way, RSS better represents populations and provides more efficient estimators than the corresponding SRS does. The parameter estimation of the KwN distribution under RSS leads to likelihood equations including nonlinear functions of the unknown parameters, therefore explicit solutions of the maximum likelihood (ML) estimators cannot be obtained. In this study, the modified maximum likelihood (MML) methodology, which is based on the principle of linearizing the nonlinear terms in the likelihood equations around the expected values of the standardized ordered statistics by using the first two terms of the Taylor series expansion, is used, see Tiku (1967) for the details of MML methodology. RSS based MML estimators of the unknown parameters of KwN distribution are then compared with the corresponding commonly used SRS based ML estimators with respect to bias and mean squares error values by means of an extensive Monte Carlo simulation study. Results reveal that RSS based MML estimators derived in this study are much more efficient than their competitors for all cases.

**Keywords:** Ranked set sampling, parameter estimation, non-normality, modified likelihood, Monte Carlo simulation



## On Estimating the Parameters of Logistic Distribution using Particle Swarm Optimization

Adil Kılıç<sup>1</sup> and Birdal Şenoğlu<sup>2</sup>

<sup>1</sup>Department of Statistics, Kırıkkale University, 71450 Kırıkkale, Türkiye

<sup>2</sup>Department of Statistics, Ankara University, 06100 Ankara, Türkiye

### Abstract

In this study, we focus on the maximum likelihood (ML) estimation of the parameters of Logistic distribution. Since the explicit solutions of the likelihood equations cannot be obtained, we resort to particle swarm optimization (PSO) proposed by Kennedy and Eberhart (1995) among various metaheuristic methods widely used in the literature to solve the problem of maximizing the likelihood function with respect to the distribution parameters. However, the performance of PSO is significantly affected by the appropriate selection of its initial settings. Here, the performances of PSO based on fixed and data-driven search spaces, one of the most important initial settings of PSO, are investigated with a comprehensive Monte Carlo simulation study under different sample sizes. Four of the search spaces considered in this study are fixed, and the remaining one is a data-driven search space based on modified ML (MML) estimators of the distribution parameters proposed by Acitas et al. (2019); see also Tiku (1967) in the context of MML methodology. Note that fixed search spaces are determined from narrow to wide to make comparisons more comprehensive. According to the simulation results, the search space based on MML estimators is determined to be the most efficient among the others. It is also seen that the efficiency of PSO decreases rapidly as the range specified for the fixed search space expands as expected.

**Keywords:** Particle Swarm Optimization, Logistic Distribution, Parameter Estimation, Monte Carlo Simulation, Efficiency

## Borsa İstanbul Bist100 Endeksi için Kurulan Modellerin Karşılaştırılması

Aytaç Pekmezci<sup>1</sup>, Kurtuluş Bozkurt<sup>2</sup>, Özge Taş<sup>3\*</sup>

İstatistik Bölümü, Fen Fakültesi, Muğla Sıtkı Koçman Üniversitesi, Muğla, Türkiye  
Ekonomi Bölümü, Söke İşletme Fakültesi, Aydın Adnan Menderes Üniversitesi, Aydın, Türkiye  
Kapadokya Meslek Yüksekokulu, Kapadokya Üniversitesi, Nevşehir, Türkiye

### Özet

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 işlemi 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.

**Anahtar Kelimeler:** Ekonomi, Tahminleme, Arma, Arima

## Yapay Zekanın Kurumsal Dönüşümdeki Rolü

Ali Osman Bilgin<sup>1\*\*</sup>, Kudret Elif Berkman<sup>2</sup>, Ahmet Akın Atasoy<sup>3</sup>

<sup>1</sup>Veri Yönetişimi Daire Başkanlığı, TÜİK, 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 bildiriye, 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 halihazırda yürüttüğü projeler arasında çevrimiçi iş ilanları için mesleklerin sınıflandırılması, açık iş ilanlarından beceri içgörülerini çıkarılması ve arama terimleri analizleri modelleri gibi çalışmalar bulunmaktadır. İlgili projeler, kurumun yapay zeka alanındaki bilgi birikimi ve tecrübesini artırırken, veri analizi ve sınıflandırma 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

