İÇİNDEKİLER

1. GİRİŞ

Binlerce yıldır insanların ihtiyaçlarını karşılamak amacıyla kullandıkları alışveriş eylemi, teknolojinin gelişmesi ile bağlantılı olarak ilerlemiştir. Günümüzde sanal ortamlarda da sıklıkla gerçekleştirilen alışveriş, beraberinde ödeme yönteminin de sanallaşmasını gerektirmiştir. Online ödeme işlemi için kullanılan yöntemlerden biri de kredi kartlarıdır.

5464 sayılı Banka Kartları ve Kredi Kartları Kanunu'nun üçüncü maddesinin e bendinde kredi kartı: "Nakit kullanımı gerekmeksizin mal ve hizmet alımı veya nakit çekme olanağı sağlayan basılı kartı veya fiziki varlığı bulunmayan kart numarasını" ifade eder şeklinde tanımlanmıştır. Bu tanımdan da yola çıkarak insanların alışverişlerinde kullanmasına, nakit para çekmesine ve son zamanlarda özellikle artan temassız ödeme olanaklarını sağladığı için kredi kartı kullanımına yönelik taleplerin arttığı gözlemlenebilmektedir. Bu durum kötü niyetli kişiler tarafınca da fark edildiğinden dolayı, kredi kartı dolandırıcılığı, fiziki kredi kartı hırsızlığı ve kart bilgilerinin çalınması gibi olumsuz davranışlar da artış göstermiştir. Dolandırıcılık eylemi sonucunda şahıslar, mal veya hizmet sunan işletmeler ve bankalar bu durumdan etkilenmektedir [1]. Bu olumsuz durumların tespiti ve önüne geçilmesi için geleneksel veya çeşitli yazılım yöntemleri kullanılmalıdır. Sanal ortamda ödeme imkanlarının artması ile, dolandırıcılık yöntemlerinin de doğru orantılı olarak artması, dolandırıcılık eyleminin gerçek zamanlı olarak tespitini ve önüne geçilmesini zorlaştırmaktadır. Bu nedenle bu çalışmada, "Kredi Kartı Dolandırıcılık Tespitinde Yapay Zeka Temelli Yöntemler" ele alınmıştır.

Kredi kartı dolandırıcılık tespitinde yapay zeka yöntemlerinin uygulanabilmesi için kredi kartı ve kullanıcının alışverişlerinden yola çıkarak elde edilen veriler kullanılarak analizler yapılmaktadır. Bu analizlerden yola çıkarak geliştirilen algoritma ile gerçek zamanlı ve yapay zeka temelli bir dolandırıcılık tespiti yöntemi geliştirilmektedir.

2. LİTERATÜR TARAMASI

Literatürde bulunan çeşitli makaleler incelendiğinde, kredi kartı dolandırıcılık tespitinde sıklıkla kullanılan algoritmaların makine öğrenme ve derin öğrenmeye yönelik algoritmalar oldu tespit edilmiştir. Bu bölümde, literatürde karşılaşılan makaleler incelenip yorumlanmıştır.

2.1. APATE: A novel approach for automated credit card transaction fraud detection using network-based extensions

Düzensiz davranışları yakalayabilecek önemli özellikleri tespit edebilmek, dolandırıcılık tespit sürecindeki en önemli adımlardan biridir. Bu çalışmada, içsel ve ağla ilgili özellikler birleştirilmiştir. İçsel özellikler, işlemi analiz eder ve işlemin normal müşteri profiline uyup uymadığını karşılaştırır. Bu özellikler, kredi kartı sahibinin geçmişte yaptığı işlemlerin RFM değerleri (Yenilik, Sıklık ve Parasal Değer) türetilerek oluşturulur. Ağ tabanlı özellikler ise, işlemler aracılığıyla ilişkili kredi kartı sahipleri ve üye işyerlerinden oluşan bir ağ oluşturup bu ağı analiz ederek her işlemi karakterize eder. Büyük bir Belçikalı kredi kartı

düzenleyicisinden elde edilen ve yaklaşık 3,3 milyon işlemden oluşan benzersiz bir veri seti ile sınırlı sayıda onaylanmış dolandırıcılık işlemi kullanılarak ağ üzerinden hileli etkiyi yaymak için toplu bir çıkarım algoritması kullanılmakta ve bir ifşa puanı türeterek her bir ağ nesnesinin şüpheliliğine karar verilmektedir. Önerilen yöntem logistic regression, neural networks, random forests modelleriyle karşılaştırılmış olup 0.98'den yüksek AUC skoru elde edilmiştir [2].

2.2. Fraud detection using self-organizing map visualizing the user profiles

Bu makalede, kullanıcı hesaplarının görselleştirilmesine ve eşik tipi tespitine dayalı bir dolandırıcılık tespit yöntemi önerilmektedir. Yaklaşımda kullanılan görselleştirme tekniği Kendi Kendini Düzenleyen Harita'dır (SOM). SOM tekniği orijinal haliyle sadece vektörleri görselleştirdiğinden ve çalışma içerisinde kullanıcı hesapları, kullanıcı sıralı etkinliklerini yansıtan bir kayıt koleksiyonunu depolayan matrisler olarak temsil edildiğinden, SOM ızgarasında matris görselleştirmesi için bir yöntem önerilmektedir. Ayrıca, SOM U matrisi temelinde bir algılama eşiği ayarı yöntemi önerilmektedir. Kullanıcı hesaplarının SOM'a yansıtılmasının ardından, eşik tipi ikili sınıflandırma algoritması kullanılarak hileli hesaplar tespit edilmiş ve belirli bir SOM'nin U matrisindeki çıkıntıyı bularak sınıflandırma eşiği ayarı için bir yöntem önerilmiştir. Görselleştirme yoluyla dolandırıcılık tespiti, 2 boyutlu bir uzaya yansıtılan yüksek boyutlu verinin mümkün olduğu gerçeğinden dolayı faydalı ve çekici bir veri analizi yaklaşımıdır. Çalışma, 1.01.2005 ile 1.03.2005 tarihleri arasında Polonya'nın Varşova şehrinde 10.000 kredi kartı sahibinin hesabından oluşturulan ve işlemde harcanan para miktarı, işlemin gerçekleştirildiği yer, işlem zamanı özellikleriyle karakterize edilmiş bir veri seti üzerinde sürdürülmüştür. SOM-clustering-based, GHSOM-based, GMM-based metodlarıyla karşılaştırılan yöntem 1.0 değerinde F1-score elde etmiştir [3].

