

Finansal Ve Medyatik Sahteciliğe Karşı Yapay Zekâ Çözümleri: Banknot Sahteciliği Tespitinde Makine Öğrenmesi Uygulamaları

Artificial Intelligence Solutions Against Financial And Media Counterfeiting: Machine Learning Applications In Banknote Counterfeit Detection

Osman Vedüd EŞİDİR¹, Gökhan BAK², Bedriye KARADAĞ³

Öz

Amaç: Yapılan çalışmada, banknot sahteciliğinin tespiti amacıyla yapay zekâ temelli makine öğrenmesi modelleri (Lojistik Regresyon, Rastgele Orman ve XGBoost) kullanılmış ve modeller karşılaştırılarak başarıları analiz edilmiştir. UCI veri seti üzerinden gerçekleştirilen analizler, modellerin yüksek doğruluk oranlarına ulaştığını ve sahte banknotların tespitinde etkili olduklarını ortaya koymuştur.

Tasarım/Yöntem: Çalışmada üç farklı makine öğrenimi modeli kullanılmıştır. Yapılan analizlerde, çeşitli veri ön işleme teknikleri, model eğitim süreçleri, özellik çıkarımı yöntemleri kullanılmıştır

Bulgular: Sonuçlar, makine öğrenmesi tabanlı yaklaşımların banknot sahteciliği tespitinde geleneksel yöntem ve modellere göre daha etkin çözümler sunduğunu ve özellikle finansal sistemin güvenliği açısından uygulamaya yönelik güçlü potansiyel taşıdığını göstermektedir. Çalışma, sadece finansal güvenliği değil; aynı zamanda iletişim ve medya alanlarında dezenformasyonun önlenmesi, basın etiğinin korunması ve görsel sahteciliğin medya ortamlarına yansımalarının anlaşılması açısından da katkı sunmaktadır. Bu yönüyle makale, yapay zekânın sadece teknik değil, aynı zamanda toplumsal ve iletişimsel boyutlarını da içeren disiplinlerarası bir çerçeve önermektedir.

Sınırlılıklar: Yapay zekâ temelli makine öğrenmesi modelleri (Lojistik Regresyon, Rastgele Orman ve XGBoost) kullanılmıştır. Çalışma, UCI'den alınmış sabit bir veri setiyle sınırlıdır.

Özgünlük/Değer: Bu çalışma, farklı makine öğrenmesi tekniklerinin sahtecilik tespitindeki başarısını ortaya koyarak, banknot doğrulama sistemlerine yönelik kapsamlı bir karşılaştırma sunmaktadır

Anahtar Kelimeler: Banknot Sahteciliği, Makine Öğrenmesi, İletişim, Medya, Görsel İçerik Analizi.

Abstract

Purpose: This study employs artificial-intelligence-driven machine-learning models—Logistic Regression, Random Forest, and XGBoost—to detect banknote forgery and comparatively analyze their performance. Analyses conducted on a UCI benchmark dataset show that the models achieve high accuracy rates and prove effective in identifying counterfeit banknotes.

Design/Methodology: Three distinct machine-learning models were implemented. The analysis incorporated various data-preprocessing techniques, feature-extraction methods, and systematic model-training procedures.

Findings: The results indicate that machine-learning-based approaches provide more effective solutions for banknote forgery detection than traditional methods and models. Beyond enhancing financial security, the study also contributes to preventing disinformation in communication and media, safeguarding press ethics, and understanding how visual forgery propagates in media environments. Accordingly, the article offers an interdisciplinary framework that encompasses not only the technical but also the societal and communicative dimensions of artificial intelligence.

Limitations: The investigation is confined to three machine-learning models—Logistic Regression, Random Forest, and XGBoost—and relies on a fixed dataset obtained from the UCI repository.

Originality/Value: By revealing the effectiveness of different machine-learning techniques in forgery detection, the study provides a comprehensive comparison for banknote authentication systems, highlighting their practical potential for financial security and related interdisciplinary domains.

Keywords: Banknote Counterfeiting, Machine Learning, Communication, Media, Visual Content Analysis.

¹Doç. Dr., Fırat Üniversitesi, İletişim Fakültesi, Gazetecilik Bölümü, Elazığ/TÜRKİYE, ovesidir@firat.edu.tr, ORCID: 0000-0003-2029-4758

²Dr., Bağımsız Araştırmacı, Adana/TÜRKİYE, gokhanbak2010@gmail.com, ORCID: 0000-0003-4520-0930

³Milli Eğitim Bakanlığı, Okul Yöneticisi, Adana/TÜRKİYE, bedriyekaradag001@outlook.com, ORCID: 0009-0003-9691-4367

1. GİRİř

Banknot sahtecilięi, finansal güvenlięi tehdit eden kritik bir sorundur. Geleneksel tespit yöntemleri, sınırlı veri işleme kapasitesine sahipken, makine öğrenmesi yöntemleri veri tabanlı karar destek sistemlerinde önemli ilerlemeler sağlamıştır (Yildiz vd., 2020). Bu bağlamda çalışmanın temel amacı, çeşitli makine öğrenmesi modellerinin banknot sahtecilięi tespitindeki etkililięini değerlendirmektir.

Banknotların yüzeylelerinde ve dokularında bulunan çıplak gözle görülmesi pek mümkün olmayan çok küçük detaylar (güvenlik ipleri, mikroyazılar, kabartmalar vs.), çok kısa aralıklarla farklı frekansta varyasyonlar oluşturmaktadır. Wavelet dönüşümü bu küçük, yerel farklılıkları yakalayarak sahte ve gerçek banknotlar arasındaki ayrımı istatistiksel olarak görünür kılar. Özellikle varyans, çarpıklık ve eğiklik gibi wavelet dönüşüm sonrası hesaplanan istatistikler bu tür sınıflandırma işlemlerinde güçlü ayırt edici özniteliklerdir.

Günümüzde dijital medya ortamlarında dolaşıma giren görsel ve metinsel içeriklerin doğruluęu, gazetecilik mesleęi açısından büyük bir önem arz etmektedir. Özellikle sahte banknotların yalnızca finansal sistem deęil, medya temsilleri aracılıęıyla kamuoyu algısı üzerinde de etkili olduęu görülmektedir. Bu bağlamda, yapay zekâ destekli makine öğrenmesi modelleri, sadece ekonomik deęil; aynı zamanda medya ve basın etięi çerçevesinde sahtecilięin tespitine katkı sunan stratejik araçlara dönüşmektedir.