## Açık Bankacılık Platformu

**Zeynep Yavuz<sup>1\*</sup>, Zeynep Nur Sandıkcı<sup>1</sup>, Önder Demir<sup>2</sup>, Kazım Yıldız<sup>2</sup>**

<sup>1</sup>Sipay Elektronik Para ve Ödeme Hizmetleri Anonim Şirketi, İstanbul, Türkiye

<sup>2</sup>Bilgisayar Mühendisliği, Teknoloji Fakültesi, Marmara Üniversitesi, İstanbul, Türkiye

### Özet

Açık bankacılık, müşterilerin farklı bankalardaki hesaplarını tek bir platform üzerinden yönetmelerini sağlayan yenilikçi bir finansal çözümdür. Dijitalleşmenin finans sektörüne getirdiği hız ve kolaylık, açık bankacılık ile daha ileri bir noktaya taşınmıştır. Bu çalışma, bireysel ve kurumsal müşterilere yönelik bankacılık hizmetlerini bir araya getirilmesini sağlayan bir açık bankacılık platformunun geliştirilmesini ele almaktadır. Platform, kullanıcıların birden fazla banka hesabını yönetmesine ve para transferi, kredi başvuruları, fatura ödemeleri gibi işlemleri tek bir arayüzden gerçekleştirmesine olanak tanımaktadır. Rutin bankacılık işlemlerini otomatikleştirerek insan hatalarını ve operasyonel yükü azaltan bu platform, kullanıcılara ihtiyaçlarına yönelik akıllı finansal çözümler sunmaktadır. Global açık bankacılık standartlarıyla uyumlu olarak tasarlanan platform, gelecekteki genişlemelere ve uluslararası finansal sistemlerle entegrasyonlara uygun, ölçeklenebilir ve modüler bir mimariye sahiptir. Bu çalışma, özellikle gelişmekte olan pazarlarda açık bankacılığın finansal şeffaflığı ve hizmet verimliliğini artırmadaki yenilikçi potansiyelini vurgulamaktadır. Bu çalışma, açık bankacılıkta veri güvenliği, müşteri gizliliği ve operasyonel verimlilik gibi kritik konulara stratejik çözümler geliştirmeyi hedeflemektedir. Müşteri verilerinin güvenilir bir şekilde üçüncü taraflarla paylaşımını sağlayacak güveni API standartlarının oluşturulması ve bu süreçte kullanılan şifreleme yöntemlerinin güçlendirilmesi, temel çözüm önerileri arasında yer almaktadır.

**Anahtar Kelimeler:** Açık Bankacılık, Dijital Dönüşüm, Finansal Teknolojiler

## Panel Data Regression Analysis and Its Application to Real Data

Gökhan Uyar<sup>1\*</sup>, Özlem Kaymaz<sup>1</sup>

<sup>1</sup>Department of Statistics, Ankara University, 06100 Ankara, Türkiye

### Abstract

Panel data analysis is a statistical method widely used in various disciplines such as social science, epidemiology, and econometrics (Yang et al.,2024; Shimada et al. 2023; Shakibaei et al.2021). It was first used in studies by Hildreth (1950) and Kuh (1959). The panel data consists of time series data and cross-sectional data that are used to measure the same unit at different points in time. In other words, it comprises observations from a subset of the same units during a specific time period. If there are T time periods ( $t= 1,2, \dots, T$ ) and N the number of individuals ( $i = 1,2, \dots, N$ ), the panel data will include a total of  $N \times T$  observation units. This analysis has several advantages. Firstly, it examines changes in both cross-sectional and time series data together. Secondly, more observations are produced when cross-sectional and time series data are combined, so increasing the number of observations. Therefore, the degrees of freedom increase, and the multicollinearity problem is reduced. As a result, the efficiency of the estimators improves. This study aims to investigate panel data models and estimation techniques in detail, real data that includes repeated measurements over time on the same observations. The most appropriate regression model, chosen to fit the data, will be used to estimate parameters, and the results will be interpreted.