2.3. Critical Analysis of Machine Learning Based Approaches for Fraud Detection in Financial Transactions

Bu makalenin amacı, finansal işlemler için bir dolandırıcılık tespit sistemine entegre edilebilecek seçilmiş makine öğrenimi ve aykırı değer tespit tekniklerini gözden geçirmektir. Karşılaştırma öncesinde dolandırıcılık analizinin ve veri setlerinin özellikleri hakkında modeli eğitime hazırlama süreci ve eğitim sonrası performans değerlendirmesi özelinde birtakım açıklamalar yapılmıştır. Ardından; Bayesian Networks, Recurrent Neural Networks, Support Vector Machines, Fuzzy Logic, Hidden Markov Model, K-Means Clustering, K-Nearest Neighbor gibi çeşitli makine öğrenme algoritmaları ve bunların dolandırıcılık tespit alanındaki mevcut uygulamaları en iyi yaklaşımı bulmak amacıyla her modelin olumlu ve olumsuz özellikleri göz önünde bulundurularak karşılaştırmalı olarak tartışılmıştır [4].

2.4. A cost-sensitive decision tree approach for fraud detection

Bu çalışmada, her bir terminal olmayan düğümde ayırma özniteliği seçilirken yanlış sınıflandırma maliyetlerinin toplamını en aza indiren, maliyete duyarlı yeni bir karar ağacı

yaklaşımı geliştirilmiş ve bu yaklaşımın performansı iyi bilinen geleneksel sınıflandırma modelleriyle (YSA, SVM) karşılaştırılmıştır. Sunulan yaklaşımda, yanlış sınıflandırma maliyetleri değişken olarak alınır. Sonuçlar, bu maliyete duyarlı karar ağacı algoritmasının, doğruluk ve gerçek pozitif oran gibi iyi bilinen performans ölçütlerine ve aynı zamanda yeni tanımlanmış bir maliyete duyarlı ölçüt özgüllüğüne göre verilen problem setinde mevcut iyi bilinen yöntemlerden daha iyi performans gösterdiğini göstermektedir. Buna göre, hile tespit sistemlerinde bu yaklaşımın uygulanması ile hileli işlemlerden kaynaklanan mali kayıplar daha da azaltılabilmektedir. Gerçek veriler üzerinde çalışılan yöntemin karar ağaçları, YSA ve SVM gibi klasik modellerden daha iyi sonuç verdiği gözlemlenmiştir [5].

2.5. An Efficient Real Time Model For Credit Card Fraud Detection Based On Deep Learning

Bu çalışma, derin öğrenme paradigmalarını araştıran karşılaştırma çalışmalarının az ve gerçek zamanlı yaklaşımlara yeterli önemin verilmemesi dolayısıyla ortaya çıkmış olup kredi kartı dolandırıcılığı problemi ile başa çıkmak amacıyla derin sinir ağı (deep neural network) teknolojisine dayanan canlı bir kredi kartı dolandırıcılık tespit sistemi önermektedir. Çalışma içerisindeki gerçek zamanlı veri akış hatlarının oluşturulması Kafka teknolojisi ile gerçekleştirilmiştir. Kullanılan veri seti, Worldline ve ULB'nin Machine Learning Group'un büyük veri madenciliği ve dolandırıcılık tespiti üzerine bir araştırma işbirliği sırasında toplanan ve analiz edilen, Eylül 2012'de Avrupa kartları tarafından iki gün içinde yapılan işlemleri içermektedir. Yapılan testler ve bazı tipik gerçek zamanlı ikili sınıflandırıcıların Deep NN Auto encoders'a karşı; Linear SVM Regression, Logistic Regression, NN Based Classification ve Non Linear Auto Regression ile oluşturulan performans bazlı karşılaştırma çalışmasından sonra, kıyaslama deneyleri Deep NN Auto encoders'ın F1 puanına dayanarak umut verici sonuçlara sahip olduğunu göstermektedir [6].

2.6. A Review of Credit Card Fraud Detection Techniques

Bu çalışmada Hindistan öncelikli olmak üzere dünya üzerinde gerçekleştirilen dolandırıcılık türleri ve bu türlere karşın kullanılabilecek makine öğrenmesi algoritmaları karşılaştırılmıştır. Karşılaştırma sonucunda her algoritmanın odak noktası ve zayıf yönleri üzerinde durulmuştur. Random Forest ve KNN gibi birkaç strateji, küçük bir veri kümesinde iyi çalışır fakat büyük bir veri kümesinde yeterince esnek değildir. SVM ve Karar Ağacı gibi bazıları ön işlemeden geçmiş ve örneklenmiş veriler üzerinde iyi sonuçlar verirken, ANN gibi algoritmalar önceki durumdan öğrenir ve Genetik Algoritma tespit etmede hızlıdır. Bulanık sistemler ve Lojistik Regresyon gibi bazı teknikler, ham örneklenmemiş verilerle daha iyi sonuç hataları verir. Çalışmada sonuç olarak, asıl üzerinde durulması gereken problemin dolandırıcılık işleminin gerçek zamanlı tespit edilmesi olduğu vurgulanmıştır [7].

2.7. Real Time Credit Card Fraud Detection on Huge Imbalanced Data using Meta-Classifiers

Bu çalışmada, dengesiz ve büyük verilerdeki dolandırıcılık eylemlerini tespit etmek amacıyla kullanılabilecek gerçek zamanlı ağaç tabanlı bir meta sınıflandırıcı TBMC sunulmaktadır. Geliştirilen meta sınıflandırıcı tabanlı model, iki düzeyde tahminlere dayalı olarak çalışmaktadır. Birinci seviyedeki tahminler Random Forest sınıflandırıcısı tarafından, ikinci seviyedeki tahminler ise Karar Ağaçları ve Gradient Boosted Trees ile oluşturulan bir topluluk tarafından gerçekleştirilmektedir. İki seviyede elde edilen sonuçlar ise, nihai tahminleri oluşturmak amacıyla birleştirilmektedir. Önerme için UCSD-FICO verileriyle yapılan deneyler sonrasında sonuçlar mevcut modellerle karşılaştırılmıştır. Deneysel sonuçlar, geliştirilen TBMC modelinin performansının arttığını göstermektedir. Bununla birlikte, geliştirilen model, dengesiz veriler için uzmanlaşmış dengeleme önlemleri olan orta düzeyde MCC ve BCR sergilemektedir [8].