Makine öğrenmesi, bilgisayar sistemlerinin verilerden öğrenme yeteneęini ifade eden bir yapay zekâ alt alanıdır ve özellikle son yıllarda büyük veri kümeleri üzerinde uygulamalı arařtırmalarda yoğun olarak kullanılmaktadır. Bu yöntemler sayesinde büyük ve karmařık veri setlerinden anlamlı bilgiler elde edilebilmekte, aynı zamanda geleceęe yönelik tahminler yapılarak karar alma süreçleri desteklenmektedir. Özellikle tahmine dayalı analitik (predictive analytics), veri yoğun sektörlerde stratejik planlama ve müşteri davranıřlarını anlamlandırmak için temel bir araç haline gelmiştir. Makine öğrenmesi, son yıllarda büyük veri analitięi ve tahmine dayalı modelleme süreçlerinde önemli bir araç haline gelmiştir. Özellikle tüketici davranıřlarının anlaşılması, öngörülerin geliştirilmesi ve stratejik karar destek sistemlerinin oluřturulmasında makine öğrenmesi model ve algoritmaları etkin ve bařarılı řekilde kullanılmaktadır (Eřidir, 2025a).

2. LİTERATÜR TARAMASI VE KAVRAMSAL ÇERÇEVE

Kapsamlı veri setlerinde gizli kalmıř çok katmanlı iliřkilerin ve desenlerin ortaya çıkarılmasında klasik istatistik yöntemlerinin kısıtlı kaldıęı literatürde vurgulanmaktadır (Speer, 2021). Bu sınırlılıęı ařmak üzere devreye giren makine öğrenimi modelleri, deęiřken evrenini geniş tutarak ve parametreler arası çoklu etkileřimleri hesaba katarak yüksek doğruluk düzeyine sahip kestirimler üretmektedir. Wu'nun (2023) bulguları da, söz konusu algoritmaların deęiřkenler arasında kurulan karmařık aęları bütüncül biçimde analiz ederek geleneksel yaklařımlara oranla daha kapsayıcı ve duyarlı sonuçlar verdięini ortaya koymaktadır.

Yapay zekâ temelli sahtecilik tespitine iliřkin literatür, büyük oranda finans ve güvenlik alanlarında yoğunlařmıştır. Ancak son yıllarda, iletiřim ve medya arařtırmalarında da yapay zekâ modellerinin, dezenformasyonun önlenmesi, içerik doğrulama ve görsel manipölasyonların tespiti gibi konularda kullanımı artmıştır (Örnek: deepfake içeriklerin teřhisi). Bu geliřmeler, banknot gibi görsel nitelikli sahtecilięin medya teknolojileriyle olan iliřkisinin daha derinlikli biçimde ele alınmasını zorunlu kılmaktadır.

Karabınar ve Can (2005) çalışmalarında, iřletmelerin para trafięinde sahte para ile karřılařmaları durumunda ortaya çıkan muhasebe ve vergi sorunlarını incelemiřlerdir. Çalışmada, sahte paranın muhasebe kayıtlarına alınmasında farklı muhasebe çözümleri önerilmiştir. Ayrıca, sahte paranın kanunen kabul edilmeyen gider sayılması nedeniyle, iřletmelerin gerçekte elde etmedikleri bir gelir üzerinden vergi ödeme yükümlölüęü ile karřılařtıkları vurgulanmıştır. Yazarlar, mevcut vergi mevzuatında sahte paraya iliřkin açık bir düzenlemenin bulunmadıęını, bu durumun iřletmelerin mağduriyetine neden olduęunu belirtmiřlerdir. Çalışmada ayrıca, sahte paranın emniyet özellikleri ve sahte parayı ayırt etme yöntemleri de ayrıntılı řekilde açıklanmıştır.

Yildiz ve arkadaşları (2020), Bosna-Hersek para birimi olan Konvertibl Mark banknotlarının sahtecilik tespiti için derin öğrenme yöntemlerini uygulamışlardır. Çalışmada 189 gerçek ve 173 sahte banknottan oluşan bir ham veri seti oluşturulmuş ve veri artırımı (data augmentation) ile veri seti genişletilerek 3360 gerçek ve 2412 sahte banknot elde edilmiştir. İnce ayar (fine-tuning) yöntemiyle önceden eğitilmiş CNN modelleri ile sahte banknotlar tespit edilmiştir. Elde edilen sonuçlara göre, VGG16 modeli %99,88 doğruluk oranı ile en yüksek performansı göstermiştir. AlexNet %99,38 ve GoogLeNet %97,36 doğruluk elde etmiştir. Araştırma, veri artırımı uygulamalarının model başarımını artırmadaki etkisini vurgulamakta ve derin öğrenme tabanlı modellerin geleneksel sahte para tespit tekniklerine kıyasla daha yüksek doğruluk ve verimlilik sunduğunu göstermektedir.

Pachón, Ballesteros ve Renza (2021) çalışmalarında, sahte banknotların tespiti amacıyla derin öğrenme tabanlı yöntemlerin performansını analiz etmişlerdir. Çalışmada, Kolombiya Peso'suna ait gerçek ve sahte banknotlardan oluşan bir veri seti kullanılmıştır. Transfer öğrenme (transfer learning) ve CNN mimarisi ile analizler yapılmıştır. Transfer öğrenme yaklaşımında AlexNet, SqueezeNet, ResNet18 ve InceptionV3 mimarileri değerlendirilmiştir. Sonuçlar, ResNet18 mimarisinin en yüksek doğruluk (%100) oranına ulaştığını; ancak önerilen özel modelin CPU ve GPU üzerinde önemli ölçüde daha hızlı çıkarım (inference) süreleri sağladığını göstermiştir. Çalışma, sahte banknot tanıma literatürüne yöntemsel çeşitlilik ve de düşük maliyetli çözümler üretme açısından katkı sunmaktadır.

Çilburunoğlu'nun (2023) çalışması, kredi kartı sahtekarlığını saptamak amacıyla makine öğrenimine dayalı çoklu modelleri karşılaştırmalı olarak incelemektedir. Araştırmada Random Forest, Yapay Sinir Ağları (YSA), Destek Vektör Makineleri (DVM), Lojistik Regresyon ile Karar Ağacı algoritmaları uygulanmış ve bunların performansları değerlendirilmiştir. Bulgular, her algoritmanın kendine özgü öğrenme mekanizmaları nedeniyle çıktılarında farklılık gösterebildiğini, ancak bütün modellerin hesaplanan doğruluk değerleri ile diğer performans göstergelerinin genel olarak yüksek düzeylerde seyrettiğini ortaya koymaktadır.

