Teknoloji Fakültesi

MARMARA ÜNİVERSİTESİ

TEKNOLOJİ FAKÜLTESİ

BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

Bitirme Projesi 1. Ara Raporu

FraudShield: Makine Öğrenmesi Tabanlı Finansal Dolandırıcılık Tespit Sistemi

---

PROJE YAZARI

Muhammed Yasin Özdemir

Berkay Zaim

DANIŞMAN

İstanbul, 2025

Teknoloji Fakültesi

MARMARA ÜNİVERSİTESİ

TEKNOLOJİ FAKÜLTESİ

BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

Bitirme Projesi 1. Ara Raporu

FraudShield: Makine Öğrenmesi Tabanlı Finansal Dolandırıcılık Tespit Sistemi

---

PROJE YAZARI

Muhammed Yasin Özdemir - 171421005

Berkay Zaim - 171421002

DANIŞMAN

Dr. Öğr. Üyesi EYÜP EMRE ÜLKÜ

İstanbul, 2025

MARMARA ÜNİVERSİTESİ
TEKNOLOJİ FAKÜLTESİ
BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

Marmara Üniversitesi Teknoloji Fakültesi Bilgisayar Mühendisliği Öğrencisi …… nın "……" başlıklı bitirme projesi çalışması, …./…./….. tarihinde sunulmuş ve jüri üyeleri tarafından başarılı bulunmuştur.

Jüri Üyeleri

Prof. Dr. Adı SOYADI (Danışman)
Marmara Üniversitesi ..... (İMZA) .....

Doç. Dr. Adı SOYADI (Üye)
Marmara Üniversitesi ..... (İMZA) .....

Dr. Öğr. Üyesi Adı SOYADI (Üye)
Marmara Üniversitesi ..... (İMZA) .....

İÇİNDEKİLER

Sayfa

SEMBOLLER/SYMBOLS

T **:** İşlem süresi (s)

R **:**Risk skoru

P **:**Pozitif tahmin sayısı (adet)

N **:**Negatif tahmin sayısı (adet)

TP **:**Doğru pozitif tahmin sayısı (adet)

FP **:**Yanlış pozitif tahmin sayısı (adet)

TN **:**Doğru negatif tahmin sayısı (adet)

FN **:**Yanlış negatif tahmin sayısı (adet)

α **:**Model eşik değeri

AUC **:**ROC eğrisi altındaki alan

μ **:**Ortalama işlem skoru

σ **:**Standart sapma

KISALTMALAR/ABBREVIATIONS

ML: Machine Learning (Makine Öğrenmesi)

PCA: Principal Component Analysis (Ana Bileşenler Analizi)

LightGBM: Light Gradient Boosting Machine

SPA: Single Page Application

API: Application Programming Interface

CI/CD: Continuous Integration / Continuous Deployment

ROC: Receiver Operating Characteristic

DB: Database

Redis: Remote Dictionary Server (Önbellekleme teknolojisi)

Git: Versiyon Kontrol Sistemi

ŞEKİL LİSTESİ

Sayfa

TABLO LİSTESİ

Sayfa

ÖZET

Dijital finansal işlemlerin yaygınlaşması, kullanıcıların hizmetlere daha hızlı erişmesini sağlarken, dolandırıcılık vakalarının da artmasına neden olmuştur. Özellikle kredi kartı üzerinden gerçekleştirilen işlemler, kötü niyetli aktörler tarafından hedef alınmakta ve geleneksel güvenlik yöntemleri bu tehditlere karşı yetersiz kalmaktadır. Bu nedenle, gerçek zamanlı, öğrenebilir ve esnek yapıda çalışan yeni nesil tespit sistemlerine ihtiyaç duyulmaktadır.