2.8. Real-time Credit Card Fraud Detection Using Machine Learning

Bu çalışma, gerçek dünyadaki işlemlerde tespit edilen dört ana dolandırıcılık olayına odaklanmaktadır. Çalışma dahilinde her bir dolandırıcılık örneğinin makine öğrenmesi algoritmaları kullanılarak incelenmesi, bu şekilde en iyi yöntemin seçilmesi ve hile türlerine göre en uygun algoritmanın seçilmesi için kapsamlı bir kılavuz oluşturulması amaçlanmıştır. İki veri kaynağının birleştirilmesiyle oluşan projede ele alınan bir diğer önemli unsur da hilelerin gerçek zamanlı tespitidir. Bunun için, gerçekleştirilen eylemin hileli olup olmadığına karar verilmesi amacıyla örnekleme yöntemleri ile değerlendirilen, çarpık dağılıma sahip veri üzerinde tahmine dayalı analitik ve bir API modülü uygulanmıştır. Bu uygulama sonucu dört dolandırıcılık modelinde (100$'ın üzerindeki İşlem, Bilinmeyen web adresi, Riskli MCC, ISO-Yanıt Kodu) en yüksek doğruluk oranlarına (sırasıyla %74, %72, %83, ve %91) makine öğrenmesi modellerinin LR, LR, NB ve SVM olduğu gözlemlenmiştir [9].

2.9. Data mining for credit card fraud: A comparative study

Bu makale, kredi kartı sahtekarlığını daha iyi tespit etme (ve dolayısıyla kontrol etme ve kovuşturma) girişiminin bir parçası olarak iki gelişmiş veri madenciliği yaklaşımını, destek vektör makinelerini ve rastgele ormanları, iyi bilinen lojistik regresyonla birlikte değerlendirir. Çalışmada, Ocak 2006-Ocak 2007 dönemine ait uluslararası bir kredi kartı operasyonundan elde edilen işlemler kullanılmaktadır. Performans değerlendirmesi için, farklı düşük örnekleme seviyelerine sahip eğitim veri kümelerinden çok daha düşük sahtekarlık oranına (% 0,5) sahip bir test veri kümesi kullanılmıştır. Duyarlılık, G-mean ve ağırlıklı-doğruluk, eğitim verilerinde daha düşük dolandırıcılık oranları ile azalırken, kesinlik ve özgüllüğün ters bir eğilim gösterdiği; F-score ve AUC'de lojistik regresyon, eğitim verilerinde değişen oranlarda dolandırıcılık ile benzer performansı korurken, RF ve SVM, AUC'de azalan bir eğilim ve F'de artan bir eğilim göstermiştir. Lojistik regresyon, farklı düşük örnekleme seviyeleri ile benzer performansı korurken, üst dosya derinliklerindeki

SVM performansı, eğitim verilerinde daha düşük dolandırıcılık oranı ile artma eğilimi göstermiştir. Rastgele ormanlar, üst dosya derinliklerinde çok daha yüksek performans göstermiştir. Böylece, üst derinliklerde daha az yanlış pozitif ile daha fazla dolandırıcılık vakasını yakamaktadırlar; bu, dolandırıcılık tespit modellerinin gerçek hayatta kullanımında önemli bir hususdur. Bu çalışmada tekniklerin parametrelerini optimize etmek için kasıtlı bir girişimde bulunulmamıştır. Parametre ayarlama DVM için önemli olabilir ve dengeli örneklemenin, dengesiz veriler üzerinde Rastgele Ormanların kullanılmasında avantajlı olduğu not edilmiştir [10].

2.10. Yapay Zeka Destekli ÇOKS Yöntemi ile Kredi Kartı Sahtekarlığının Tespiti

Bu makalede; öncelikle kredi kartının günlük yaşamdaki kullanıma ve dolandırıcılık terimlerine değinilmiştir. Dolandırıcılık eyleminin önlenmesi ve ayırt edilebilmesindeki önem vurgulanmıştır. Kredi kartı sahteciliğini tespite yönelik araştırmalarda çoğunlukla makine öğrenmesi algoritmalarından ve farklı sınıflandırma algoritmalarının kullanıldığı görülmektedir. Fakat kredi kartı sahteciliğini tespit etmek amacıyla karar ağacı, KNN ve naive bayes classifierlarının bir arada kullanıldığı bir çalışmanın literatürde mevcut olmadığı belirtilmiştir. Bu gözlemden yola çıkarak bu makaleye yönelik yapılan çalışma kapsamında kredi kartı dolandırıcılığa yönelik karar ağacı, KNN ve naive bayes makine öğrenmesi algoritmalarından yararlanan ve Çoğunluk Oyu ile Karar Verme Sistemi (ÇOKS) olarak adlandırılan yeni bir sezgisel algoritma geliştirilmiştir. Bu algoritmanın ortak karar verme mekanizması için de bir sayısal devre tasarımı lojik fonksiyonu olan çoğunluk fonksiyonundan faydalanılmıştır. ÇOKS'nin etkinliği her biri 30 farklı özniteliğe sahip 284,807 kredi kartı işleminin yer aldığı bir veri kümesi üzerinde test edilmiştir. Yürütülen testler finansal güvenliği hedefleyen bu yeni yöntemin %99,93 doğruluk oranı, %95,60 kesinlik oranı ve %80,0 ROC AUC değeri ile veri kümesindeki bir işlemi sahte veya yasal işlem olarak sınıflandırabilmeyi başarmıştır [11].