Altan ve Zafer (2024) çalışmalarında, kredi kartı sahteciliğinin tespitinde denetimli makine öğrenmesi modellerini analiz etmişlerdir. Çalışmada, Ocak 2023 dönemine ait bir kamu bankasının 13.050 adet kredi kartı işleminin gerçek verileri kullanılmıştır. Python programlama dili ile Random Forest, Logistic Regression, K-En Yakın Komşu, Karar Ağaçları ve Gradient Boosting modelleri ile analizler yapılmıştır. Modellerin performansları doğruluk, precision, recall, F1 skoru ve ROC-AUC metrikleriyle değerlendirilmiştir. En yüksek doğruluk oranı %93,1 ile Karar Ağaçları modelinde elde edilmiştir. Gerçek ve güncel veriler ile gerçekleştirilen çalışmada, bankaların kredi kartı sahteciliği tespitinde operasyonel ve maliyet yükünü azaltabilecek etkili makine öğrenmesi çözümleri sunulmuştur. Gerçek veriler kullanılarak modeller geliştirilmiştir.

Woldehana ve arkadaşları (2025) çalışmalarında, sahte ve gerçek Etiyopya banknotlarının sınıflandırılması için açıklanabilir yapay zeka destekli derin öğrenme tabanlı bir yöntem önermişlerdir. Çalışmada, 4000 adet banknot görüntüsünden oluşan bir veri seti oluşturulmuştur. Eğitim ve test için veri seti %80 ve %20 oranında bölünmüştür. DenseNet121 modeli %99,87 doğruluk oranı ile en yüksek başarıyı göstermiştir. Çalışmada SHAP (Shapley Additive Explanations) ve TF-EXPLAIN açıklanabilir yapay zeka teknikleri kullanılarak modelin karar mekanizmaları analiz edilmiştir. Banknotlardaki su işaretleri ve merkez bölgelerin sınıflandırmada belirleyici olduğu tespit edilmiştir. Geliştirilen mobil uygulama ile kullanıcılar çevrimdışı ortamda banknot doğrulaması yapabilmektedir.

3. METODOLOJİ LİTERATÜR TARAMASI

Araştırmada üç ayrı makine öğrenmesi tekniği ele alınmıştır. Analitik süreçte kapsamlı veri ön işleme prosedürleri, model eğitimi stratejileri ve özellik mühendisliği yaklaşımları uygulanmıştır. Büyük veri ekosisteminin çeşitliliği, üretim hızı ve hacmi, metodolojik bakımdan önemli güçlükler doğurmaktadır (Eşidir, 2025a). Makine öğrenmesi, veri temelli problemlere uygun algoritmalar aracılığıyla modeller kurmayı hedefleyen hesaplamalı bir disiplindir. Bu alan, veriden öğrenerek elde ettiği bilgiyi öngöründe bulunmak ve karar süreçlerini desteklemek üzere kullanan yapay zekâ alt dalı olarak konumlanmaktadır. Hangi makine öğrenmesi algoritmasının tercih edileceği, verinin nitelikleri ve ele alınan sorunsalın karmaşıklık düzeyi tarafından belirlenmektedir.

Bu çalışmada, banknotların sahte mi gerçek mi olduğunun belirlenmesi amacıyla UCI Machine Learning Repository üzerinden temin edilen Banknote Authentication veri seti kullanılmıştır.

Veri seti, wavelet dönüşümü ile çıkarılan dört adet öznitelik (varyans, arpıklık, eğiklik ve entropi) ile her bir banknotun gerçek veya sahte olduğunu gösteren ikili bir sınıf etiketinden oluşmaktadır. Öncelikle, veri seti üzerinde temel betimleyici istatistikler hesaplanmıştır. Wavelet dönüşümü (Türkede "Dalgaık Dönüşümü" olarak da adlandırılır), bir sinyalin veya verinin farklı zaman ve frekans öleklerinde analiz edilmesini saėlayan matematiksel bir tekniktir.

3.1. Yazılım ve Donanım Altyapısı

Veri analizi ve modelleme süreçleri Python 3.12.7 programlama dili kullanılarak gerçekleştirilmiştir. alıřma süresince NumPy, Pandas, Scikit-learn, TensorFlow ve Keras gibi yaygın veri bilimi kütüphaneleri kullanılmıştır. Kodlama ve görselleřtirme işlemleri, Jupyter Notebook 7.2.2 platformu üzerinde yürütölmüřtür. Jupyter Notebook; kod yazma, veri görselleřtirme ve analiz süreçlerinin entegre bir ortamda sunulmasını saėlamaktadır. Kullanılan sistemde Windows 11 Pro işletim sistemi ve 16 GB RAM'e sahip Intel Core i7 işlemcili bir donanım tercih edilmiştir. Bilgisayar sistemi, veri setinin hızlı ve verimli biçimde işlenebilmesine olanak tanımıştır.

3.2. Veri Seti ve Ön İşlemler

Veri setinin yüklenmesi ve ön işlenmesi, analizin ilk adımıdır. alıřmada, banknotların sahte mi gerçek mi olduğunu belirlenmesi amacıyla UCI Machine Learning Repository üzerinden temin edilen Banknote Authentication veri seti kullanılmıştır. Kullanılan veri seti, Varyans, Simetri, Eğrilik ve Entropi özelliklerinden oluşmaktadır. Öznitelikler StandardScaler ile öleklendirilerek model eğitimine hazır hale getirilmiştir. Veri dağılımı analiz edilmiş ve sınıflar arasında anlamlı ayrımlar olduğu gözlemlenmiştir. Veri setinde toplamda 761 adet Gerçek (0) ve 610 adet Sahte (1) banknot bulunmaktadır. Dolayısıyla veri setinin dengeli dağılım gösterdiği söylenebilir. Eğitim sürecinde herhangi bir sınıf dengesi düzeltilme yöntemine ihtiyaç duyulmamıştır. Veri setinde 0 gerçek ve 1 sahte olmak üzere iki sınıf bulunmaktadır.