**Keywords:** Panel Data, Cross-Section, Time Series, Panel Regression Model

## İş İlanlarından Ötesi: Meslek ve Beceri Etiketleme ile Geleceğin İşgücünü Şekillendirmek

**Ahmet Kurnaz<sup>1\*</sup>**

<sup>1</sup>Siyaset Bilimi ve Kamu Yönetimi Bölümü, Çanakkale Onsekiz Mart Üniversitesi, Çanakkale, Türkiye

### Özet

Küresel işgücü açığı, 473 milyon kişi olarak tahmin edilmekte olup, 268 milyon kişi gerekli kriterleri karşılayamadıkları için işgücü dışında kalmaktadır. Bu duruma ek olarak, yetenek açığı son 17 yılın en yüksek seviyesine ulaşmış ve firmaların %77'si istedikleri nitelikte çalışanları bulmakta zorlandıklarını belirtmiştir. Uluslararası İşgücü Örgütü (ILO), bu işgücü açığının kapanmasının önündeki en büyük engellerden birinin nitelik uyumsuzluğu olduğunu vurgulamaktadır. Gelişmiş ekonomilerde hükümetler, ikiz dönüşüm kapsamında işgücü piyasasına yönelik politikalar geliştirmek için emek arzını ve piyasanın beceri taleplerini anlamaya yönelik araçlar geliştirmektedir. Son dönemde bu çabalar, özellikle üretken yapay zekanın ve yeşil mutabakatın işgücü üzerindeki potansiyel etkilerinin incelenmesine odaklanmıştır. Bu politikaların temel amacı, işgücüne gerekli becerilerin kazandırılması yoluyla istihdam kayıplarının önlenmesi, adil geçişin sağlanması ve üretkenliğin artırılmasıdır. Bu bağlamda, nitelik uyumsuzluğunun azaltılması için eğitim kurumları ile istihdam piyasasının talepleri arasındaki ilişkinin güçlendirilmesi büyük önem arz etmektedir. Bu amaca yönelik olarak gerçekleştirilen projeler arasında Avrupa Birliği (AB), Amerika Birleşik Devletleri (ABD) ve OECD merkezli çalışmalar dikkat çekmektedir. Ülkemizde de 12. Kalkınma Planı yükseköğretim ve mesleki eğitimin, işgücü piyasasının beceri talepleri doğrultusunda dönüşümüyle birlikte istihdama katılımı güçlü bir şekilde vurgulamaktadır. Bu doğrultuda Mesleki Alanların Yönelimi Araştırması (MAYA) ile işverenlerin çalışanlarda görmek istedikleri beceri ve yetkinlikleri uluslararası standartlarla uyumlu bir şekilde ortaya çıkararak işgücü, yükseköğretim ve mesleki eğitim politikalarını desteklemesi hedeflenmektedir. Böylece dikey ve yatay nitelik uyumsuzluğunun azaltılarak eğitime ve insan kaynağına yapılan yatırımlar daha verimli hale getirilebilecektir. MAYA projesi, hesaplamalı sosyal bilimler metodolojisine dayanmaktadır. Proje kapsamında, TÜİK tarafından toplanan çevrimiçi iş ilanları R ve Python büyük veri kütüphaneleri, Doğal Dil İşleme (DDİ) yaklaşımları ve Geniş Dil Modelleri (GDM - large language models) kullanılarak işlenmiştir. Analizler sonucunda meslekler Uluslararası Standart Meslek Sınıflandırması (International Standard Classification of Occupations - ISCO-08) ve iş ilan metinleri de Avrupa Beceriler, Yetkinlikler, Nitelikler ve Meslekler Standardı (Europe Skills, Competencies, Qualifications and Occupations - ESCO) içindeki becerilerle uyumlu biçimde etiketlenmiştir. Ayrıca, metin madenciliği teknikleri ile iş ilanları üzerinden mesleklerin ve beceri talebinin zaman içindeki dönüşümü analiz edilmektedir. Çalışma kapsamında, dijital ve yeşil becerilerin yanı sıra, iş veren tarafından aranan fiziksel, bilişsel ve sosyal beceriler de belirlenerek, iş gücü piyasasında hangi tür becerilerin öne çıktığının ve hangi becerilere yönelik talebin arttığının tespit edilmesi amaçlanmaktadır. ESCO'nun yanı sıra dijital, yeşil ve yapay zeka becerileri ile etiketlemeler oluşturulan kapsamlı sözlükler aracılığıyla ve modern yapay zeka teknolojileri kullanılarak yapılmaktadır.