2.11. Credit Card Fraud Detection using Machine Learning Algorithms

Bu makale, kredi kartı dolandırıcılıklarının kolaylaşmasından ve çevrimiçi ödeme kullanılan sitelerin artmasının dolandırıcılık riskini de artırdığından bahsetmektedir. Dolandırıcılık oranlarındaki artış, farklı makine öğrenimi yöntemlerini kullanarak kredi kartı sahtekarlıklarını tespit ve analiz etmeye yönlendirmiştir. Makale kapsamında yapılan çalışmaların temel amacı, müşterilerin geçmiş işlem detaylarını analiz etmek amacıyla Streaming Transaction Data için yeni bir dolandırıcılık tespit yöntemi geliştirmektir. Bu kapsamda öncelikle kart sahiplerinin işlem tutarlarına göre, işlemler farklı gruplara ayrılır. Daha sonra farklı gruplardan yapılan işlemleri bir araya getirmek için kayan pencere stratejisi kullanılarak, gruplar ayarlanır Sonrasında ise farklı sınıflandırıcılar, gruplar üzerinde ayrı ayrı eğitilir. Böylelikle en iyi sınıflandırma yöntemi seçilir ve kavram kayması problemini çözmek için bir geri besleme mekanizması oluşturulmuş olur. Sonuç olarak bu makalede, müşteriler işlemlerine göre gruplanır ve her kart sahibi için bir profil oluşturularak kişilerin davranış kalıplarının çıkartıldığı yeni bir dolandırıcılık tespiti yöntemi geliştirilmiştir. Daha sonra bu veriler üç farklı gruba ayrılarak farklı sınıflandırma teknikleri uygulanır ve daha sonra her

sınıflandırıcı türü için puanlama yapılır. Araştırma sonunda Lojistik regresyon, karar ağacı ve random forest algoritmalarının daha iyi sonuç veren algoritmalar olduğunu gözlemlenmiştir [12].

2.12. Credit Card Fraud Detection Using Autoencoder Neural Network

Bu makalede öncelikle kredi kartı dolandırıcılıklarının artma nedenlerinden ve kredi kartı dolandırıcılık nedenlerinden bahsedilmiştir. Bunun asıl sebebinin güvenlik açıkları olduğu belirtilmiştir. Bundan dolayı kredi kartı dolandırıcılık tespitinin önemine değinilmiştir. Tespit için kullanılan yöntemlerin genellikle big data ve makine öğrenmesi algoritmaları olduğu bilinmektedir. Burada kullanılan classification problemlerinin en büyük sorunu büyük çaplı veri setlerindeki dengesiz verilerdir. Çünkü bu veri setlerinde asıl ilgilenilmek istenen konu dolandırıcılık verileriyken, yasal olan verilerin sayıları çok çok daha fazladır. Bu nedenle bu makalede gürültüden arındırılmış outoencoder ve oversampling üzerine durulmuştur. Yeni azınlık sınıfı örneklerini sentezlemek için oversampling algoritması kullanılır, ancak bu algoritmanın problemi gürültü ihtimalidir. Bu makalede bu gibi gürültü sorunlarına değinilmiş, yalnızca azınlık sınıfı örneğini yanlış sınıflandırma maliyetiyle oversampling yapmayan aynı zamanda örneklenen veri kümesini gürültüden arındırabilen ve sınıflandırabilen bir gürültü giderme autoencoder sinir ağı (DAE) algoritması geliştirilmiştir [13].

2.13. Spotting Collective Behaviour of Online Frauds in Customer Reviews

Bu makalede; öncelikle dolandırıcılık faliyetlerinden ve özellikle mail veya siteler aracılığıyla insanları kandırmaya çalışan toplu dolandırıcılık faliyetlerinden bahsedilmiştir. Sonrasında ise grup spam tespitinin, grubun belirsiz tanımlanması, gruplar arası dinamiklerin çeşitlive etiketli grup düzeyinde spam verilerinin az olması nedenleriyle bireysel sahtekarlık tespitinden daha zor olduğu anlatılmaktadır. Bu makalede, dolandırıcılık işlemini tespit etmek için unsupervised bir yöntem olan DeFrauder'ı önerilmektedir. DeFrauder yöntrmi; İlk olarak, temel ürün inceleme grafiğinden yararlanır ve gözden geçirilen veriler arasında çok yönlü işbirliğini modelleyen çeşitli davranışsal sinyalleri birleştirerek olması muhtemel dolandırıcılık gruplarını tespit eder. Ardından, gözden geçirenleri bir yerleştirme alanına eşler ve her gruba bir spam puanı atar, böylece oldukça benzer davranışsal özelliklere sahip spam göndericilerden oluşan gruplar yüksek spam puanı elde eder. Ve Buna göre dolandırıcılık puanı verilmiş olur.

Dört gerçek dünya veri setindeki beş temel ile karşılaştırıldığında, DeFrauder veri setlerinde %17,11 daha yüksek NDCG@50 (ortalama olarak) ile en iyi temel çizgiden daha iyi performans göstererek üstün performans gösterir. Veri setleri olarak ise; Amazon, Playstore, YelpNYC ve YelpZip verisetlerinden yararlanılmıştır [14].

2.14. Kredi Kartı Dolandırıcılık Tespitinin Makine Öğrenmesi Yöntemleri İle Tahmin Edilmesi

Bu makalede kredi kartı, kredi kartı kullanımı ve dünyadaki yerinden genel olarak bahsedilmiş ve haliyle artan dolandırıcılık girişimlerinin tespitinin önemine değinilmiştir. Günümüzde herhangi bir kredi kartı bilgilerine ulaşmanın oldukça kolaylaşması kredi kartı dolandırıcılarına fırsat sunmaktadır. Bunun önüne geçebilmek için, kredi kartı ile gerçekleştirilen hesap hareketleri değişikliğinde zaman ve harcamaların analiz edilmesi sayesinde dolandırma amaçlı verilerin analizi ile dolandırıcılığı önlemeye yönelik yöntemler geliştirilebilinir edebilir. Bu makale kapsamında yapılan çalışmalarda Kaggle veritabanından elde edilen Kredi Kartı Dolandırıcılık Teşhis veri seti kullanılarak Çok Katmanlı Yapay Sinir Ağı ve Naive Bayes yöntemleri ile modelleme yapılmıştır.

Bu çalışma sonucunda kişilerin kredi kartlarını kullanma zaman aralıkları Naive Bayes ve çok katmanlı yapay sinir ağları yöntemleriyle analiz edilerek yaptıkları işlemin farklı kişi tarafından yapıldığını tespit etmek amaçlanmıştır. Sonuç olarak birbirlerine kıyasla çok katmanlı yapay sinir ağlı ile daha yüksek bir başarı oranı edilmiştir [15].