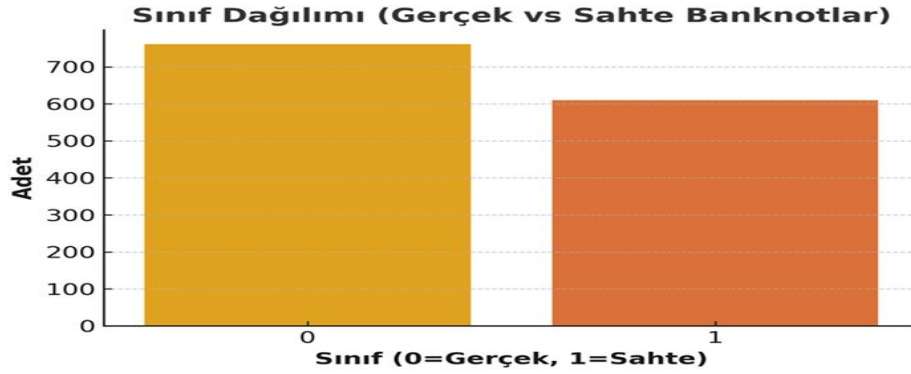
Tablo 1: Veri Setinden Alınan Örnek Kayıtlar

Varyans (Variance)	arpıklık (Skewness)	Basıklık (Curtosis)	Entropi (Entropy)	Sınıf (Class)
-0,24745	1,9368	-2,4697	-0,80518	1
4,4549	2,4976	1,0313	0,96894	0
-0,52645	-0,24832	-0,45613	0,41938	1
-0,39816	5,9781	1,3912	-1,1621	0
-2,588	3,8654	-0,3336	-1,2797	1
3,82	10,9279	-4,0112	-5,0284	0
0,11806	0,39108	-0,98223	0,42843	1
3,9922	-4,4676	3,7304	-0,1095	0
2,0051	-6,8638	8,132	-0,2401	0
0,63655	5,2022	-5,2159	-6,1211	1

Tablo 1'de, veri setinden seçilmiş 5 gerçek ve de 5 sahte banknota ait örnek kayıtlar gösterilmiştir. Her gözlemin varyans, arpıklık, basıklık ve entropi deėerleri sunulmuřtur.

3. ARAřTIRMANIN YÖNTEMİ

Bu alıřma, medya ekonomisi ve işletmeciliėi alanındaki akademik üretimin dinamik yapısını, kuramsal temellerini ve yöntemsel eğilimlerini bibliyometrik analiz yöntemiyle sistematik olarak deėerlendirmeyi amaçlamaktadır. Arařtırma, uluslararası akademik yayınları kapsayan geniş bir evren üzerinde detaylı bir literatür taraması gerçekleştirilmesi ve elde edilen verilerin VOSviewer gibi metin madenciliėi uygulamalarıyla görselleřtirilmesine dayanmaktadır.

Şekil 1: Veri Setindeki Gerçek ve Sahte Banknotların Dağılımı

Şekil 1’de görüldüğü üzere veri setindeki gerçek ve sahte banknotların sayıları birbirlerine yakındır dolayısı ile veri setinin dengeli bir şekilde dağıldığını söylemek mümkündür. Modelleme sürecinde sınıf dengesizliği kaynaklı sapmalar daha az olacaktır. Veri setinde toplamda 761 adet Gerçek (0) ve 610 adet Sahte (1) banknot bulunmaktadır.

3.3. Veri Ön İşleme

Veri ön işleme, bir veri kümesini analiz ile model kurulumuna elverişli hâle getirme safhasıdır. Bu kademede yürütülen adımlar, geliştirilecek modellerin hem doğruluğunu hem de genel performansını doğrudan şekillendirir (Eşidir, 2025a). Çalışmada, ön işleme sırasında ilk olarak kategorik nitelikler sayısal göstergelere dönüştürülmüş, ardından veri kümesi %80 eğitim ve %20 sınama alt kümesine ayrılmıştır. Sayısal öznitelikler Standart Ölçekleyici yöntemiyle yeniden ölçeklendirilerek homojen bir dağılım elde edilmiştir. Böyle bir ölçekleme uygulaması, öğrenme sürecini hızlandırmakta ve değişkenler arasındaki karşılaştırılabilirliği güçlendirmektedir; zira birbirinden farklı ölçeklerdeki özniteliklerin doğrudan modele sokulması, katsayı öğreniminde dengesizliklere yol açabilir. Ölçeği eşitleyen bu müdahale, algoritmanın örüntüleri daha verimli biçimde irdelemesine imkân tanımaktadır.

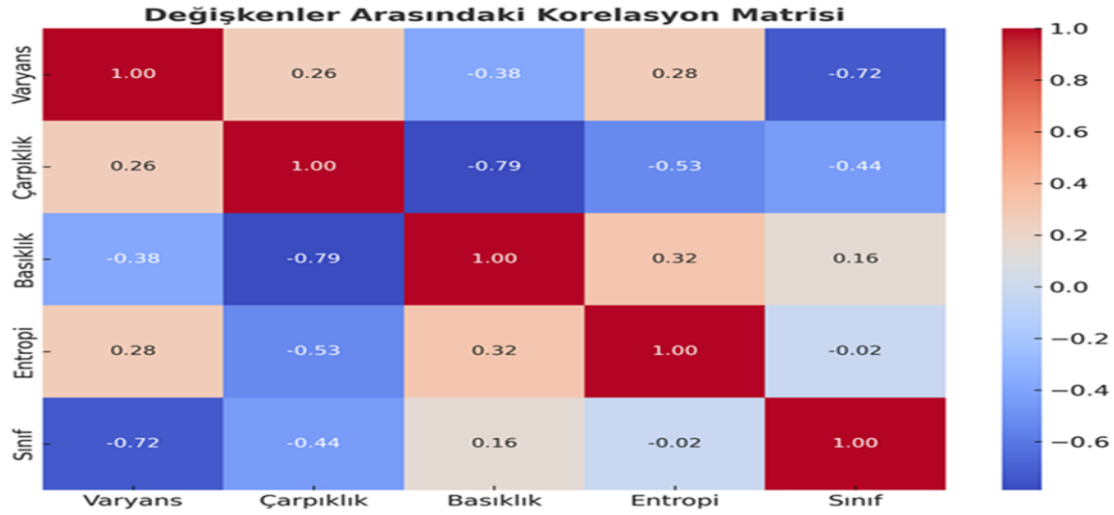
3.4. Modellerin Performans Değerlendirme Kriterleri