Bu çalışma kapsamında, kredi kartı işlemlerinde dolandırıcılık riskini belirlemeye yönelik bir model geliştirilmiştir. Model, istatistiksel örüntülerin tespiti ve sınıflandırılması aşamalarını birlikte ele alan bir yapıda tasarlanmıştır. İşlem verileri üzerinden elde edilen zaman temelli ve davranışsal özellikler modele entegre edilmiştir. Ayrıca, veri setindeki dengesizlik problemi için çeşitli örnekleme ve ağırlıklandırma yöntemleri uygulanmıştır.

Geliştirilen sistemde, farklı algoritmaların birlikte çalışabildiği modüler bir yapı oluşturulmuş ve bu yapı üzerinden sınıflandırma işlemleri gerçekleştirilmiştir. Tüm süreçler, uçtan uca veri işleme hattı içinde yapılandırılmış ve sonuçların tekrar üretilebilirliği gözetilmiştir. Çalışma, finansal sistemlerde karşılaşılan tehditlere karşı veri temelli, güncellenebilir ve bütünleşik bir yaklaşım sunmayı hedeflemektedir.

ABSTRACT

The widespread use of digital financial services has increased accessibility while simultaneously leading to a rise in fraudulent activities. In particular, credit card transactions have become primary targets for malicious actors, and traditional security methods have proven insufficient against evolving threats. As a result, there is a growing need for real-time, adaptive, and intelligent detection systems capable of identifying fraudulent behavior.

This study proposes a model designed to detect the risk of fraud in credit card transactions. The model is structured to combine pattern recognition and classification within an integrated framework. Time-based and behavioral features derived from transaction data are incorporated into the model to enhance its decision-making capability. Additionally, data imbalance is addressed through sampling techniques and class-based weighting strategies.

A modular system architecture has been developed to enable the integration of different algorithms working together within a unified pipeline. All processes, including data transformation, feature extraction, modeling, and evaluation, are organized in an end-to-end flow to ensure consistency and reproducibility. The proposed system aims to offer a data-driven, updatable, and holistic solution to the emerging challenges of financial fraud in digital environments.