2.15. Credit Card Fraud Detection Using Machine Learning: A Study

Bu makalede kredi kartının tanımından ve kullanım alanlarından bahsedilerek genel bir giriş yapılmıştır. Kredi kartı dolandırıcılıkları kategorilerine ayrılmış ve açıklanmıştır. Bunlar; Kart sahibi ile, satıcı tarafından ve zorla el koyarak dolandırıcılık şeklindedir. Dünyanın hızla dijitalleşmeye doğru ilerlemesi ve para işlemlerinin nakit kullanmadan gerçekleşmeye artarak devam etmesiyle birlikte kredi kartı kullanımı oldukça artmıştır. Bununla ilişkili dolandırıcılık faaliyetleri artmakta ve bu da banka ve şirketler için büyük bir kayba yol açmaktadır. Bu nedenle dolandırıcılık işlemleri, dolandırıcılık olmayan normal işlemlerden analiz edip tespit etmek gerektiğinden bahsedilmiştir. Bu makalede dolandırıcılık olaylarını farkedip önüne geçmek için kullanılabilecek yöntemlerden bahsedilmiştir. Bu metodolojiler arasında Gizli Markov Modeli (HMM), Karar Ağaçları, Random Trees, Bayesian Belief Networks, Genetik algoritmalar, Lojistik Regresyon, Destek Vektör Makineleri (SVM), KNN, Fuzzy Clustering ve Neural Networks bulunmaktadır. Makalede bu tekniklerin kapsamlı bir analizi sunulmaktadır.

Sonuç olarak bu makale, kredi kartı dolandırıcılık işlemini anlamamızı ve dolandırıcılık işlemini, normal işlemlerden ayırmamıza yardımcı olabilecek algoritmaları tanımamızı, aralarındaki farkları, avantaj ve dezavantajlarıyla anlatarak doğru algoritmayı seçme konusunda yardımcı olmaktadır [16].

2.16. Credit Card Fraud Detection in e-Commerce: An Outlier Detection Approach

Bu makalede dolandırıcılık ve spam e-posta belirleme görevlerindeki zorluğun uygun denetimli öğrenme modellerini eğitmek için gereken tüm olası kalıpların olmamasından kaynaklandığı vurgulanmıştır. Bu sorun, dolandırıcılık örneklerinin az ve aynı zamanda zamanla değiştiği durumlar olduğundan dolayı daha da göze çarpmaktadır. Dolandırıcılık

modelindeki değişiklik, dolandırıcıların sahtekarlığı önlemek için alınan önlemleri atlatmak için yeni yollar geliştirmeye devam etmesinin sonucu olarak ortaya çıkmaktadır. Sınırlı veri ve sürekli değişen yöntemler makinelerin öğrenmesini zorlaştırmaktadır. Bu makale kapsamında yasal davranışın zamanla değişmediğini ve yasal davranışı temsil eden veri noktalarının farklı gruplamalar altında tutarlı uzamsal imzaya sahip olduğunu varsayılmıştır. Bu hipoteze dayanarak, bir clustering yöntemleri topluluğu kullanarak her bir veri noktasına bir tutarlılık puanı atayarak büyük veri kümelerindeki aykırı değerleri tespit eden bir yaklaşım önerilmiştir. Bu çalışmadaki asıl hedef, büyük veri kümelerindeki aykırı değerleri tespit edebilen ve değişen dolandırıcılık kalıplarına karşı dayanıklı yeni bir yöntem önermektir. Ayrıca bu makalede, aykırı değer saptama yöntemlerini değerlendirmek için yaygın olarak kullanılan bir metrik olsa da, ROC eğrisi altındaki alanın doğru metrik olmadığı savunulmaktadır. Aykırı değer algılama sorunlarının sınıfların çarpık dağılımına sahip olduğundan, kesinlik-hatırlama eğrileri daha uygundur çünkü kesinlik, yanlış pozitifleri gerçek negatiflerle (aykırı değerler) karşılaştırır ve bu nedenle sınıf dengesizliği sorunundan etkilenmez. Makalenin devamında kesinlik-hatırlama eğrisinin altındaki alanın bir değerlendirme ölçüsü olarak ROC'den daha iyi olduğu belirtilmektedir.

Veri Setleri olarak; Landsat uydu veri setinin değiştirilmiş versiyonu, ann-tiroid veri setinin değiştirilmiş versiyonu ve Kaggle aracılığıyla mevcut olan büyük bir gerçek dünya kredi kartı dolandırıcılık tespit veri setleri kullanılmıştır [17].

2.17. Detection of Credit Card Fraud in E-Commerce Using Data Mining

Bu makalede e-ticaret sitelerinin en çok tercih edilen yönteminin kredi kartları ile olduğu, kredi kartı bilgilerinin çalınması halinde de dolandırıcılık işlemlerinin gerçekleşebileceği vurgulanmıştır. Dolandırıcılık işleminin önüne geçmek için siparişlerin incelenmesi gerektiği üzerine durulmuştur. Dolandırıcılık şüphesi olan siparişler, sadece dolandırılan kişi ve banka için değil aynı zamanda alışveriş siteleri için de büyük endişe kaynağıdır. Sahtekarlık işlemleri sadece müşterileri değil, aynı zamanda şirketleri ve bankaları da etkiler. Bu nedenle, e ticaret siteleri siparişleri kategorize etmeli ve şüpheli işlemlere karşı önlemler almalı ve üzücü bir durum yaşanmadan onu engelleyebilmelidir. E-ticaret sitelerinde müşteriler hakkında daha az bilgi olması nedeniyle sınıflandırma işlemini yapabilmek daha zordur. Bu makale kapsamında yapılan çalışmada, bir e-ticaret sitesinin gerçek sipariş verileri incelenmiş ve şüpheli işlemler belirlenmiştir. Öncelikle, tüm sipariş verileri analiz edilip, filtrelenmiştir. Sınıflandırma için en iyi değişkenler değişken seçim algoritmaları ile belirlenmiştir. Daha sonra sınıflandırma algoritmaları uygulanmış ve %92 başarı oranı ile şüpheli siparişler belirlenmiştir. Karşılaştırmalı veri madenciliği yöntemleri olarak Naive Bayesian, Karar Ağaçları ve Yapay Sinir Ağı kullanılmıştır.

Sonuç olarak bu makale kapsamında e-ticaret sitesi için veri elde etmek amacıyla manuel olarak dolandırıcılık işlemleri yapılmıştır. Bu veriler sayesinde yapılan çalışmalar sonucunda elde edilen sınıflandırma verileri ile e-ticaret sitelerinin iş süreçlerini büyük ölçüde iyileştirmesi hedeflenmiştir [18].

