R&S - RESEARCH STUDIES ANATOLIA JOURNAL

https://dergipark.org.tr/rs | Vol:9 Issue:1; pp:80-89

DERİN ÖĞRENME TABANLI YÖNTEMLER KULLANILARAK E-TİCARET PLATFORMLARINDA ANOMALİ KULLANICI DAVRANIŞLARININ TESPİTİ

Detection of Anomalous User Behaviors in E-Commerce Platforms Using Deep Learning–Based Methods

Serkan METİN[1]

[1]*Dr. Öğr. Üyesi, Malatya Turgut Özal Üniversitesi, Sosyal ve Beşeri Bilimler Fakültesi, Yönetim Bilişim Sistemleri Bölümü, serkan.metin@ozal.edu.tr*, https://orcid.org/0000-0003-1765-7474

Araştırma Makalesi/Research Article

Makale Bilgisi
Geliş/Received: 19.01.2026
Kabul/Accepted: 04.03.2026

DOI:
https://doi.org/rs.1867252

IZLIK:
https://izlik.org/JA49JK98EP

Anahtar Kelimeler
E-ticaret, Anomali Tespiti, Derin Öğrenme

ÖZ

E-ticaret platformlarının hızlı büyümesiyle birlikte, kullanıcı davranışları ve işlem kayıtları üzerinden ortaya çıkabilen anormal durumların tespiti, güvenlik ve operasyonel verimlilik açısından kritik bir gereksinim hâline gelmiştir. Bu çalışmada, e-ticaret sistemlerinde anomali tespiti amacıyla farklı derin öğrenme ve makine öğrenmesi tabanlı yaklaşımlar karşılaştırmalı olarak incelenmiştir. Çalışmada kullanılan veri seti, Kaggle platformundan elde edilen ve 17.049 gözlem ile 18 değişkenden oluşan gerçekçi bir e-ticaret veri tabanıdır. Veri seti; müşteri demografik bilgileri, sipariş detayları, ödeme yöntemleri, teslimat süresi, web sitesi etkileşimleri ve müşteri memnuniyeti gibi hem davranışsal hem de işlemsel nitelikte özellikler içermektedir.

Araştırma kapsamında Autoencoder, LSTM-Autoencoder, Deep Support Vector Data Description (Deep SVDD), GAN tabanlı yeniden yapılandırma yaklaşımı (GAR-AD) ve Isolation Forest olmak üzere beş farklı anomali tespit yöntemi uygulanmıştır. Modeller, AUROC, AUPRC, en iyi F1 skoru ve operasyonel senaryoları yansıtmak amacıyla Top-K (Top %1, %5 ve %10) metrikleri üzerinden değerlendirilmiştir. Deneysel sonuçlar, Autoencoder modelinin yüksek ayrıştırma gücü sağladığını; GAN tabanlı GAR-AD yaklaşımının ise AUPRC ve Best-F1 metrikleri açısından daha dengeli ve tutarlı bir performans sunduğunu göstermektedir. Top-K analizleri, özellikle GAR-AD modelinin sınırlı inceleme kapasitesi senaryolarında yüksek doğruluk ve lift değerleri ürettiğini ortaya koymuştur.

Keywords
E-commerce, Anomaly Detection, Deep Learning

ABSTRACT

With the rapid growth of e-commerce platforms, detecting anomalous transactions and user behaviors from large-scale operational data has become a critical requirement for ensuring system security and operational efficiency. This study presents a comparative analysis of several deep learning– and machine learning–based anomaly detection approaches applied to e-commerce systems. The dataset used in this research was obtained from the Kaggle platform and consists of 17,049 observations with 18 attributes, including customer demographic information, order and payment details, delivery time, website interaction metrics, and customer satisfaction scores. This rich combination of behavioral and transactional features provides a realistic representation of real-world e-commerce environments.

The experimental evaluation includes five anomaly detection methods: Autoencoder, LSTM-Autoencoder, Deep Support Vector Data Description (Deep SVDD), a GAN-based reconstruction approach (GAR-AD), and Isolation Forest. Model performance was assessed using AUROC, AUPRC, best F1 score, and Top-K metrics (Top 1%, 5%, and 10%) to reflect practical investigation capacity constraints in real-world e-commerce operations. The results indicate that the Autoencoder model achieves strong discriminatory performance, while the GAN-based GAR-AD approach provides more balanced and robust results in terms of AUPRC and best F1 score. In particular, the Top-K analysis demonstrates that GAR-AD yields superior precision and lift values under limited inspection budgets.

Atıf/Citation: Metin, S. (2026). Derin öğrenme tabanlı yöntemler kullanılarak e-ticaret platformlarında anomali kullanıcı davranışlarının tespiti, *R&S - Research Studies Anatolia Journal, 9*(1). 80-89.

Sorumlu yazar/Corresponding author: Serkan METİN, serkan.metin@ozal.edu.tr

1. Giriş

İnternetin hızlı gelişimi ve bilgisayar bilimi ile teknolojisindeki sürekli ilerlemelerle birlikte, internet tabanlı çevrim içi ticaret hızla gelişmiş ve günlük yaşamın kolaylığını artırmıştır (Zhao vd., 2019). Günümüzde insanlar geleneksel alışveriş yerine giderek daha fazla internet üzerinden alışveriş yapmaktadır(Hernandez vd., 2017). Bu da e-ticaret kavramının oluşmasını sağlamıştır.

E-ticaret, "electronic commerce" ifadesinin kısaltması olup Türkçeye "elektronik ticaret" olarak çevrilmektedir (Mendonça, 2016). Elektronik ticaret ya da e-ticaret, ürün ve hizmetlerin çevrim içi ortamda alım, satım ve diğer ticari işlemlerini kapsayan geniş bir faaliyet yelpazesini ifade eder (Gupta, 2014). E-ticaret platformları, alıcılar ve satıcılar arasında küresel ölçekte sorunsuz dijital işlemleri mümkün kılarak çevrim içi perakendeyi köklü bir şekilde dönüştürmüştür (Becker, 2025). E-ticaret sistemleri, özellikle son yıllarda teknolojideki büyük ilerleme ve internet üzerinden sunulan hizmetlerdeki gelişmelerle birlikte satış hacminde önemli bir artış yaşamıştır (Abdul vd., 2021). E-ticaretin üstel büyümesi, küresel ekonomiyi tartışmasız biçimde dönüştürmüş; işletmeler ve tüketiciler için benzersiz kolaylık, verimlilik ve ölçeklenebilirlik sunmuştur (Karunaratne, 2023).