Makine öğrenmesi modellerinin performanslarını nesnel biçimde değerlendirebilmek amacıyla çeşitli metrikler ve analiz yöntemleri kullanılmaktadır. Bu çalışmada model performansları; Kesinlik (Precision), Duyarlılık (Recall), F1 Skoru ve Doğruluk (Accuracy) olmak üzere dört temel metrik üzerinden analiz edilmiştir. Söz konusu metrikler, farklı senaryolarda modelin güçlü ve zayıf yönlerinin belirlenmesi açısından kritik önem taşımaktadır. Kesinlik, modelin pozitif sınıf tahminlerinde ne ölçüde isabetli olduğunu; Duyarlılık ise gerçek pozitif örnekleri ne ölçüde doğru tanıdığını göstermektedir (Eşidir, 2025b). F1 Skoru, kesinlik ve duyarlılık arasında denge kurarak sınıflandırma başarısını tek bir ölçü altında özetlerken; Doğruluk, modelin genel başarı düzeyini ifade etmektedir. Bu metrikler, özellikle dengesiz veri setlerinde modelin sınıflandırma performansını bütüncül bir şekilde değerlendirmek için kullanılmaktadır.

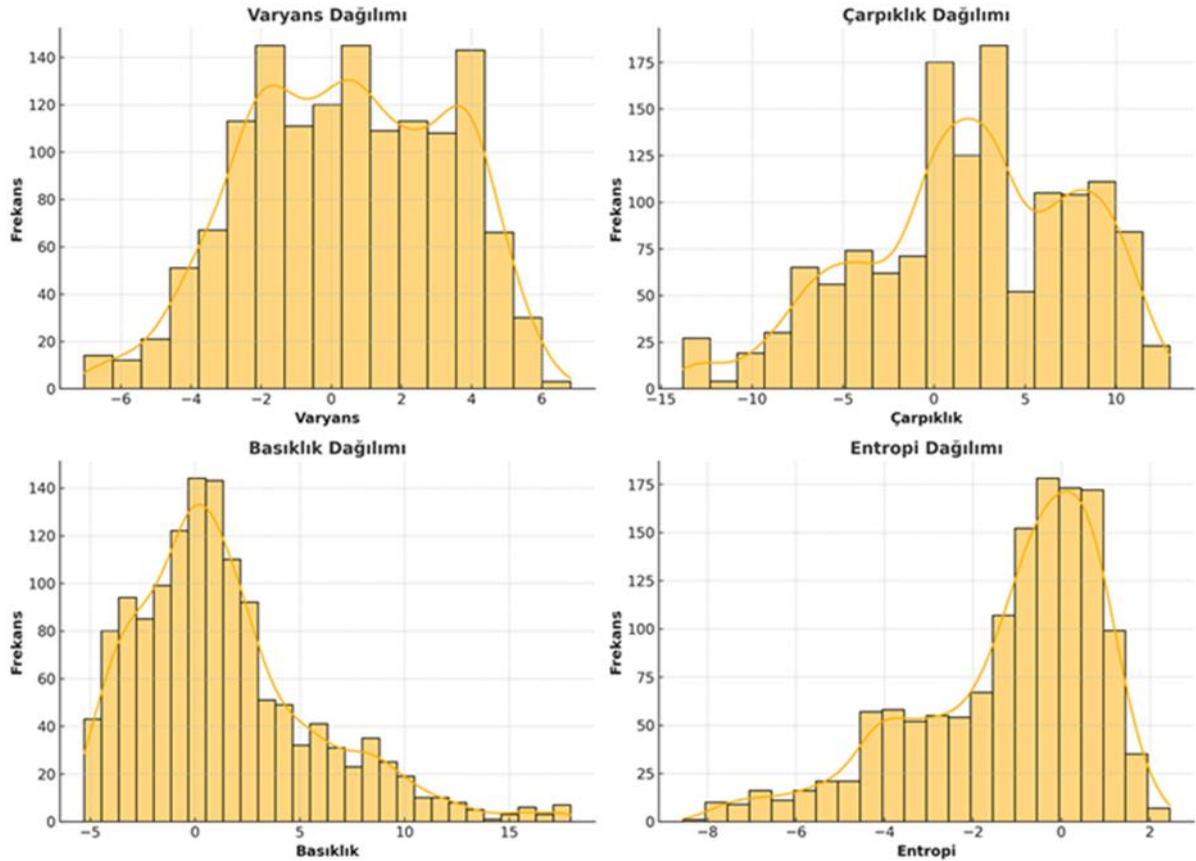
4. ARAŞTIRMANIN ANALİZİ VE BULGULARI

4.1. Görsel İstatistiksel Analiz ve Sonuçlar

Bu bölümde, görsel betimleyici analizler sunularak, temel istatistiksel eğilimlerin anlaşılmasına katkı sağlanmıştır. Amaç, veri setinde yer alan özniteliklerin genel dağılımını ve birbirleriyle olan ilişkilerini görselleştirerek, makine öğrenmesi modelleri öncesi veri yapısına dair sezgisel çıkarımlarda bulunmaktır.

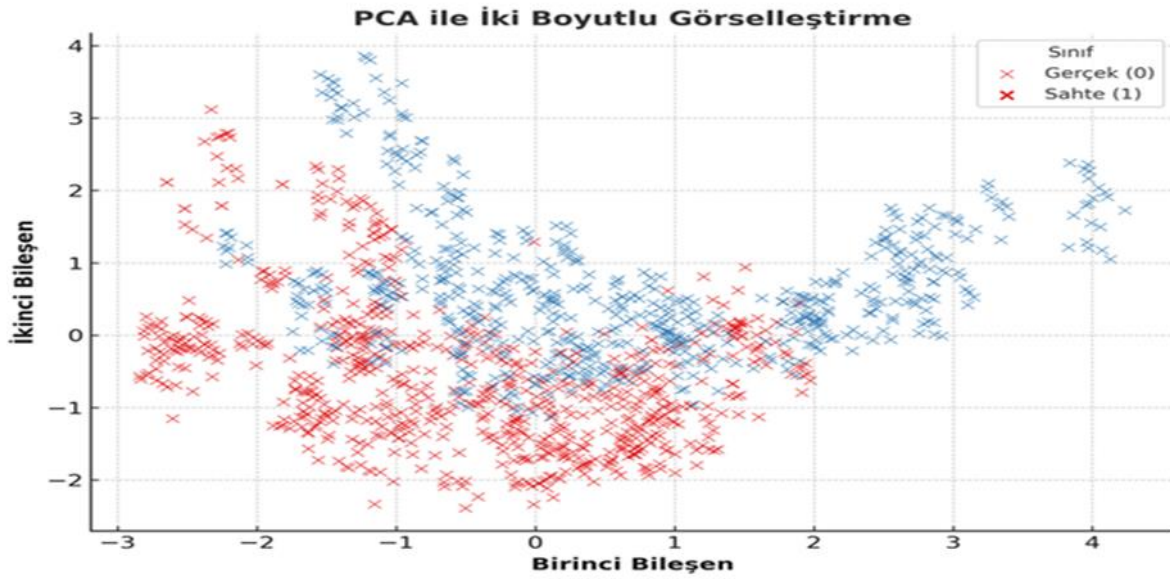
Şekil 2: Değişkenler Arasındaki Korelasyon Matrisi