3. GELİŞTİRİLEN YAKLAŞIM VE BULGULAR

Tablo 3.1. Yapay Zeka Temelli Gerçek Zamanlı Dolandırıcılık Tespit Sistemi Gösterimi

Öncelikle ilk ve en önemli adımlardan biri olan çalışılacak veri setlerinin bulunması üzerine Kaggle [19], Google Dataset Search [20], UCI Machine Learning Repository [21], OpenML [22], DataHub [23], Papers with Code [24], EU Open Data Portal [25], Awesome Public Datasets [26] gibi çeşitli kaynaklardan detaylı bir araştırma süreci gerçekleştirilmiştir. Bu doğrultuda bulunan veri setlerinin özellikleri aşağıdaki tabloda (Tablo 3.2.) gösterilmiştir. Hem e-ticaret ve bankacılık olmak üzere çeşitli sektörde kullanılan hem de akademik kaynaklardan elde edilen veri setlerinin yapısal olarak incelenmesi ve işleme sürecinden önce veriler üzerinde yorum yapılabilmesi amacıyla veri setinin tanınması amaçlanmıştır. Tanıma süreci; satır ve sütun özelliklerinin/sayılarının, veri tiplerinin, eksik/aykırı yahut tekrar eden değerlerin, veri dengesizliğinin ve verilerin betimsel istatistik değerlerinin incelenmesi vb. şeklinde ilerlemiştir. Bir sonraki aşamada ise elde edilen bilgiler görselleştirilerek veri setleri hakkında genel bilgi edinilmesi sağlanmıştır.

İsim	Sektör	Boyut (satır, sütun)	Özellikler	Gizli mi?	Referanslar
Credit Card Fraud Detection [27]	Bankacılık	(284807, 31)	Time, Amount, Class, V1-V28*	Evet	[28-43]
Credit Card Transactions [44]	Bankacılık	(19999999, 15)	User, Card, Year, Month, Day, Time, Amount, Use Chip, Merchant Name, Merchant City, Merchant State, Zip, MCC, Errors, Is Fraud?	Hayır	[45]
Simulated Credit Card Transactions generated using Sparkov [46]	Bankacılık	(1296675, 23), (555719, 23)	trans_date_trans_time, cc_num, merchant, category, amt, first, last, gender, street … lat, long, city_pop, job, dob, trans_num, unix_time, merch_lat, merch_long, is_fraud	Hayır	[45-48]
Fraud E- Commerce [49]	E-Ticaret	(151112, 11), (138846, 3)	user_id,, signup_time, purchase_time, purchase_value, device_id, source, browser, sex, age, ip_address, class	Hayır	[48]
IEEE CIS Fraud Detection [50]	E-Ticaret	(590540,433) (506691,432)	TransactionDT, TransactionAMT, ProductCD, card1- card6, addr, dist, email domain, C1-C14*, D1-D15*, M1-M9*, Vxxx …	Evet	[42]
Synthetic Data from a Financial Payment System [51]	Bankacılık	(594643, 10)	step, customer, age, gender, zipcodeOri, merchant, category, zipMerchant, amount, fraud	Hayır	[52-53]
Synthetic Financial Datasets For Fraud Detection [54]	Bankacılık	(6362620, 11)	PAYMENT, step, type, amount, nameOrig, oldbalanceOrg, newbalanceOrig, nameDest, oldbalanceDest, newbalanceDest, isFraud, isFlaggedFraud	Hayır	[55-56]

* Gizlilik nedeniyle sütunların ne olduğu bilinmemekte.

Tablo 3.2. Seçilen veri setlerinin karşılaştırılması

Makine öğrenmesi projelerinde en yüksek performansı sağlayabilmek adına üzerinde çalışılan veri setinin algoritmaya en düzgün şekilde verilmesi elzemdir. Bu şartı sağlayabilmek adına ise veri setlerinin belirli ön-işleme adımlarına tabi tutulması gerekmektedir. Bununla birlikte, dolandırıcılık durumlarının dengesiz olması nedeniyle veri setinin dengelenmesi gerekmektedir. Bu ön-işleme adımlarından; gürültülü/eksik/aykırı verilerin tanımlanması ve çeşitli yöntemlerle manipüle edilmesi, veri setinin çeşitli tekniklerle (oversampling, undersampling, SMOTE vb.) dengeli hale getirilmesi şeklinde bahsedilebilir.

Gerçek hayatta kullanılan veri setlerinde istenilen yahut işe yarar tüm veriler ilk bakışta görülmeyebilir. Tüm sütunların herhangi bir eleme işleminden geçmeksizin algoritmaya verilmesi de model performansını ve çalışma zamanını olumsuz etkilemektedir. Modeli en optimal haline getirebilmek ve veri setlerinden özellik seçebilmek amacıyla belirli özellik seçme yöntemleri uygulanıp veri setleri eğitime hazır hale getirilmiştir.

Veri setinin eğitime uygun hale getirilmesi sonrasında, literatür taramasında detaylıca araştırılmış olan makaleler tablo haline getirilmiş olup (Tablo 3.3.) karşılaştırmak amacıyla en fazla kullanılan algoritmalar renklendirilmiştir. Akademik çalışmalarda en fazla kullanılan

algoritmalar **Logistic Regression, Neural Network, Random Forest Model, Support Vector Machines, K-Nearest Neighbor, Decision Tree, Naive Bayes** olarak belirlenmiş olup performans metrikleri bu algoritmalar özelinde değerlendirilmiştir. Daha önce bahsedildiği üzere veri setinin dengesiz olması dolayısıyla "accuracy" değeri performans metriklerinden biri olarak gösterilmemiş olup model başarısı **F1-Score, K-Fold Cross Validation, ROC-AUC** gibi çeşitli yöntemlerle ölçülmüştür.