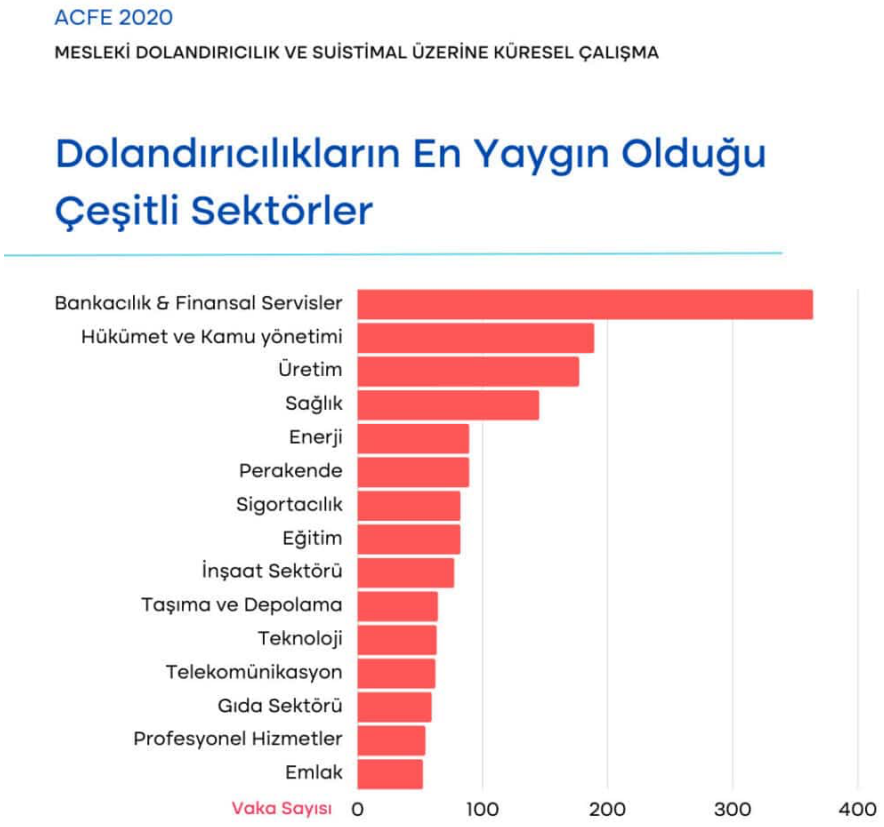
1. GİRİŞ

Dijital finansal teknolojilerdeki hızlı ilerleme, bireylerin ve kurumların finansal hizmetlere erişimini ciddi anlamda kolaylaştırmış, bununla birlikte finansal işlemlerin hacminde kayda değer bir artış meydana getirmiştir. Ancak bu gelişmeler, dolandırıcılık faaliyetlerinin de paralel şekilde çeşitlenip artmasına zemin hazırlamıştır. Kredi kartlarının yaygınlaşması, çevrim içi ödeme sistemlerinin kullanımı ve dijital bankacılığın benimsenmesiyle birlikte dolandırıcılık teknikleri daha karmaşık, organize ve tahmin edilmesi zor bir yapıya bürünmüştür. Literatürde dijital finansın finansal işlemler üzerindeki olumlu etkileri sıkça vurgulanmakla birlikte, bu dönüşümün beraberinde getirdiği yeni risk alanları da dikkat çekmektedir [1].

Finansal dolandırıcılık, yalnızca bireysel kullanıcıları değil; aynı zamanda işletmeleri ve finansal sistemin bütünlüğünü tehdit eden kritik bir sorundur. Bu tehditler, maddi zararların ötesine geçerek, güven kaybı, hukuki süreçlerde artış ve operasyonel risklerin yükselmesi gibi çok yönlü olumsuz sonuçlar doğurabilmektedir. Özellikle geleneksel güvenlik altyapılarına sahip kurumlar, modern saldırı tekniklerine karşı yeterli esnekliği ve tepki hızını gösterememektedir [2]. Bu nedenle, dolandırıcılıkla mücadelede geleneksel yöntemlerin ötesine geçilerek, veri temelli ve sürekli öğrenebilen yeni sistemlerin geliştirilmesi kaçınılmaz hâle gelmiştir.

Mevcut dolandırıcılık tespit sistemleri çoğunlukla sabit kurallara dayalı olarak çalışmakta ve yalnızca bilinen tehditleri tanımlamada etkili olmaktadır. Oysa günümüzde karşılaşılan dolandırıcılık türleri, çok daha dinamik ve kompleks bir yapı sergilemektedir. Bu durum, yanlış alarmların artmasına ve bazı tehditlerin gözden kaçmasına neden olmaktadır [3]. Bu sebeple, dolandırıcılığı daha isabetli şekilde tespit edebilecek yeni nesil yöntemlere duyulan ihtiyaç giderek artmaktadır.

Nitekim ACFE (2020) tarafından yayımlanan küresel rapora göre, dolandırıcılık vakalarının en sık görüldüğü sektörün açık ara bankacılık ve finansal hizmetler olduğu belirlenmiştir (Şekil 1.1). Bu durum, sektörel bazda risk düzeylerinin farklılaştığını ve özellikle dijitalleşme oranı yüksek alanlarda daha yoğun güvenlik önlemleri alınması gerektiğini ortaya koymaktadır.[14]



Şekil 1.1 Dolandırıcılığın en yaygın görüldüğü sektörler

Son dönemde öne çıkan makine öğrenmesi tabanlı yaklaşımlar, bu alandaki eksiklikleri giderme potansiyeli taşımaktadır. Bu yöntemler, geçmiş işlem verilerinden yola çıkarak sıra dışı davranış kalıplarını tanımlayabilmekte; bilinmeyen dolandırıcılık örüntülerini ortaya çıkararak, sürekli öğrenme yetenekleri sayesinde kendilerini geliştirebilmektedir. Denetimli ve denetimsiz öğrenme algoritmalarının bu alandaki başarıları, geleneksel sistemlere kıyasla daha güçlü bir çözüm sunduğunu göstermektedir [4][5].