Bugün elektronik ticaret, güçlü kavramı ve süreçleriyle modern yaşamı derinden değiştirdiği için olağanüstü bir olgu olarak değerlendirilmektedir. Birçok işletme için e-ticaret bir öncelik hâline gelmiştir. Ürün ve/veya hizmetler için daha düşük maliyetli ve daha işlevsel bir dağıtım süreci sunarak mevcut pazar konumlarını geliştirme ve güçlendirme fırsatı sağlamıştır (Taher, 2021). E-ticaret ile işletmeler, "Tam Zamanında Üretim" ve "Tam Zamanında Teslimat" sistemlerini ticaret ortaklarıyla entegre ederek dünya çapındaki stratejik yetkinliklerini artırmaktadır (Jain vd., 2021).

E-ticaretin popülaritesi artmaya devam ettikçe, e-ticaret siparişlerindeki anormal verilerin incelenmesi giderek daha önemli hale gelmektedir (Wang vd., 2023). E-ticaret faaliyetlerinde anomali tespiti uygulaması, e-ticaret finansal faaliyetlerinde dolandırıcılık yapan hesapların belirlenmesi, anormal kullanıcıların hızlı ve doğru biçimde tanımlanması, anormal davranışların uyarılması veya durdurulması açısından kritik rol oynamaktadır (Wu vd., 202).

Bu kapsamda anomali tespiti için veri madenciliği, makine öğrenmesi ve derin öğrenme yöntemleri yaygın olarak kullanılmaktadır. Derin öğrenme modelleri, güçlü özellik öğrenme yetenekleri sayesinde büyük miktardaki karmaşık ve yapılandırılmamış veriden otomatik olarak yüksek seviyeli soyut özellikler çıkarabilmekte, böylece anomali tespitinin doğruluğunu ve sağlamlığını artırmaktadır (Li vd., 2024).

Bu makalenin geri kalanı şu şekilde düzenlenmiştir: İkinci bölümde, anomali kavramı ve tespit yöntemleri ele alınmıştır. Üçüncü bölümde, anomali tespitine yönelik literatürdeki çalışmalar incelenmiştir. Dördüncü bölümde, yöntemler ve bulgulardan elde edilen sonuçlar verilmiştir. Sonuç bölümünde, çalışmanın sonuçlarını ve gelecekteki araştırmalar için önerilerde bulunulmuştur.

2. Anomali Kavramı ve Tespit Yöntemleri

Anomaliler, normal örneklerden farklı veri özelliklerine sahip veri örüntüleridir (Liu vd., 2012). Anomaliler her zaman bir saldırı olarak sınıflandırılamaz; daha önce bilinmeyen, şaşırtıcı bir davranış da olabilir. Zararlı olabilir ya da olmayabilir. (Agrawal ve Agrawal, 2015). Anomali tespiti, verilen bir veri kümesindeki anormal veya olağandışı verileri belirleyen önemli bir veri analizi görevidir. Veri madenciliği araştırmalarında ilgi çekici bir alan olup, verideki nadir ve dikkat çekici örüntülerin keşfedilmesini içerir (Mohiuddin vd, 2016).

Anomali tespitinin önemli olmasının nedeni, anomalilerin dikkat çekici olması ve çoğu uygulama alanında ilgi çekici bilgiler içermesidir. Anomali tespit teknikleri; video gözetimi, kredi kartı dolandırıcılığı tespiti, yüzey kusur tespiti, tıbbi tanı gibi çok geniş bir uygulama alanına sahiptir (Minhas ve Zelek, 2019). E-ticaret bağlamında bu anomaliler, dolandırıcılık faaliyetleri, şüpheli kullanıcı davranışları, düzensiz işlemler veya sistemlerin yanlış kullanımı gibi durumları kapsayabilir (Mah vd., 2025).

Anomaliler üç ana kategoriye ayrılır. Eğer tek bir veri örneği, veri kümesinin geri kalanına göre anormal kabul edilebiliyorsa, bu örnek noktasal anomali olarak adlandırılır ve en basit anomali türü olarak değerlendirilir. Eğer bir veri örneği belirli bir bağlamda anormal, ancak başka bir bağlamda normal ise, buna bağlamsal anomali denir. Eğer ilişkili veri örneklerinden oluşan bir küme, tüm veri kümesi açısından anormal ise, buna kolektif anomali denir (Nassif vd. 2021). Mevcut anomali tespiti yaklaşımları beş kategoriye ayrılabilir: olasılık temelli, mesafe temelli, yeniden yapılandırma temelli, alan (domain) temelli ve bilgi kuramı temelli (Wang, vd., 2018).

E-ticarette anomali tespiti için kullanılan geleneksel yaklaşımlar çoğunlukla kural tabanlı sistemlere veya tek tek ürün özelliklerini ya da kullanıcı davranışlarını izole biçimde analiz eden basit makine öğrenmesi modellerine dayanmaktadır (Shao vd., 2025). E-ticaret bağlamında, büyük ölçekli gerçek zamanlı veri üretilmektedir (Jiang ve Yu, 2008). Ancak bu yöntemler, e-ticaret ekosistemlerinin bu karmaşık ve büyük ölçekteki verilerden örüntü yakalamakta sıklıkla yetersiz kalmaktadır. Anomali tespiti, veri madenciliği, makine öğrenmesi, bilgisayarla görme ve istatistik gibi çeşitli topluluklarda giderek daha önemli roller üstlenmektedir (Pang vd., 2021). Veri madenciliği, bu verilerdeki önceden bilinmeyen örüntüleri tespit ederek veri kümelerindeki gizli bilgilerin ortaya çıkarılmasında ve böylece daha etkili karar vermede yardımcı olur. Veri madenciliği büyük veri kümelerinden yararlı bilgileri ortaya çıkarmak için kullanılır. Anlamsız devasa veri kümelerinde gizli örüntüleri bulmak için veri analizi, tahmin ve sınıflandırma için farklı yaklaşımlar kullanılmaktadır (Tanır ve Ramazanoz, 2023).

3. Literatür Taraması

Sakthivanitha vd. (2025), e-ticaret sistemlerinin güvenliğini ve operasyonel etkinliğini artırmak amacıyla derin öğrenme (DL) kullanan gerçek zamanlı anomali tespiti yaklaşımlarını incelemektedir. Önerilen çerçeve; Autoencoder (AE), Uzun Kısa Süreli Bellek (LSTM) modelleri ve Grafik Sinir Ağlarından (GNN) oluşmaktadır. Topluluk (ensemble) yaklaşımı, bu üç model türünün her birinden elde edilen anormallik çıktıları birleştirerek genişletilmiş bir gradyan artırma sınıflandırıcısı aracılığıyla bütünleşik anomali skorları üretir. Bu topluluk modeli, Kaggle Online Retail Veri Seti üzerinde değerlendirilmiş ve doğruluk %96, kesinlik (precision) %92, duyarlılık (recall) %89 ve F1-Skoru %90 ile diğer bireysel modellere kıyasla tüm ölçütlerde en yüksek hizmet kalitesi metriklerini sağlamıştır.