Makale / Algoritmalar	logistic regression	neural network	random forest model	Support Vector Machines	K-Nearest Neighbor	Decision Tree	Naive Bayes
APATE: A novel approach for automated credit card transaction fraud detection using network-based extensions [1]	x	x	x				
Critical analysis of machine learning based approaches for fraud detection in financial transactions [3]				x	x		
A cost-sensitive decision tree approach for fraud detection [5]		x		x			
An efficient real time model for credit card fraud detection based on deep learning [6]	x	x		x			
A Review of Credit Card Fraud Detection Techniques [7]	x	x	x	x	x	x	
Real time credit card fraud detection on huge imbalanced data using meta-classifiers [8]			x			x	
Real-time credit card fraud detection using machine learning [9]	x			x			x
Data mining for credit card fraud: A comparative study [10]	x		x	x			
Yapay Zeka Destekli ÇOKS Yöntemi ile Kredi Kartı Sahtekarlığının Tespiti [11]					x		x
Credit Card Fraud Detection using Machine Learning Algorithms [12]	x		x			x	
Kredi Kartı Dolandırıcılık Tespitinin Makine Öğrenmesi Yöntemleri ile Tahmin Edilmesi [15]		x					x
Credit Card Fraud Detection using Machine Learning: A Study [16]	x	x	x	x	x	x	
Detection of Credit Card Fraud in E-Commerce Using Data Mining [18]		x				x	x
Approaches to Fraud detection on credit card transactions using artificial intelligence methods [19]			x	x	x	x	x

Tablo 3.3. Literatürde en fazla değinilen makine öğrenmesi algoritmaları

4. SONUÇ

Her projede olduğu gibi yapay zeka temelli dolandırıcılık tespiti çalışmasının da kendine has belirli noktaları ve veri dengesizliği, boyut indirgemesi, performans metrikleri gibi kaçırılmaması gereken önemli detayları bulunmaktadır. Projemizin bu dönem yapılan çalışmasında, bahsedilen detayların göz önünde bulundurulması ve model eğitimi aşamasından önce bu doğrultuda hareket edilmesi esas alınmıştır. Bu bağlamda son haline

getirilen veri setlerinin de önceki aşamalarda seçilen makine öğrenmesi algoritmaları ile eğitilmesi amaçlanmaktadır.

5. KAYNAKÇA

[1] F. Kaya, "Türkiye'de Kredi Kartı Uygulaması," 2009.

[2] V. Van Vlasselaer *et al.*, "APATE: A novel approach for automated credit card transaction fraud detection using network-based extensions," *Decision Support Systems*, vol. 75, pp. 38–48, Jul. 2015, doi: 10.1016/j.dss.2015.04.013.

[3] D. Olszewski, "Fraud detection using self-organizing map visualizing the user profiles," *Knowledge-Based Systems*, vol. 70, pp. 324–334, Nov. 2014, doi: 10.1016/j.knosys.2014.07.008.

[4] T. Amarasinghe, A. Aponso, and N. Krishnarajah, "Critical Analysis of Machine Learning Based Approaches for Fraud Detection in Financial Transactions," *Proceedings of the 2018 International Conference on Machine Learning Technologies - ICMLT '18*, 2018, doi: 10.1145/3231884.3231894.

[5] Y. Sahin, S. Bulkan, and E. Duman, "A cost-sensitive decision tree approach for fraud detection," *Expert Systems with Applications*, vol. 40, no. 15, pp. 5916–5923, Nov. 2013, doi: 10.1016/j.eswa.2013.05.021.

[6] Y. Abakarim, M. Lahby, and A. Attioui, "An Efficient Real Time Model For Credit Card Fraud Detection Based On Deep Learning," vol. 7, 2018.

[7] K. Pandey, P. Sachan, and N. G. Ganpatrao, "A Review of Credit Card Fraud Detection Techniques," *2021 5th International Conference on Computing Methodologies and Communication (ICCMC)*, Apr. 2021, doi: 10.1109/iccmc51019.2021.9418024.

[8] M. Kavitha and M. Suriakala, "Real Time Credit Card Fraud Detection on Huge Imbalanced Data using Meta-Classifiers," 2017.

[9] Thennakoon, Anuruddha, et al. "Real-time credit card fraud detection using machine learning", 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence). IEEE, 2019.

[10] S. Bhattacharyya, S. Jha, K. Tharakunnel, and J. C. Westland, "Data mining for credit card fraud: A comparative study," *Decision Support Systems*, vol. 50, no. 3, pp. 602–613, Feb. 2011, doi: 10.1016/j.dss.2010.08.008.

[11] M. F. Keskenler, D. Dal, and T. Aydın, "Yapay Zeka Destekli ÇOKS Yöntemi ile Kredi Kartı Sahtekarlığının Tespiti," *El-Cezeri Fen ve Mühendislik Dergisi*, May 2021, doi: 10.31202/ecjse.908260.

[12] V. N. Dornadula and S. Geetha, "Credit Card Fraud Detection using Machine Learning Algorithms," *Procedia Computer Science*, vol. 165, pp. 631–641, 2019, doi: 10.1016/j.procs.2020.01.057.

[13] P. Jiang, J. Zhang, and J. Zou, "Credit Card Fraud Detection Using Autoencoder Neural Network," 2019.

[14] S. Dhawan, S. Charan, R. Gangireddy, S. Kumar, and T. Chakraborty, “Spotting Collective Behaviour of Online Frauds in Customer Reviews,” 2019.

[15] A. YILMAZ and M. SELİMOGLU, “Kredi Kartı Dolandırıcılık Tespitinin Makine Öğrenmesi Yöntemleri ile Tahmin Edilmesi,” *Beykent Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, Feb. 2021, doi: 10.20854/bujse.873804.

[16] P. Tiwari, S. Mehta, N. Sakhuja, J. Kumar, and A. Singh, “Credit Card Fraud Detection Using Machine Learning: A Study Technical Report,” 2021.

[17] U. Porwal and S. Mukund, “Credit Card Fraud Detection in e-Commerce: An Outlier Detection Approach,” May 2019.

[18] Y Kırelli, S Arslankaya, and M Zeren, “Detection of Credit Card Fraud in E-Commerce Using Data Mining.,” *Avrupa Bilim ve Teknoloji Dergisi*, pp. 522–529, Oct. 2020.

[19] Kaggle. URL: https://www.kaggle.com (Son Erişim Tarihi: 29.12.2022).

[20] Google Dataset Search. URL: https://datasetsearch.research.google.com/ (Son Erişim Tarihi: 29.12.2022).

[21] UCI Machine Learning Repository. URL: https://archive.ics.uci.edu/ml/index.php (Son Erişim Tarihi: 29.12.2022).

[22] OpenML. URL: https://www.openml.org/ (Son Erişim Tarihi: 29.12.2022).

[23] DataHub. URL: https://datahub.io/ (Son Erişim Tarihi: 29.12.2022).

[24] Papers With Code. URL: https://paperswithcode.com/ (Son Erişim Tarihi: 29.12.2022).

[25] The Publications Office of the European Union. URL: https://data.europa.eu/en (Son Erişim Tarihi: 29.12.2022).

[26]Awesome Public Datasets URL: https://github.com/awesomedata/awesome-public-datasets (Son Erişim Tarihi: 29.12.2022).