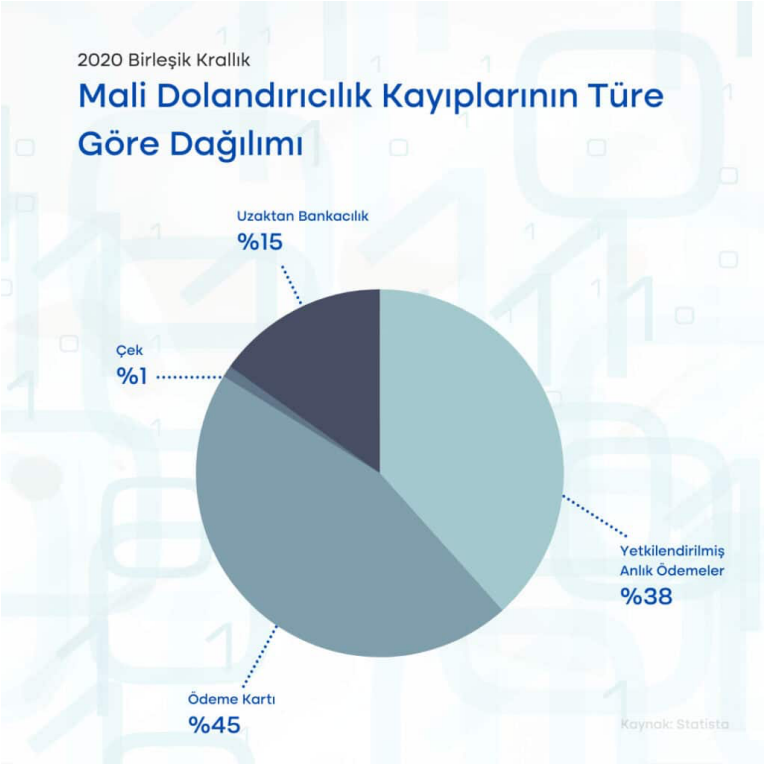
Bu çalışmada önerilen yaklaşım, finansal dolandırıcılıkla mücadelede veri odaklı ve öğrenebilir modellerin etkin kullanımını merkeze alarak, daha doğru ve etkili bir tespit sistemi oluşturmayı hedeflemektedir. Böylece günümüz dijital finansal ortamında karşılaşılan tehditlere yönelik yenilikçi ve sürdürülebilir çözümler geliştirilmesi amaçlanmaktadır. Çalışma, bireysel ve kurumsal tehditleri kapsayan çok yönlü bir dolandırıcılık perspektifinden hareketle, literatürde öne çıkan temel eksikliklere katkı sunmayı hedeflemektedir [6][7].

1.1. Proje Çalışmasının Amacı ve Önemi

Finansal dolandırıcılık vakalarının çeşitlenmesi ve karmaşıklaşması, mevcut güvenlik sistemlerinin yetersizliğini gün yüzüne çıkarmakta ve daha esnek, öğrenebilir ML ve gerçek zamanlı çözümlere olan ihtiyacı açık bir şekilde ortaya koymaktadır. Literatürde sıklıkla vurgulanan üç temel sorun, mevcut yaklaşımların bu alanda neden yetersiz kaldığını özetlemektedir: (i) gelişmiş analiz yöntemlerinin eksikliği, (ii) sadece belirli boyutlara odaklanan parçalı analiz yaklaşımları ve (iii) dolandırıcılık örüntülerinin evrimine karşı sistemlerin düşük adaptasyon kapasitesi [1][2].