Not: Bu proje TÜBİTAK ARDEB-3005 tarafından desteklenmektedir (Proje No: 124K091).

**Anahtar Kelimeler:** İşgücü, işsizlik, doğal dil işleme, çevrimiçi iş ilanları, beceriler.

## TCMB Açık Veri Platformu: EVDS Geçmişi ve Geleceği

Avni Arıkan<sup>1\*</sup>, Atakan Eren Ölmez<sup>2</sup>

<sup>1</sup>Veri Yönetişimi ve İstatistik Genel Müdürlüğü, TCMB, 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.

## Mega Trendler, Mikro Analizler: Büyük Veri Yaklaşımları

**Ömer Faruk Uyar<sup>1\*</sup>, Ömer Kiraz<sup>1</sup>**

<sup>1</sup>Türkiye İstatistik Kurumu, Ankara, Türkiye

### Özet

Bu bildiri, Türkiye İstatistik Kurumu'nun (TÜİK) büyük veri ve yapay zeka teknolojilerini kullanarak gerçekleştirdiği kurumsal dönüşüm süreçlerini kapsamlı bir şekilde incelemektedir. Özellikle, büyük veri kavramı, hızla büyüyen ve çeşitlenen veri kümeleriyle başa çıkmak için yeni analitik yöntemlere ve teknik altyapılara ihtiyaç duyan bir alan olarak tanımlanmaktadır. Günümüzde, veri kaynaklarının artması ve bu verilerin hızla çoğalması, kurumların veri yönetiminde karşılaştığı zorlukları beraberinde getirmiştir. Bu bağlamda, büyük veri teknolojileri bu tür verilerin analiz edilmesinde ve stratejik karar alma süreçlerinde kritik bir rol oynamaktadır. TÜİK'in yürüttüğü büyük veri ve yapay zeka projeleri, veri toplama, işleme ve analiz süreçlerini dönüştürerek, daha hızlı, daha kapsamlı ve daha güncel istatistiklerin üretilmesine olanak tanımaktadır. Bu projeler sayesinde TÜİK, politika yapıcılara ve karar vericilere daha iyi destek sağlayarak, ekonomik ve sosyal göstergelerin daha doğru ve zamanında analiz edilmesine katkıda bulunmaktadır. Bu çalışmada, TÜİK'in büyük veri ve yapay zeka projeleriyle elde ettiği başarılar ve karşılaştığı zorluklar ele alınarak, dijital dönüşüm süreçlerinin verimliliğe ve etkinliğe olan etkileri değerlendirilmektedir. Özellikle, büyük veri analitiği sayesinde elde edilen içgörüler, Türkiye'nin istatistiksel veri üretiminde önemli bir değişim yaratmıştır. Büyük verinin, geleneksel veri işleme yöntemleriyle yönetilmesi zor olan büyük hacimli, hızlı üretilen ve çeşitli formatlardaki veri kümelerini ifade ettiği göz önüne alındığında, bu veri setleri, yapılandırılmış, yarı yapılandırılmış veya yapılandırılmamış olarak sınıflandırılmaktadır. Büyük veri analitiği sayesinde, istatistiksel verilerin toplama ve işleme süreçlerinde hız kazanarak, verimlilik ve kalite açısından önemli iyileştirmeler sağlanmıştır. Bu sayede hem ulusal hem de uluslararası düzeyde daha etkili politika geliştirme süreçlerine katkıda bulunmakta ve veri tabanlı karar alma mekanizmalarını güçlendirmektedir. Çalışma, TÜİK'in büyük veri ve yapay zeka uygulamalarının potansiyelini ortaya koyarak, gelecekteki veri yönetimi stratejilerine ışık tutmayı amaçlamaktadır. Bu bağlamda, elde edilen bulgular, diğer kamu kurumları ve özel sektör için de önemli dersler içermekte olup, büyük veri uygulamalarının etkili bir şekilde entegrasyonu için öneriler sunulmaktadır.

**Anahtar Kelimeler:** Büyük Veri, Analiz, Dijital dönüşüm, Türkiye İstatistik Kurumu