Bozbura vd. (2019), e-ticaret anormalliklerinin tespitinde Hareketli Ortalama, Otoregresif Entegre Hareketli Ortalama (ARIMA), Kalman Filtresi, Zaman Serisi Ayrıştırması, Holt-Winters yöntemi, Markov Geçişli Model (MSM) ile LSTM'yi birleştiren hibrit algoritma (MSM+LSTM), durum bilgili (stateful) LSTM ve durum bilgisiz (stateless) LSTM yöntemlerini etiketli veri kümeleri üzerinde uygulamışlardır.

Sun (2025), kullanıcı davranışsal özellikleri ve makine öğrenmesi yöntemlerini kullanarak e-ticaret iade tahmini üzerine kapsamlı bir inceleme yapmıştır. Çalışma, farklı ürün kategorilerinde iade olasılığını tahmin etmek amacıyla tüketici etkileşim kalıplarını, satın alma geçmişini ve demografik faktörleri analiz eden bir tahmin çerçevesi geliştirmektedir. Gerçek dünya e-ticaret veri kümeleri üzerinde yapılan kapsamlı deneyler aracılığıyla, rastgele orman (random forest), gradyan artırma (gradient boosting) ve sinir ağları dâhil olmak üzere birçok makine öğrenmesi algoritması değerlendirilmiştir. Sonuçlar, önerilen yaklaşımın iade tahmininde %89,3 doğruluk sağladığını ve temel yöntemlere kıyasla yanlış pozitif oranlarını %23 azalttığını göstermektedir.

Bu makalede Mawlud ve Ahmad (2018), anomali tespiti için veritabanı yapısı modellenmiş ve bu yapıya dayalı olarak daha verimli bir anormal veritabanı tespit yaklaşımı olan (PMK-means) algoritması önerilmiştir.

Sharma vd., (2020), anomali tespiti için denetimsiz kullanıcı davranışı modellemesini sunmaktadır. Önerilen yaklaşım, oturum etkinliklerine dayalı olarak kullanıcı davranışını modellemek ve anormal veri noktalarını belirlemek amacıyla LSTM tabanlı bir Otomatik Kodlayıcı (Autoencoder) kullanmaktadır. Önerilen yöntem iki aşamalı bir süreci izlemektedir. İlk olarak, anormal olmayan veri kümesi üzerinde autoencoder kullanılarak yeniden yapılandırma hatası hesaplanmakta, ardından bu hata değeri, aykırı değerleri normal veri noktalarından ayırmak için eşik değer olarak kullanılmaktadır. Belirlenen aykırı değerler daha sonra anomali olarak sınıflandırılmaktadır. Deneysel sonuçlar, en iyi durumda modelin %90,17 doğruluk, %91,03 doğru pozitif ve

%9,84 yanlış pozitif oranı ürettiğini göstermektedir. Bu sonuçlar, önerilen yaklaşımın otomatik anomali tespitinde etkili bir şekilde kullanılabileceğini ortaya koymaktadır.

Bu araştırmada Jiang vd., (2025), bulut altyapılarında gerçek zamanlı kullanıcı davranışı anomali tespiti ve proaktif tehdit erken uyarısı için yeni bir derin öğrenme çerçevesi sunmaktadır. Önerilen sistem, normal kullanıcı kalıplarını modellemek ve güvenlik ihlallerine işaret eden sapmaları belirlemek amacıyla dikkat mekanizmalarıyla birlikte LSTM tabanlı zamansal dizi analizini entegre etmektedir. Çok kiracılı bulut veri kümeleri üzerinde yapılan kapsamlı değerlendirmeler, anormal davranışların tespitinde %96,8 doğruluk, %94,2 kesinlik (precision) ve %95,7 duyarlılık (recall) ile üstün bir performans sergilendiğini göstermektedir.

Becker (2025), kullanıcı etkileşimleri, işlemler ve yorum ağlarının grafik yapısından yararlanarak e-ticaret platformlarında dolandırıcılık tespitini geliştirmek amacıyla grafik sinir ağı (GNN) tabanlı bir anomali tespit çerçevesi önermektedir. Gerçek dünya e-ticaret veri kümeleri üzerinde yapılan deneyler, GNN tabanlı modelin F1 skoru, kesinlik (precision), duyarlılık (recall) ve yanlış pozitif oranının azaltılması açısından geleneksel dolandırıcılık tespit yaklaşımlarından daha iyi performans gösterdiğini ortaya koymaktadır. Çerçeve, 0,91 F1 skoru ile anormal faaliyetleri başarıyla tespit etmiştir.

Zhao vd., (2019), e-ticaret işlem sistemlerinde hem sistem davranışını hem de kullanıcı davranışını analiz etmeye odaklanmaktadır. Kullanıcı davranışları dört gruba ayrılmıştır: web gezinme davranışı, klavye vuruş (tuşlama) davranışı, ağ işlemi davranışı ve mobil terminal davranışı. Bu yöntemlerin performans ve korelasyon analiz teknikleri ayrıntılı biçimde tartışılmış ve karşılaştırılmıştır. Son olarak, çevrim içi işlem sistemleri için davranış analizi alanındaki gelecekteki araştırmalara yönelik fikirler sunulmuştur.

Lai ve Yang (2023), çevrimiçi davranış anomalilerini tespit etmek ve incelemek için yeniden yapılandırılabilir, birden fazla derin öğrenme zaman serisi tahmin modeli içeren Olay Tabanlı Davranış Analizi (Event-Based Behavior Analysis, EBBA) adlı bir çerçeve önermektedir. Taiwan'daki birinci seviye bir e-ticaret şirketinin gerçek erişim günlükleri kullanılarak eğitilen Tekrarlayan Sinir Ağları (RNN) ve Uzun/Kısa Süreli Bellek (LSTM) içeren bir EBBA prototipi geliştirilmiştir. Ön deneysel sonuçlar, RNN ve LSTM modellerinin çevrimiçi kötüye kullanım olaylarının tahmininde sırasıyla %86,4 ve %92,2 doğruluk oranlarına, ayrıca sırasıyla sıfır ve %0,0385 gibi son derece düşük yanlış izin verme oranlarına ulaştığını göstermektedir.