Bu kapsamda yürütülen bu proje, finansal işlemlerdeki anormal davranışları tespit edebilen ve aynı zamanda dolandırıcılık vakalarını sınıflandırarak risk seviyelerini belirleyebilen bir model geliştirmeyi amaçlamaktadır. Önerilen yaklaşımda, PCA algoritması ile istatistiksel olarak olağandışı işlem örüntüleri belirlenirken, bu örüntüler LightGBM algoritması ile dolandırıcılık riski açısından sınıflandırılacaktır. Bu sayede yalnızca bilinen tehditlere değil, aynı zamanda daha önce tanımlanmamış şüpheli davranışlara da etkin şekilde yanıt verebilen bir yapının oluşturulması hedeflenmektedir [3][4].

Dolandırıcılığın türlere göre dağılımına bakıldığında ise, ödeme kartı ve yetkilendirilmiş anlık ödemelerin en fazla zarara yol açan türler olduğu görülmektedir (Şekil 1.2). Bu dağılım, sistemin neden ödeme davranışlarını merkeze alan bir yapı üzerine kurgulandığını ortaya koymaktadır [14].



Şekil 1.2. Mali dolandırıcılık kayıplarının türe göre dağılımı

Çalışmanın özgün yönü, literatürdeki çok sayıda parçalı çözümün ötesine geçerek, gerçek dünya uygulamalarına entegre edilebilir bütüncül bir analiz mimarisi sunmasında yatmaktadır. Önerilen sistem, yalnızca teknik doğruluk sağlamayı değil; aynı zamanda sürekli güncellenebilirlik ve yapısal esneklik yoluyla dijital dolandırıcılık ortamının değişken doğasına uzun vadeli uyum sağlayabilmeyi hedeflemektedir. Buna ek olarak, sistemin bireysel, organize ve senaryo bazlı dolandırıcılık türlerini ayrıştırarak öğrenebilmesi, çok boyutlu tehditler karşısında da etkili olmasını mümkün kılmaktadır [5].

Güncel araştırmalar, makine öğrenimi ML uygulamalarının finansal güvenlik alanında özellikle düşük yanlış alarm oranı, yüksek esneklik ve hibrit modelleme kapasitesiyle dikkat çektiğini göstermektedir [6]. Özellikle birden fazla algoritmanın birlikte kullanıldığı hibrit yaklaşımlar, dolandırıcılık tespitinde anlamlı performans artışlarına katkı sağlamaktadır. Bu çerçevede söz konusu proje, yalnızca mevcut sistemlerin ötesine geçmeyi değil, aynı zamanda hem akademik literatürdeki boşluklara katkı sunmayı hem de sektörel uygulamalara uygun bir çözüm modeli ortaya koymayı amaçlamaktadır.

Bu bağlamda çalışmanın temel araştırma soruları şu şekilde belirlenmiştir:

- Özellik temelli anomali tespiti ve sınıflandırma yaklaşımı, dolandırıcılığın erken ve isabetli tespitinde ne kadar etkilidir?

- PCA ve LightGBM algoritmalarının birleşimi, doğruluk ve işlem süresi (*T*) bakımından nasıl avantajlar sunmaktadır?
- Öğrenebilen bir sistem, gelişen dolandırıcılık tekniklerine karşı ne ölçüde uyum sağlayabilir?

Bu yönüyle proje yalnızca bir yazılım çözümü değil; aynı zamanda dijital finansal sistemlerin güvenliğine yönelik sürdürülebilir, veri odaklı ve yapay zekâ temelli bir katkı sunmayı hedeflemektedir.

1.2. Literatür Özeti

Finansal dolandırıcılık tespitine yönelik literatür, dijitalleşme sürecinin hızlanmasıyla birlikte yeni tehdit türlerinin ortaya çıktığını ve geleneksel güvenlik mekanizmalarının bu tehditleri önlemede yetersiz kaldığını ortaya koymaktadır. Bu bağlamda, çok sayıda araştırma yeni nesil veri analitiği ve yapay zekâ destekli yaklaşımlara yönelmiştir.