Şekil 2'deki Korelasyon matrisi, değişkenler arasındaki doğrusal ilişki düzeylerini göstermektedir. Korelasyon matrisi incelendiğinde, giriş değişkenleri arasında güçlü doğrusal ilişkiler bulunmamaktadır; her bir değişken sahteciliği farklı bir yönüyle yakalamaya yönelik bağımsız bilgiler sunmaktadır.

Şekil 3: Veri Setindeki Değişkenlerin Dağılımı

Şekil 3'te veri setindeki değişkenlerin dağılımı gösterilmiştir. Değişkenlerin dağılımı incelendiğinde, bazı değişkenlerde asimetri ve uç değer etkileri gözlemlenmiştir. Bu durum sınıflandırma modellerinin başarısını etkileyebilir.

Şekil 4: PCA İle İki Boyutlu Görselleştirme



Şekil 4’te PCA sonuçlarına göre, banknotların sınıf ayrımı makul düzeyde temsil edilmektedir. Özellikle sahte banknotlar belirli bir kümelenme eğilimi göstermektedir. Toplam varyansın büyük bölümünün yakalandığı ve sahte banknotların bu düzlemde kümелendiği görülmektedir. Bu da, orijinal dört boyutlu özellik uzayında sınıfların ayrılabilir olduğuna işaret etmektedir.

5. MAKİNE ÖĞRENMESİ MODELLERİ

Makine öğrenmesi, bilgisayar sistemlerinin veri analizi yoluyla öğrenmesini ve bu süreçte elde edilen bilgiler doğrultusunda tahmin yapabilme yetisi kazanmasını amaçlayan bir bilim dalıdır. Çözüm süreçlerinde istatistiksel ve matematiksel modeller kullanılarak veri kümesinin yapısına ve problem türüne uygun algoritmalar seçilmektedir (Eşidir, 2025c).

5.1. Lojistik Regresyon

Lojistik regresyon, bağımlı değişkenin kategorik olduğu durumlarda başvurulacak istatistik temelli bir sınıflandırma yöntemidir. İki sınıflı problemlerde yaygın biçimde kullanılan bu yaklaşım, gözlemlerin belirli bir kategoriye ait olma olasılığını hesaplar. Yöntem, doğrusal regresyonla elde edilen birleşik skoru lojistik (sigmoid) dönüşümden geçirerek çıktıları $[0,1]$ aralığında sınırlandırır ve böylelikle ikili sınıflandırmayı mümkün kılar (Aslan, 2025).

5.2. Rastgele Orman (Random Forest)

Rastgele Orman (Random Forest), topluluk öğrenmesi (ensemble learning) yaklaşımına dayanan ve sınıflandırma ile regresyon problemlerinde yaygın olarak kullanılan güçlü bir makine öğrenmesi modelidir. Model, farklı alt örneklem kümeleri üzerinden oluşturulan çok sayıda karar ağacından oluşmakta ve nihai karar, bu ağaçların oylama yöntemiyle alınmaktadır. Bu yapı sayesinde, tek bir karar ağacında oluşabilecek aşırı öğrenme (overfitting) problemi azaltılarak modelin genelleme yetisi artırılmaktadır.

Sınıflandırma ve regresyon problemlerinde yaygın olarak kullanılmakta olup, birden fazla karar ağacını bir araya getirerek her bir ağacın bireysel zayıflıklarını dengelemekte ve daha kararlı, güçlü ve genelleştirilebilir bir model oluşturmaktadır. Bu algoritma, bootstrap örnekleme yöntemiyle eğitim veri setinden birçok alt örneklem (subsample) oluşturur. Her bir bootstrap örneği, orijinal veri kümesinden rastgele seçilen örnekleri içermekte olup, bu örnekler kullanılarak bağımsız karar ağaçları inşa edilir. Bu süreçte, her karar ağacı veri setinin farklı bir alt kümesi üzerinde eğitilir. Ayrıca, her düğümdeki bölünme işlemi tüm özellikler yerine rastgele seçilen bir alt küme üzerinden gerçekleştirilir. Bu yöntem, modelin varyansını düşürerek aşırı öğrenmeyi (overfitting) önlemekte ve farklı karar ağaçları arasındaki korelasyonu azaltarak modelin genelleme performansını artırmaktadır.

Sınıflandırma problemlerinde, nihai tahmin genellikle oylama (majority voting) yöntemiyle belirlenmektedir (Oguine vd., 2021).

5.3. XGBoost (Extreme Gradient Boosting)

XGBoost, gradyan artırma (gradient boosting) algoritmasının optimize edilmiş ve geniş veri setlerinde yüksek doğruluk sağlayan bir versiyonudur. Model, art arda oluşturulan zayıf öğreniciler (genellikle karar ağaçları) ile önceki tahmin hatalarını azaltmayı hedefler. Her iterasyonda model, bir önceki öğrenicinin hatalarına odaklanarak öğrenme sürecini iyileştirir ve bu sayede güçlü bir sınıflandırıcı elde edilir.

XGBoost, hem sınıflandırma hem de regresyon görevlerinde yaygın olarak tercih edilen, son derece hızlı çalışıp yüksek doğruluk düzeyleri sunması nedeniyle veri bilimi yarışmaları ile büyük ölçekli uygulamalarda öne çıkan bir yöntemdir. Son yıllarda makine öğrenmesi literatüründe en başarılı ve en sık başvurulanan modeller arasında yer alan bu algoritma; spam e-posta saptama, reklam eşleştirme, sahtekârlık belirleme ve anomali izleme gibi pek çok problem alanında etkili çıktılar üretmektedir (Chen ve Guestrin, 2016). XGBoost’un sağladığı hassas kestirimler ve alternatif makine öğrenmesi yaklaşımlarına kıyasla sunduğu hesaplama avantajları, özellikle hacimli veri kümeleri ve karmaşık problem yapıları söz konusu olduğunda modelin popülaritesini artırmıştır (Abar, 2020).