[27]Credit Card Fraud Detection URL: https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud (Son Erişim Tarihi: 29.12.2022).

[28] A. Dal Pozzolo, O. Caelen, Reid A. J. and G. Bontempi, “Calibrating Probability with Undersampling for Unbalanced Classification.”, In Symposium on Computational Intelligence and Data Mining (CIDM), IEEE, 2015.

[29] A. Dal Pozzolo, A. Boracchi, O. Caelen, L. Borgne, Yann-Ael, S. Waterschoot, G. Bontempi, “Learned lessons in credit card fraud detection from a practitioner perspective”, Expert systems with applications,41,10,4915-4928, 2014.

[30] A. Dal Pozzolo, A. Boracchi, Giacomo, O. Caelen, C. Alippi, G. Bontempi, “Credit card fraud detection: a realistic modeling and a novel learning strategy”, IEEE transactions on neural networks and learning systems,29,8,3784-3797,IEEE, 2018.

[31] A. Dal Pozzolo, “Adaptive Machine learning for credit card fraud detection” ULB MLG PhD thesis (supervised by G. Bontempi), 2015.

[32] F. Carcillo, A. Dal Pozzolo, Y. Le Borgne, O. Caelen, Y. Mazzer, G. Bontempi, "Scarff: a scalable framework for streaming credit card fraud detection with Spark," Information fusion,41, 182-194, Elsevier, 2018.

[33] F. Carcillo, Y. Le Borgne, O. Caelen, G. Bontempi, "Streaming active learning strategies for real-life credit card fraud detection: assessment and visualization," International Journal of Data Science and Analytics, 5,4,285-300, Springer International Publishing, 2018.

[34] B. Lebichot, Y. Le Borgne, Liyun He, F. Oblé, G. Bontempi "Deep-Learning Domain Adaptation Techniques for Credit Cards Fraud Detection," INNSBDDL 2019: Recent Advances in Big Data and Deep Learning, pp 78-88, 2019.

[35] Fabrizio Carcillo, Y. Le Borgne, O. Caelen, F. Oblé, G. Bontempi, "Combining Unsupervised and Supervised Learning in Credit Card Fraud Detection Information Sciences," 2019.

[36] Y. Le Borgne, G. Bontempi, "Reproducible Machine Learning for Credit Card Fraud Detection - Practical Handbook," 2021.

[37] B. Lebichot, G. Paldino, W. Siblini, L. He, F. Oblé, G. Bontempi, "Incremental learning strategies for credit cards fraud detection," IInternational Journal of Data Science and Analytics, 2021.

[38] M. Al-Shabi, "Credit Card Fraud Detection Using Autoencoder Model in Unbalanced Datasets." Journal of Advances in Mathematics and Computer Science, 1-16, 10.9734/jamcs/2019/v33i530192, 2019.

[39] T. Sarkar, "XBNet: An extremely boosted neural network," *Intelligent Systems with Applications*, vol. 15, p. 200097, Sep. 2022, doi: 10.1016/j.iswa.2022.200097.

[40] G. Pang, C. Shen, and A. van den Hengel, "Deep Anomaly Detection with Deviation Networks," *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Jul. 2019, doi: 10.1145/3292500.3330871.

[41] Y. Xu, H. Dong, M. Zhou, J. Xing, X. Li, and J. Yu, "Improved Isolation Forest Algorithm for Anomaly Test Data Detection," *Journal of Computer and Communications*, vol. 09, no. 08, pp. 48–60, 2021, doi: 10.4236/jcc.2021.98004.

[42] D. Nugent, "Privacy-Preserving Credit Card Fraud Detection using Homomorphic Encryption," 2022.

[43] M. Al-Shabi, "Credit Card Fraud Detection Using Autoencoder Model in Unbalanced Datasets Some of the authors of this publication are also working on these related projects: Difference Equation View project Optical MINs View project," 2019.

[44]Credit Card Transactions. URL: https://www.kaggle.com/datasets/ealtman2019/credit-card-transactions?select=User0_credit_card_transactions.csv (Son Erişim Tarihi: 29.12.2022).

[45] I. Padhi *et al.*, "Tabular Transformers For Modeling Multivariate Time Series," 2021.

[46]Sparkov Data Generation. URL: https://github.com/namebrandon/Sparkov_Data_Generation (Son Erişim Tarihi: 29.12.2022).

[47] E. Altman, “Synthesizing Credit Card Transactions,” 2019.

[48] P. Grover *et al.*, “FDB: Fraud Dataset Benchmark,” Aug. 2022.

[49]Amazon Fraud Dataset Benchmark. URL: https://github.com/amazon-science/fraud-dataset-benchmark (Son Erişim Tarihi: 29.12.2022).

[50]IEEE-CIS Fraud Detection. URL: https://www.kaggle.com/competitions/ieee-fraud-detection/data (Son Erişim Tarihi: 29.12.2022).

[51] Synthetic data from a financial payment system. URL: https://www.kaggle.com/datasets/ealaxi/banksim1/code (Son Erişim Tarihi: 29.12.2022).

[52] E. Alonso Lopez-Rojas, S. Axelsson, and E. Alonso, “BankSim: A Bank Payment Simulation for Fraud Detection Research CyberAIMs View project BigData@BTH -Scalable resource-efficient systems for big data analytics View project BANKSIM: A BANK PAYMENTS SIMULATOR FOR FRAUD DETECTION RESEARCH,” 2014.

[53] Lopez-Rojas, Edgar Alonso ; Axelsson, Stefan Banksim: A bank payments simulator for fraud detection research Inproceedings 26th European Modeling and Simulation Symposium, EMSS 2014, Bordeaux, France, pp. 144–152, Dime University of Genoa, 2014, ISBN: 9788897999324.

[54]Synthetic Financial Datasets For Fraud Detection. URL: https://www.kaggle.com/datasets/ealaxi/paysim1 (Son Erişim Tarihi: 29.12.2022).

[55] B. Stojanović *et al.*, “Follow the Trail: Machine Learning for Fraud Detection in Fintech Applications,” *Sensors*, vol. 21, no. 5, p. 1594, Feb. 2021, doi: 10.3390/s21051594.

[56] E. A. Lopez-Rojas , A. Elmir, and S. Axelsson, “PaySim: A financial mobile money simulator for fraud detection,” In: The 28th European Modeling and Simulation Symposium-EMSS, Cyprus, 2016.