Li vd., (2024), varyasyonel otokodlayıcı (VAE) tabanlı bir e-ticaret veri anomali tespit yöntemi önermektedir. VAE, verinin potansiyel dağılımını öğrenerek özellikleri etkili biçimde çıkarabilmekte ve anormal verileri tanımlayabilmektedir. Deneysel sonuçlar, önerilen yöntemin özellikle doğruluk, duyarlılık (recall), F1 skoru ve ROC-AUC gibi temel performans göstergelerinde geleneksel yöntemlerden daha iyi sonuçlar verdiğini göstermektedir.

Bi vd., (2016), Temel Bileşenler Analizi (PCA) tabanlı yeni bir anomali tespit modeli önerilmektedir. Önerilen şema, kullanıcının davranışlarını ortaya çıkarmaya yönelik bir yöntem sunmakta ve kullanıcının erişimini temsil eden seçilmiş özellikleri analiz etmektedir. PCA yöntemi, anomali tespit modeline entegre edilerek, sistemi tasarım açısından daha tutarlı hale getirmekte, kullanıcının davranışlarını daha kapsamlı biçimde tanımlamakta ve algoritmanın verimliliğini ile kararlılığını artırmaktadır.

Wang vd. (2018), üreteci karşıt ağlar (GAN) tabanlı yeni bir alan tabanlı anomali tespit yöntemi önermektedir. Üreticinin daha fazla anomali üretmesini sağlamak ve normal veri dağılımına yakınsamasını önlemek için minimum olasılık düzenlemesi (minimum likelihood regularization) önerilmektedir. Anomali skorlarının uygun şekilde birleştirilmesinin, ayrıştırıcının (discriminator) kararlılığını etkili bir şekilde artırdığı gösterilmiştir. Önerilen yöntem, Cifar10 ve UCI veri setlerinde diğer anomali tespit yöntemlerine göre önemli iyileşmeler sağlamıştır.

4. Araştırmanın Metdolojisi

4.1. Araştırmanın konusu, amacı ve problemi

Araştırmanın konusu, bu çok boyutlu müşteri ve sipariş verileri üzerinde anomali tespiti yöntemleri kullanılarak olağandışı davranış ve işlemlerin belirlenmesidir. Araştırmanın amacı, potansiyel sahtekârlık girişimleri, sıra dışı satın alma kalıpları, olağan dışı indirim kullanımları, aşırı sipariş tutarları veya düşük memnuniyetle sonuçlanan riskli işlemler gibi anormal durumları tespit etmektir. Araştırmanın problemi ise, büyük hacimli ve çok değişkenli e-ticaret verileri içerisinde geleneksel istatistiksel yöntemlerle kolayca fark edilemeyen anomalilerin makine öğrenmesi tabanlı anomali tespiti yaklaşımlarının kullanılmasıyla tespitinin yapılmasıdır.

4.2. Evren, örneklem ve ölçek geliştirme

Çalışmanın evrenini, Kaggle veritabanında bulunan https://www.kaggle.com/datasets/umuttuygurr/e-commerce-customer-behavior-and-sales-analysis-tr?select=ecommerce_customer_behavior_dataset_v2.csv adresinden indirilen veri seti oluşturmaktadır.

Veri seti; müşteri demografik özellikleri (yaş, cinsiyet, şehir), sipariş bilgileri (ürün kategorisi, birim fiyat, satın alınan adet, indirim tutarı, toplam sipariş tutarı), ödeme ve erişim kanalları, teslimat süresi, site etkileşimleri (görüntülenen sayfa sayısı, tekrar gelen müşteri durumu) ve müşteri memnuniyet puanı gibi 18 değişken ve 17.049 gözlem içermektedir. Veri setini oluşturan sayısal ve kategorik sütunlara ait temel istatistiksel değerler Tablo 1 ve Tablo 2'de verilmiştir.

Tablo 1. Sayısal Değişkenler Ait Tanımlayıcı İstatistikler

Değişken	Ortalama	Std. Sapma	Min	25%	Medyan	75%	Max
Müşteri Yaşı	39.9	11.5	18	30	40	50	65
Birim Fiyat	120.5	68.3	10	65	115	175	350
Satın Alınan Adet	2.9	1.7	1	1	3	4	10
İndirim Tutarı	24.6	18.2	0	10	20	35	120
Toplam Sipariş Tutarı	325.8	240.7	30	140	290	460	1800
Oturum Süresi (dk)	12.5	6.1	1	8	12	17	45
Görüntülenen Sayfa Sayısı	6.8	3.4	1	4	6	9	25
Müşteri Memnuniyet Puanı	4.1	0.7	1	4	4	5	5

Tablo 2. Kategorik Değişkenler Ait İstatistikler

Şehir	Sayı	Tekrar Gelen Müsteri	Sayısı	Ödeme Yöntemi	Sayı
İstanbul	4.402	Evet	15039	Kredi Kartı	6801
Ankara	2.422	Hayır	2010	Dijital Cüzdan	3276
İzmir	2.072	**Ürün Kategorisi**	**Sayı**	Banka Transferi	1763
Bursa	1.721	Ev & Bahçe	2060	Kapıda Ödeme	888
Adana	1.326	Kitaplar	2206	Banka Kartı	4321
Antalya	1.246	Güzellik	2212	**Kullanılan Cihaz Türü**	**Sayı**
Gaziantep	1.183	Spor	2248	Mobil	9543
Konya	1.120	Giyim	2103	Masaüstü Bilgisayar	5845
Kayseri	851	Electronik	2074	Tablet	1661
Eskişehir	706	Moda	2056	**Cinsiyet**	**Sayı**
		Oyuncaklar	2090	Erkek	8436
				Kadın	8613

4.3. Verilerin Analizi

Çalışmada veriler üzerinde analiz yapmak için Isolation Forest, Autoencoder, LSTM Autoencoder, GAR-AD ve DeepSVDD yöntemleri kullanılmıştır.

Isolation Forest Algoritması, denetimsiz bir makine öğrenmesi algoritması olarak, örnek veri içindeki anormal verileri tespit eder ve izole noktaları belirlemek için mesafe veya yoğunluk hesaplamasına gerek duymaz. Algoritma, örnek verinin dağılımını ve yoğunluğunu analiz ederek bir örneğin izole olup olmadığını doğru şekilde

belirler (Zhang ve Liu, 2022). Isolation Forest, veriyi ikili ağaç yapısında olduğu gibi iki parçaya böler; bölme işlemi ya belirli bir ağacın yüksekliği sınırına ulaşana kadar ya da veri daha fazla bölünemeyene kadar devam eder (Al Farizi vd., 2021).