Wang ve arkadaşları (2023), dijital finansın kurumlar üzerindeki etkilerini inceledikleri çalışmalarında, dijitalleşmenin finansman erişimini artırmakla birlikte, yeni dolandırıcılık risklerini de beraberinde getirdiğini belirtmişlerdir [1]. Özellikle kart dolandırıcılığına yönelik araştırmalar, teknolojik uyum eksikliklerinin saldırganlar için açık kapı oluşturduğunu ve sistemlerin sürekli güncellenebilir bir yapıya sahip olması gerektiğini vurgulamaktadır [2].

Ryman-Tubb ve çalışma arkadaşları (2021), geleneksel kural tabanlı dolandırıcılık tespit sistemlerinin, yüksek yanlış alarm üretme eğiliminde olduğunu ve bilinmeyen tehditleri saptamakta yetersiz kaldığını ortaya koymuşlardır [3]. Bu sınırlamaları aşmak amacıyla önerilen yeni yöntemler, makine öğrenimi ML tabanlı yaklaşımları merkeze almaktadır. Bello (2023), bu tekniklerin gelişmiş analiz kapasitesi sunduğunu ve gerçek zamanlı tespit sistemlerinin başarımını artırabileceğini ifade etmektedir [4].

Yapılan çalışmalar, Principal Component Analysis PCA gibi boyut indirgeme yöntemlerinin anomali tespitinde; Light Gradient Boosting Machine LightGBM gibi ağaç tabanlı algoritmaların ise sınıflandırma süreçlerinde yüksek doğruluk sağladığını göstermektedir [5]. Sun ve arkadaşlarının (2023) çalışması, AutoEncoder ile düşük boyutlu özellik çıkarımı sonrasında LightGBM algoritmasının sınıflandırma başarımını artırdığını ortaya koymuştur [6].

Liu ve arkadaşları tarafından geliştirilen CoDetect modeli, hem ağ yapısını hem de varlık özelliklerini birlikte değerlendirerek çift yönlü bir analiz sunmakta ve dolandırıcılık tespitinde bütüncül bir yaklaşım önermektedir. Her ne kadar bu çalışma doğrudan ağ analizi yapmıyor olsa da, CoDetect'in "özellik temelli anomali tespiti" yaklaşımı, önerilen sistem ile benzerlik göstermektedir [7].

Bununla birlikte, yakın dönemli bir IEEE çalışması, farklı algoritmaların birlikte kullanıldığı hibrit modellerin, özellikle sigorta sektöründe dolandırıcılığı tespit etmede daha başarılı sonuçlar verdiğini ortaya koymuştur. Bu durum, çalışmada önerilen PCA + LightGBM kombinasyonunun literatürde karşılık bulduğunu desteklemektedir [8].

Ayrıca KPMG Türkiye tarafından yayımlanan 2024 tarihli raporda, üretken yapay zekâ uygulamalarının finansal işlemlerdeki riski azaltma ve dolandırıcılık eğilimlerini önceden tahmin etme amacıyla giderek daha fazla kullanıldığı belirtilmiştir. Bu gelişme, makine öğrenimi tabanlı modellerin yalnızca operasyonel değil; aynı zamanda yönetsel karar süreçlerine de stratejik katkı sunduğunu göstermektedir [9].

Sonuç olarak literatür, geleneksel sistemlerin günümüz tehditlerine karşı yetersiz kaldığını; veri temelli, öğrenebilir ve yorumlanabilir modellerin dolandırıcılık tespitinde yeni bir standart oluşturduğunu ortaya koymaktadır. Bu çalışma, literatürde belirtilen açıkları kapatmayı amaçlayan; gerçek zamanlı çalışan, öğrenebilen ve anomali örüntülerini etkin biçimde analiz edebilen bir yapı sunarak alana özgün katkı sağlamayı hedeflemektedir.