XGBoost, artan gradyan yöntemini temel alan bir topluluk öğrenmesi modelidir; bu yaklaşım, seçilen kayıp fonksiyonunu minimize etmek amacıyla zayıf öğrenicileri ardışık biçimde kümüle ederek öngörü isabetini yükseltir (Chen ve Guestrin, 2016). Yüksek ölçeklenebilirlik ve parametrik esneklik özellikleri, yöntemin gradyan artırma kütüphaneleri içinde hızla öne çıkmasını sağlamış ve onu yaygın biçimde kullanılan bir topluluk stratejisi hâline getirmiştir (Aslan, 2025). Torbalama ile istifleme gibi yaygın topluluk yaklaşımlarının yanında, boosting kategorisinde konumlanan bu model, ardışık yapı sayesinde tahmin hatalarını sistematik biçimde azaltır (Wang vd., 2020). “Extreme Gradient Boosting” ifadesinin kısaltması olan XGBoost, çok sayıda düşük kapasiteli öğreniciyi aşamalı olarak birleştirip yüksek performanslı tahminleyicilere dönüştürmeyi hedefler (Liang vd., 2020).

6. MAKİNE ÖĞRENMESİ MODELLERİNİN PERFORMANSLARI

Makine öğrenmesi yaklaşımları, klasik modellere göre daha geniş bir değişken çeşitliliğini kapsamlı biçimde inceleyebilmekte ve çok boyutlu ilişkileri başarıyla modelleyebilme olanağı sunmaktadır. Üç farklı makine öğrenmesi modeli kullanılarak elde edilen tahmin performansları karşılaştırmalı olarak analiz edilmiştir. Makine öğrenmesi modellerinin banknot doğrulama verisindeki performanslarına ilişkin bulgular detaylı biçimde değerlendirilmektedir. Elde edilen metrikler, modellerin sınıflandırma başarılarını karşılaştırmalı olarak analiz etmeye olanak sağlamıştır. Modellerin performans metrikleri Tablo 2’de görülmektedir.

Tablo 2: Makine Öğrenmesi Modellerinin Performans Metrikleri

Model	Doğruluk	F1 Skoru	Hassasiyet	Duyarlılık
Lojistik Regresyon	0,993	0,992	0,984	1
Rastgele Orman	0,996	0,996	0,992	1
XGBoost	0,996	0,996	0,992	1

Tablo 2’de sunulan sonuçlar, banknot doğrulama probleminde kullanılan üç farklı makine öğrenmesi modelinin oldukça yüksek başarı düzeylerine ulaştığını göstermektedir. Lojistik Regresyon modeli %99,3 doğruluk ve %99,2 F1 skoru ile güçlü bir performans sergilerken, Rastgele Orman ve XGBoost modelleri %99,6 doğruluk ve %99,6 F1 skoru ile bu başarıyı bir adım öteye taşımıştır. Tüm modellerin duyarlılık oranlarının 1,00 olması, modellerin tüm gerçek örnekleri doğru bir şekilde sınıflandırdığını ve hiçbir gerçek pozitif örneği kaçırmadığını ortaya koymaktadır. Hassasiyet değerleri ise %98,4 ile %99,2 arasında değişmiş, bu da modellerin yanlış pozitif oranlarının oldukça düşük olduğunu göstermiştir.

Karşılaştırmalı sonuçlar incelendiğinde, Rastgele Orman ve XGBoost modelleri, doğruluk, F1 skoru ve hassasiyet bakımından Lojistik Regresyon modeline kıyasla daha üstün bir performans

sergilemiştir. Bu durum, özellikle hata toleransının düşük olduğu sahtecilik tespiti gibi uygulamalarda bu iki modelin tercih edilmesi gerektiğine işaret etmektedir. Genel olarak tüm modeller, veri setindeki sınıflar arasında ayırım yapmada son derece başarılı olmuş, düşük varyansla yüksek doğruluk sağlamış ve endüstriyel ölçekte güvenle kullanılabilecek bir sınıflandırma başarısı sunmuştur.

7. SONUÇ VE ÖNERİLER

Yapılan çalışma, banknot sahteciliği tespitinde makine öğrenmesi modellerinin etkinliğini ortaya koymuş ve istatistiksel açıdan sağlam performanslar elde edildiğini göstermiştir. Çalışmada, banknotların sahte mi gerçek mi olduğunun belirlenmesine yönelik olarak wavelet dönüşümüyle elde edilen varyans, çarpıklık, eğiklik ve entropi değişkenleri üzerinden istatistiksel analizler gerçekleştirilmiştir. Bulgular, tüm değişkenlerde gerçek ve sahte banknotlar arasında istatistiksel olarak anlamlı farklılıklar bulunduğunu göstermiştir ($p < 0.05$).

Özellikle varyans ve entropi değişkenlerinde gözlenen anlamlı farklılıklar, sahte banknot yüzeylerindeki düzensizliklerin ve dalgalanmaların, gerçek banknotlara kıyasla daha belirgin olduğunu ortaya koymuştur. Bu durum, wavelet tabanlı özellik çıkarımının sahteciliği tespit etmede etkin bir yöntem olduğunu desteklemektedir. Çarpıklık ve eğiklik değişkenlerinde elde edilen bulgular ise, sahte banknotlarda yüzey dokularının daha asimetrik ve uç değerlere sahip olduğunu göstermiştir. Bu bulgular, sahte banknot üretim süreçlerindeki yapısal düzensizliklerin veri temelli analizlerle ortaya konulabileceğini göstermektedir.

Gelecekte yapılacak çalışmalar, banknot sahteciliği gibi fiziksel manipülasyonların medya ortamlarında dijital dezenformasyonla birlikte ele alınmasını sağlayarak daha bütüncül yaklaşımlar geliştirebilir. İletişim ve gazetecilik disiplinlerinin katkısıyla, yapay zekâ modelleri yalnızca sahteciliği tespit etmekle kalmayıp, aynı zamanda bu tür olayların kamuoyuna nasıl yansıtıldığına dair analizler için de kullanılabilir.