Autoencoder, normal veri örüntülerini öğrenerek yeniden yapılandırma hatası üzerinden anomalileri tespit eden denetimsiz bir derin öğrenme yöntemidir. Model, denoising yapıda eğitilmiş olup, giriş verilerine rastgele gürültü eklenerek orijinal verinin yeniden üretilmesi hedeflenmiştir. Bu yaklaşım, modelin yalnızca gürültüsüz ve normal örüntüleri öğrenmesini teşvik eder.

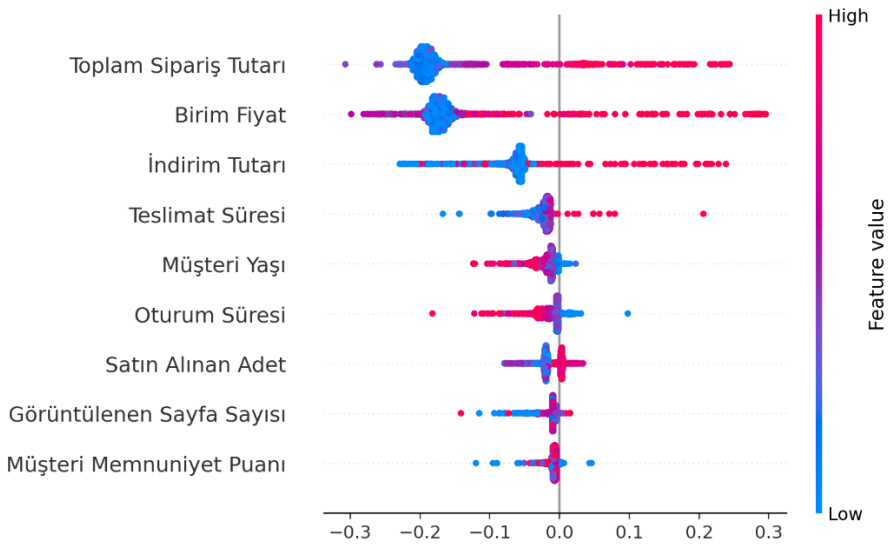
LSTM-autoencoder'lar uzun vadeli bağımlılıkları yakalayabilme ve bağlamsal bilgiyi modelleyebilme yeteneğine sahiptir (Lachekhab vd., 2024). LSTM Tabanlı Autoencoder, özellikle zamansal veya sıralı veriler için tasarlanmış olup, uzun vadeli bağımlılıkları yakalayabilen Long Short-Term Memory (LSTM) mimarisini kullanmaktadır. Bu yapı, özellikle kullanıcı davranışları veya işlem sıraları gibi zamana bağlı kalıpların bulunduğu senaryolarda anomali tespiti için uygundur.

Deep Support Vector Data Description (Deep SVDD), verinin büyük çoğunluğunun normal olduğu varsayımıyla, derin bir ağ aracılığıyla tüm normal örnekleri özellik uzayında tek bir merkez etrafında yoğunlaştırmayı hedefleyen denetimsiz bir anomali tespit yöntemidir. Eğitim sırasında, ağ çıktılarının bu merkeze olan uzaklığı optimize edilir. Test aşamasında ise, merkezden uzaklığı yüksek olan örnekler anomali olarak sınıflandırılır.

GAN tabanlı yeniden yapılandırma yaklaşımı (GAR-AD), bu yöntemde generator ağı, her bir girdi örneğini yeniden üretirken; discriminator ağı, gerçek örnekler ile yeniden yapılandırılmış örnekleri ayırt etmeyi öğrenmektedir. Generator, adversarial kayıp ile L2 yeniden yapılandırma kaybından oluşan birleşik bir amaç fonksiyonu ile eğitilmektedir. Eğitim tamamlandıktan sonra, anomali skorları standartlaştırılmış özellikler üzerinde hesaplanan yeniden yapılandırma hatası olarak tanımlanmakta olup, daha yüksek hata değerleri daha yüksek anomali olasılığını göstermektedir.

Çalışmada modellerin uygulanmasından önce anomali skorlarının açıklanabilirliğini sağlamak amacıyla skorları hedef değişken olarak kullanan bir surrogate model (Random Forest Regressor) eğitilmiştir. Bu model üzerinden SHAP (SHapley Additive exPlanations) yöntemi uygulanarak her bir özelliğin anomali skoruna katkısı hesaplanmıştır. Şekil 1'de, her bir değişkenin etki dereceleri gösterilmektedir.

Şekil 1. Bağımsız Değişkenlerin Model Çıkışına Etki Oranları



Veri üzerine uygulanan yöntemlere ait performans metrik değerleri Tablo 3'de verilmiştir.

Tablo 3. Model Performans Değerleri

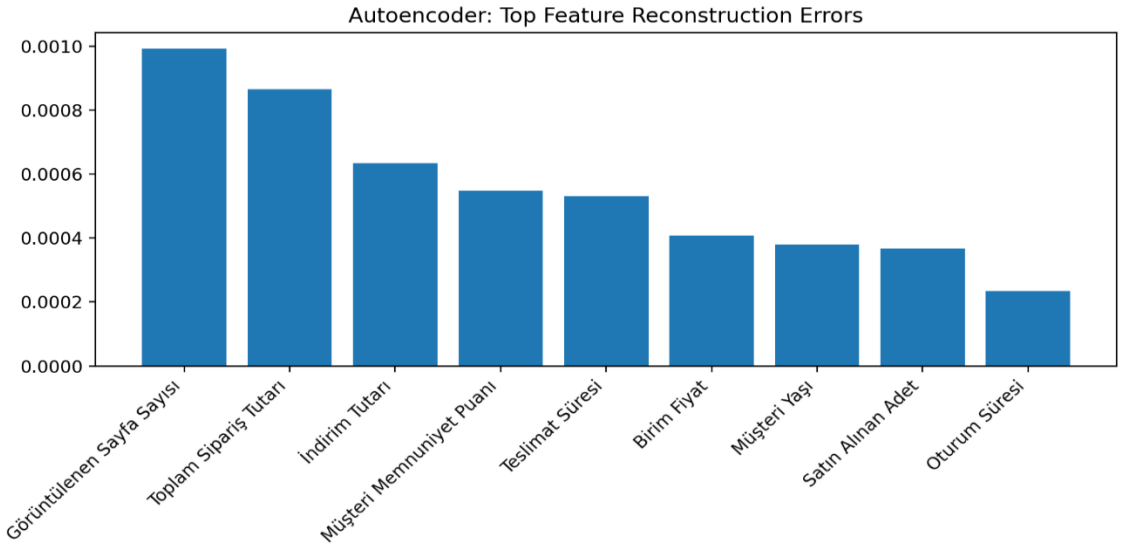
Model	AUROC	AUPRC	BestF1	K_frac	K	Precision_K	Recall_K	F1_K	Lift_K
Autoencoder	0,90	0,62	0,60	0,01	170,00	0,96	0,32	0,48	32,19
	0,90	0,62	0,60	0,05	852,00	0,39	0,65	0,49	13,00
	0,90	0,62	0,60	0,10	1704,00	0,22	0,74	0,34	7,38
Deep SVDD	0,85	0,54	0,55	0,01	170,00	0,98	0,32	0,49	32,58
	0,85	0,54	0,55	0,05	852,00	0,33	0,55	0,41	11,00
	0,85	0,54	0,55	0,10	1704,00	0,20	0,65	0,30	6,54
GAR-AD	0,88	0,65	0,66	0,01	170,00	1,00	0,33	0,50	33,36

	0,88	0,65	0,66	0,05	852,00	0,39	0,66	0,49	13,16
	0,88	0,65	0,66	0,10	1704,00	0,22	0,72	0,33	7,24
Isolation Forest	0,67	0,05	0,12	0,01	170,00	0,06	0,02	0,03	2,16
	0,67	0,05	0,12	0,05	852,00	0,06	0,09	0,07	1,84
	0,67	0,05	0,12	0,10	1704,00	0,07	0,23	0,11	2,35

Tablo 3 incelendiğinde, Autoencoder'ın AUROC açısından en yüksek ayrıştırma gücünü sağladığını (AUROC=0.8977) göstermektedir. Bununla birlikte, dengesiz sınıf dağılımına sahip anomali tespiti problemlerinde daha anlamlı kabul edilen AUPRC metriği açısından en yüksek performans GAR-AD modeli tarafından elde edilmiştir (AUPRC=0.6513). Benzer şekilde, eşik seçimine dayalı en iyi F1 skoru (BestF1=0.6609) da GAR-AD yönteminden elde edilmiştir.

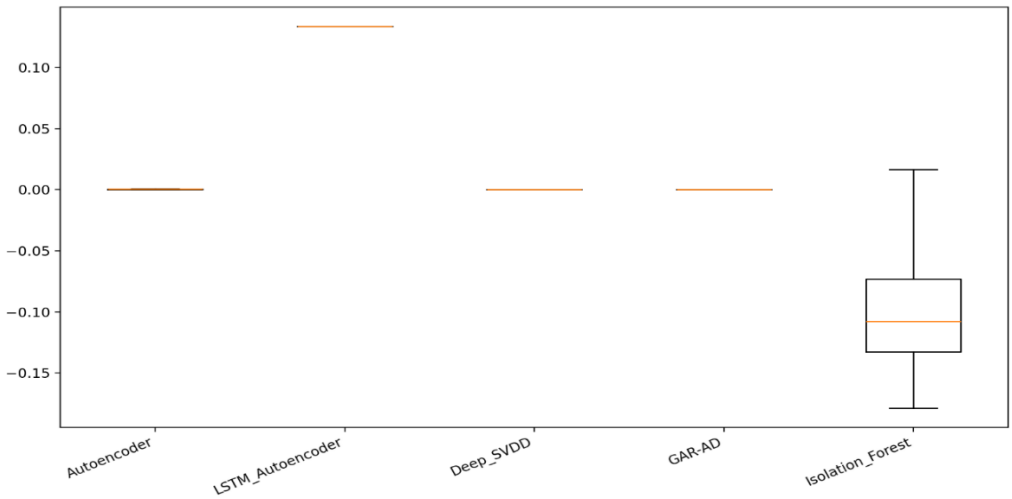
Kfrac, Top-K analizinde kullanılan veri oranını ifade eder. E-ticarette anomali tespitinden çoğunlukla en riskli %1, %5 ve %10'luk siparişler incelenir. Bu yüzden Precision_K ve Lift_K değerleri kritik öneme sahiptir. Top-K analizinde GAR-AD özellikle Top %1 ve Top %5 dilimlerinde en yüksek precision ve lift değerlerine sahiptir. Top %10'da ise Autoencoder daha yüksek anomali yakalama oranı ulaşmıştır. Bu bulgular eşliğinde, GAR-AD'in e-ticaret anomali tespiti uygulamalarında daha dengeli sonuçlar ürettiği görülmüştür.

Şekil 2. Autoencoder için Özellik Bazlı Yeniden Yapılandırma Hatası



Şekil 2, Autoencoder modelinin her bir sayısal değişken için ürettiği ortalama yeniden yapılandırma hatalarını göstermektedir. Yeniden yapılandırma hatası, modelin normal davranışı öğrenirken hangi değişkenlerde daha fazla zorlandığını ve dolayısıyla anomali skorunu hangi özelliklerin daha güçlü biçimde etkilediğini ortaya koymaktadır. Şekil 2'de görüldüğü üzere Görüntülenen Sayfa Sayısı değişkeni, en yüksek yeniden yapılandırma hatasına sahiptir.

Şekil 3. Anomali Tespit Modelleri İçin Anomali Skoru Dağılımları



Şekil 3, Autoencoder, LSTM-Autoencoder, Deep SVDD, GAR-AD ve Isolation Forest modellerinin ürettiği anomali skorlarının dağılımlarını kutu grafiği aracılığıyla karşılaştırmalı olarak göstermektedir. Grafik incelendiğinde, modeller arasında hem skor ölçeği hem de dağılım genişliği açısından belirgin farklılıklar olduğu görülmektedir. Autoencoder, Deep SVDD ve GAR-AD modelleri için skorların medyan değerleri yaklaşık olarak 0.00 seviyesinde ve çeyrekler arası aralık (IQR) oldukça dar kalmaktadır. Bu modellerde skorların büyük bir bölümü yaklaşık −0.002 ile 0.002 aralığında yoğunlaşmıştır. Bu durum, veri setindeki örneklerin büyük çoğunluğunun modeller tarafından normal olarak değerlendirildiğini ve anomalilerin yalnızca dağılımın üst kuyruğunda yer alan az sayıdaki gözlem üzerinden ayrıştığını göstermektedir. LSTM-Autoencoder modelinde ise skorların medyan değeri yaklaşık 0.12–0.13 civarındadır. Buna karşılık Isolation Forest modelinin skor dağılımı belirgin biçimde daha geniştir. Medyan skor değeri yaklaşık −0.11 civarındayken, alt ve üst çeyrekler yaklaşık olarak −0.14 ile −0.08 aralığında yer almaktadır. Ayrıca skorların minimum değeri yaklaşık −0.18 seviyesine kadar düşerken, maksimum değerlerin 0.01–0.02 civarına ulaştığı gözlemlenmektedir. Bu geniş dağılım, Isolation Forest'ın örnekler arasında daha yüksek bir ayrım ürettiğini göstermektedir. Modeller 17.049 kayıt arasından 853 anormal işlem belirlemiştir. Anomali olarak belirlenen işlemlere ait ortak özellikler Tablo 4'de verilmiştir.