Çalışma, UCI'den alınmış sabit bir veri setiyle sınırlıdır; gerçek hayatta karşılaşılabilecek farklı aşınma seviyeleri veya baskı hataları gibi faktörler bu veride temsil edilmemektedir. Gerçek hayat koşullarında aşınmış, yıpranmış veya özel koşullarda basılmış banknotlar gibi daha çeşitli veri türleriyle çalışılması gerekmektedir. Ayrıca veri setinin basit yapısı nedeniyle, analizlerde kullanılan makine öğrenmesi modellerinin ayrıştırma gücü tam olarak sınanamamış olabilir. Elde edilen sonuçlar, bankacılık sektöründe otomatik sahte para tarama sistemlerinin geliştirilmesinde kullanılabilir.

Bu çalışma, farklı makine öğrenmesi tekniklerinin sahtecilik tespitindeki başarısını ortaya koyarak, banknot doğrulama sistemlerine yönelik kapsamlı bir karşılaştırma sunmaktadır. Gelecekte yapılacak benzer çalışmalarda, farklı para birimleri ve daha büyük çaptaki veri setleriyle modeller sınanarak testler gerçekleştirilebilir. Ek olarak, derin öğrenme veya görüntü işleme teknikleri ile sahtecilik tespitinde insan gözünün yakalamasının mümkün olmadığı detaylar tespit edilebilir.

Etik Beyan: Bu çalışmada “Etik Kurul” izni alınmasını gerektiren bir yöntem kullanılmamıştır.

Yazar Katkı Beyanı: 1. Yazarın katkı oranı %34, 2. Yazarın katkı oranı %33, 3. Yazarın katkı oranı ise %33'dür.

Çıkar Beyanı: Yazarlar arasında çıkar çatışması yoktur.

Ethics Statement: In this study, no method requiring the permission of the “Ethics Committee” was used.

Author Contributions Statement: 1st author's contribution rate is 34%, 2nd author's contribution rate is 33%, 3rd author's contribution rate is 33%.

Conflict of Interest: There is no conflict of interest among the authors.

KAYNAKA

- Abar, H. (2020). XGBoost ve Mars yntemleriyle altın fiyatlarının kestirimi. *EKEV Akademi Dergisi*, 83, 427-446.
- Altan, G., & Zafer, M. R. (2024). Denetimli makine ğrenmesi yntemleri ile kredi kartı sahteciliğini tahmin etme: Karşılařtırmaalı analiz. *İktisat Politikası Arařtırmaları Dergisi*, 11(2), 242-262. <https://doi.org/10.26650/JEPR1433315>
- Aslan, K. (2025). *Yapay zekâ makine ğrenmesi ve veri bilimi kursu: Sınıfta yapılan rnekler ve zet notlar*. C ve sistem programcıları derneğı, İstanbul.
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785 – 794. <https://doi.org/10.1145/2939672.2939785>
- ilburunoğlu, K. (2023). Kredi kartı dolandırıcılık tespitinde makine ğrenme algoritmalarının karşılařtırmaalı analizi (Yksek Lisans Tezi). İstanbul Gedik niversitesi Eėitim Enstits.
- Eřidir, K. A. (2025a). Makine ğrenimi modelleri ile yetiřkin eėitimi analizi: Modellerin karşılařtırmaalı performansı. *Elektronik Sosyal Bilimler Dergisi*, 24(2), 946-964. <https://doi.org/10.17755/esosder.1589887>
- Eřidir, K. A. (2025b). TİK mikro verileri ile ocuk iřgc tahmini: Makine ğrenimi modellerinin performans analizleri. *Fırat niversitesi Mhendislik Bilimleri Dergisi*, 37(1), 453-466. <https://doi.org/10.35234/fumbd.1607609>
- Eřidir, K. A. (2025c). Trkiye'nin kimyasal madde ithalatının gelecek tahmini: Makine ğrenmesi ve topluluk ğrenme yntemleri performans analizi. *Fırat University Journal of Social Sciences*, 35(1), 261-278. <https://doi.org/10.18069/firatsbed.1580620>
- Karabınar, S., & Can, A. V. (2005). Sahte para olayının vergi ve muhasebe aısından incelenmesi: Sorunlar ve neriler. *Muhasebe ve Finansman Dergisi*, (25), 113-123.
- Liang, W., Luo, S., Zhao, G., & Wu, H. (2020). Predicting hard rock pillar stability using gbdt, xgboost, and lightgbm algorithms. *Mathematics*, 8(5), 765. <https://doi.org/10.3390/math8050765>
- Oguine, O. C., Oguine, K. J., Okorie, C. I., & Oguine, M. B. (2021). Comparative analysis and forecasting on the death rate of COVID-19 patients in Nigeria using random forest and multinomial Bayesian epidemiological models. *Journal of Clinical Case Studies Reviews & Reports*, (3), 182. [https://doi.org/10.47363/JCCSR/2021\(3\)182](https://doi.org/10.47363/JCCSR/2021(3)182)
- Pachón, C. G., Ballesteros, D. M., & Renza, D. (2021). Fake banknote rRecognition using deep learning. *Applied Sciences*, 11(3), 1281. <https://doi.org/10.3390/app11031281>
- Speer, A. B. (2021). Empirical attrition modelling and discrimination: Balancing validity and group differences. *Human Resource Management Journal*, 34(1), 1-19. <https://doi.org/10.1111/1748-8583.12355>
- Wang, L., Wang, X., Chen, A., Jin, X., & Che, H. (2020). Prediction of type 2 diabetes risk and its effect evaluation based on the xgboost model. *Healthcare*, 8(3), 247. <https://doi.org/10.3390/healthcare8030247>
- Woldehana, Y. D., Yimer, M. A., & Molla, A. M. (2025). An Explainable counterfeit and genuine ethiopian banknote classification. *Journal of Emerging Computer Technologies*, 5(1), 24-35. <https://doi.org/10.57020/ject.1579598>

- Wu Y. (2023). Job embeddedness review: Presentation, measurement and development. *Advances in Economics, Management and Political Sciences*, 47(1), 169- 174. <https://doi.org/10.54254/2754-1169/47/20230393>
- Yildiz, A., Almisreb, A. A., Dzakmic, Š., Tahir, N. M., Turaev, S., & Saleh, M. A. (2020). Banknotes counterfeit detection using deep transfer learning approach. *International Journal of Advanced Trends in Computer Science and Engineering*, 9(5), 8115–8122. <https://doi.org/10.30534/ijatcse/2020/172952020>