Tablo 4. Anomali Davranış Gösteren Müşterilerin Ortak Özellikleri

Özellik	Genel Ortalama	Anomali Ortalama
Müşteri Yaşı	34.95	35.86
Birim Fiyat	447.90	1288.46
Satın Alınan Adet	3.01	3.02
İndirim Tutarı	69.79	294.67
Toplam Sipariş Tutarı	1277.44	3988.09
Oturum Süresi	14.54	13.94
Görüntülenen Sayfa Sayısı	9.00	9.11
Teslimat Süresi	6.50	7.88
Müşteri Memnuniyet Puanı	3.90	3.78

5. Sonuç ve Öneriler

Bu çalışmada, e-ticaret sistemlerinde ortaya çıkan anormal işlem ve kullanıcı davranışlarının tespiti amacıyla, Kaggle veritabanından indirilen 17.049 gözlem ve 18 değişkenden oluşan, hem davranışsal hem de işlemsel özellikleri içeren veriseti kullanımıştır. Bu çok boyutlu yapı, özellikle yüksek sınıf dengesizliği ve karmaşık örüntülerin bulunduğu anomali tespiti problemleri için uygun bir test ortamı sunmaktadır. Kullanılan yöntemlerin performanslarının incelendiğinde, Autoencoder tabanlı yaklaşım, AUROC=0,90 değeri ile normal ve anormal örnekleri ayırt etme açısından en yüksek ayrıştırma oranına ulaşmıştır. AUPRC ve Best-F1 metrikler dikkate alındığında, GAR-AD (GAN tabanlı yeniden yapılandırma) yaklaşımının daha dengeli bir performans sergilediği gözlemlenmiştir. GAR-AD modeli, AUPRC=0,65 ve Best-F1=0,66 değerleri ile hem yanlış pozitifleri sınırlayabilmiş hem de anormal örneklerin yakalanmasında daha istikrarlı sonuçlar üretmiştir. Top %1 ve Top %5 dilimlerinde GAR-AD modelinin Precision_K ve Lift_K metrikleri açısından diğer yöntemlere kıyasla daha başarılı sonuçlar elde ettiği görülmüştür. Isolation Forest yöntemi ise tüm metriklerde görece düşük performans sergilemiş ve bu veri seti bağlamında derin öğrenme tabanlı yöntemlere kıyasla sınırlı bir ayırt edicilik sunmuştur.

Gelecek çalışmalarda, anomali tespit modellerinin çıktılarının sabit ağırlıklar yerine öğrenilebilir mekanizmalarla bütünleştirilmesi model doğruluk oranlarının artırılmasına olanak sağlayabilir. Müşteri–ürün–sipariş etkileşimlerinin açık biçimde temsil edildiği grafik tabanlı modellerin (GNN, graph autoencoder) kullanılması da anaomali tespitinde önemli başarılar elde edebilir.

Kaynakça

Abdul Hussien, F. T., Rahma, A. M. S., & Abdulwahab, H. B. (2021). An e-commerce recommendation system based on dynamic analysis of customer behavior. *Sustainability,* 13(19), 10786. https://doi.org/10.3390/su131910786

Agrawal, S., & Agrawal, J. (2015). Survey on anomaly detection using data mining techniques. *Procedia Computer Science,* 60, 708–713. https://doi.org/10.1016/j.procs.2015.08.220

Al Farizi, W. S., Hidayah, I., & Rizal, M. N. (2021). Isolation Forest Based Anomaly Detection: *A Systematic Literature Review. 2021 8th International Conference on Information Technology, Computer and Electrical Engineering (ICITACEE),* Semarang, Indonesia, 118-122, https://doi.org/10.1109/ICITACEE53184.2021.9617498.

Becker, L. (2025). Anomaly detection in e-commerce platforms via graph neural networks. *Asia Pacific Economic and Management Review*, 2(2), 1–10. https://doi.org/10.62177/apemr.v2i2.208

Bi, M., Xu, J., Wang, M., & Zhou, F. (2016). Anomaly detection model of user behavior based on principal component analysis. *Journal of Ambient Intelligence and Humanized Computing*, 7, 547–554. https://doi.org/10.1007/s12652-015-0341-4

Bozbura, M., Tunc, H. C., Kusak, M. E., & Sakar, C. O. (2019). Detection of e-commerce anomalies using LSTM-recurrent neural networks. In Proceedings of the 8th International Conference on Data Science, *Technology and Applications*, 217–224. https://doi.org/10.5220/0007924502170224

Gupta, A. (2014). E-commerce: Role of e-commerce in today's business. *International Journal of Computing and Corporate Research*, 4(1), 1–8.

Hernandez, S., Alvarez, P., Fabra, J., & Ezpeleta, J. (2017). Analysis of users' behavior in structured e-commerce websites. *IEEE Access*, 5, 11941–11958. https://doi.org/10.1109/ACCESS.2017.2707600

Jain, V., Malviya, B., & Arya, S. (2021). An overview of electronic commerce (e-commerce). *Journal of Contemporary Issues in Business and Government,* 27(3), 665–670. https://doi.org/10.47750/cibg.2021.27.03.090

Jiang, X., Jia, R., & Zhang, F. (2025). Deep learning-based user behavior anomaly detection and threat early warning in cloud computing environments. *Academia Nexus Journal,* 4(3), 1–24.

Jiang, Y., & Yu, S. (2008). Mining e-commerce data to analyze the target customer behavior. *In First International Workshop on Knowledge Discovery and Data Mining*, 406–409. https://doi.org/10.1109/WKDD.2008.90

Karunaratne, T. (2023). Machine learning and big data approaches to enhancing e-commerce anomaly detection and proactive defense strategies in cybersecurity. *Journal of Advances in Cybersecurity Science, Threat Intelligence, and Countermeasures,* 7(12), 1–16.

Lachekhab, F., Benzaoui, M., Tadjer, S.A., Bensmaine, A., & Hamma, H. (2024). LSTM-Autoencoder Deep Learning Model for Anomaly Detection in Electric Motor. *Energies,* 17, 2340. https://doi.org/10.3390/en17102340

Lai, A. I. C., & Yang, H. C. (2023). Deep learning based behavior anomaly detection within the context of electronic commerce. *In 2023 IEEE International Conference on Intelligence and Security Informatics (ISI),* 1–6. https://doi.org/10.1109/ISI58743.2023.10297230

Li, J., Liu, S., & Zou, J. (2024). E-commerce data anomaly detection method based on variational autoencoder. *In 2024 3rd International Conference on Artificial Intelligence, Internet of Things and Cloud Computing Technology (AIoTC),* 231–234. https://doi.org/10.1109/AIoTC63215.2024.10748279

Liu, F. T., Ting, K. M., & Zhou, Z. H. (2012). Isolation-based anomaly detection. *ACM Transactions on Knowledge Discovery from Data,* 6(1), 1-39. https://doi.org/10.1145/2133360.2133363

Mah, P. M., Skalna, I., & Pelech-Pilichowski, T. (2025). AI-driven anomaly detection in e-commerce services: A deep learning and NLP approach to the isolation forest algorithm trees. *Journal of Theoretical and Applied Electronic Commerce Research*, 20(214), 1–32. https://doi.org/10.3390/jtaer20030214

Mawlud, A. T., & Ahmad, S. S. (2018). E-commerce transactions of behaviors anomaly detection system. *Iraqi Journal for Information Technology*, 9(1), 137–152.

Mendonça, H. G. (2016). E-commerce. *Journal of Innovation, Projects and Technologies*, 4(2), 240–251.

Minhas, M. S., & Zelek, J. (2019). Anomaly detection in images. *Computer Vision and Pattern Recognition*, 1–7. https://doi.org/10.48550/arXiv.1905.13147

Mohiuddin, A., Mahmood, A. N., & Hu, J. (2016). A survey of network anomaly detection techniques. *Journal of Network and Computer Applications,* 60, 19–31. https://doi.org/10.1016/j.jnca.2015.11.016

Nassif, A. B., Talib, M. A., Nasir, Q., & Dakalbab, F. M. (2021). Machine learning for anomaly detection: A systematic review. *IEEE Access*, 9, 78658–78700. https://doi.org/10.1109/ACCESS.2021.3083060

Pang, G., Shen, C., Cao, L., & van den Hengel, A. (2021). Deep learning for anomaly detection: A review. *ACM Computing* Surveys, 54(2), 1-38. https://doi.org/10.1145/3439950

Sakthivanitha, M., Priscila, S. S., & Praveen, B. M. (2025). Real-time anomaly detection using deep learning in e-commerce platform. *Procedia Computer Science,* 269, 309–320. https://doi.org/10.1016/j.procs.2025.08.283

Shao, Z., Wang, X., Ji, E., Chen, S., & Wang, J. (2025). GNN-EADD: Graph neural network-based e-commerce anomaly detection via dual-stage learning. *IEEE Access*, 13, 8963–8976. https://doi.org/10.1109/ACCESS.2025.3526239

Sharma, B., Pokharel, P., & Joshi, B. (2020). User behavior analytics for anomaly detection using LSTM autoencoder – Insider threat detection. *In Proceedings of the 11th International Conference on Advances in Information Technology*, 1–10. https://doi.org/10.1145/3406601.3406610)

Sun, M. (2025). Research on e-commerce return prediction and influencing factor analysis based on user behavioral characteristics. *Pinnacle Academic Press Proceedings Series,* 3, 15–28. https://doi.org/10.71222/ghj64w02

Taher, G. (2021). E-commerce: Advantages and limitations. *International Journal of Academic Research in Accounting, Finance and Management Sciences,* 11(1), 153–165. https://doi.org/10.6007/IJARAFMS/v11-i1/8987

Tanır, D., & Ramazanov, S. (2023). Veri madenciliği yöntemleri ile Türkiye'de fertlerin e-ticaret kullanımını etkileyen faktörlerin analizi. *Karamanoğlu Mehmetbey Üniversitesi Sosyal ve Ekonomik Araştırmalar Dergisi,* 25(44), 46–65.

Wang, C., Zhang, Y. M., & Liu, C. L. (2018). Anomaly detection via minimum likelihood generative adversarial networks. *In 2018 24th International Conference on Pattern Recognition (ICPR),*1121–1126. https://doi.org/10.1109/ICPR.2018.8545381

Wang, W., Li, H., Zhang, T., Wang, Z., & Lai, Y. (2023). Two-layer generalization boosting model for anomaly detection of e-commerce orders. *In Sixth International Conference on Computer Information Science and Application Technology (CISAT 2023).* https://doi.org/10.1117/12.3004538

Wu, J., Qiu, Z., Zeng, Z., Xiao, R., Rida, I., & Zhang, S. (2024). Graph autoencoder anomaly detection for e-commerce application by contextual integrating contrast with reconstruction and complementarity. *IEEE Transactions on Consumer Electronics*, 70(1), 1623–1630. https://doi.org/10.1109/TCE.2024.3352186

Zhang, L., & Liu, L. (2022). Data Anomaly Detection Based on Isolation Forest Algorithm. *2022 International Conference on Computation, Big-Data and Engineering (ICCBE),* Yunlin, Taiwan, 87-89. https://doi.org/10.1109/ICCBE56101.2022.9888169

Zhao, P., Ding, Z., Wang, M., & Cao, R. (2019). Behavior analysis for electronic commerce trading systems: A survey. *IEEE Access*, 7, 108703–108728. https://doi.org/10.1109/ACCESS.2019.2933247

Etik, Beyan ve Açıklamalar

1. Etik Kurul izni ile ilgili;

☒ Bu çalışmanın yazar/yazarları, Etik Kurul İznine gerek olmadığını beyan etmektedir.

☐ Bu çalışmanın yazar/yazarları, ……. Üniversitesi ……Etik Kurulu'nun tarih …….., sayı …….. ve karar …… ile etik kurul izin belgesi almış olduklarını beyan etmektedir.

2. Bu çalışmanın yazar/yazarları, araştırma ve yayın etiği ilkelerine uyduklarını kabul etmektedir.

3. Bu çalışmanın yazar/yazarları kullanmış oldukları resim, şekil, fotoğraf ve benzeri belgelerin kullanımında tüm sorumlulukları kabul etmektedir.

4. Bu çalışmanın benzerlik raporu bulunmaktadır.