2. MATERYAL VE YÖNTEM

Günümüzde finansal dolandırıcılığın karmaşıklığı, geleneksel güvenlik yaklaşımlarını yetersiz bırakmakta ve daha dinamik, veri odaklı çözümlere olan ihtiyacı artırmaktadır. Bu bağlamda geliştirilen FraudShield platformu, farklı veri kaynaklarını bir araya getiren, çok katmanlı analiz yöntemleri kullanan ve gerçek zamanlı karar mekanizmalarıyla donatılmış bir yapı sunmaktadır. Sistem, makine öğrenimi algoritmaları, grafik analiz teknikleri ve çoklu veritabanı mimarileriyle entegre çalışarak bireysel ve organize dolandırıcılık örüntülerini yüksek doğrulukla tespit etmeyi amaçlamaktadır.

Bu bölümde, FraudShield'in araştırma yapısı, kullanılan yöntemler, analiz teknikleri ve bu yöntemlerin proje hedeflerine nasıl hizmet ettiği detaylı olarak sunulmaktadır. Ayrıca, yöntemin literatürdeki karşılığı ve neden bu tasarımın tercih edildiği ilgili çalışmalar ışığında açıklanmıştır.

2.1. Araştırma Tasarımı

Bu çalışma, finansal işlem verileri üzerinde gerçekleşen dolandırıcılıkları erken ve doğru şekilde tespit edebilecek bir analiz yapısının geliştirilmesini amaçlamaktadır. FraudShield olarak adlandırılan bu sistem; veri odaklı, öğrenebilir ve gerçek zamanlı işleyebilen bir model sunmayı hedeflemektedir.

Araştırmanın temel amacı, işlem verilerinden elde edilen örüntüler aracılığıyla olağandışı davranışları tanımlayabilen bir yapı oluşturmaktır. Geliştirilen yapı, işlem davranışlarındaki sapmaları belirleyerek dolandırıcılık riski taşıyan durumları sınıflandırmaya odaklanmaktadır.

Bu kapsamda:

- **Bağımlı değişken**, sistemin dolandırıcılık vakalarını doğru şekilde tespit edebilme başarımıdır. Bu başarı, doğruluk oranı, hata oranı ve F1-Skoru gibi değerlendirme ölçütleriyle analiz edilecektir.
- **Bağımsız değişkenler**, işlem miktarı, işlem sıklığı, zaman bilgisi gibi kullanıcıya ve işleme özgü parametreleri kapsamaktadır.

Araştırma tasarımı iki temel analiz aşamasından oluşmaktadır:

2.1.1. Anomali Tespiti

İlk aşamada, işlem verilerindeki olağandışı davranışların tespiti için **PCA** yöntemi uygulanmaktadır. Bu yöntem, veri setindeki örüntüleri boyut indirgeme yoluyla sadeleştirerek aykırı durumların ön analizini sağlamaktadır. PCA sayesinde sistem, dolandırıcılık riski barındıran işlemlerin öncelikle anomali olarak değerlendirilmesini mümkün kılar.

2.1.2. Sınıflandırma

İkinci aşamada, işlem verilerinin dolandırıcılık içerip içermediğini belirlemek amacıyla **LightGBM** algoritması kullanılmaktadır. Bu algoritma, geçmiş verilerden öğrenme yoluyla yeni işlemleri değerlendirmekte ve sınıflandırma kararı vermektedir. Özellikle karar ağaçlarına dayalı yapısı sayesinde sınıflandırma doğruluğunu artırmakta ve işlem süresi açısından etkin çözümler sunmaktadır.

Proje süreci; öncelikle veri altyapısının oluşturulması, ardından analiz bileşenlerinin geliştirilmesi ve son aşamada modelin test edilmesini içeren aşamalı bir yapıda

ilerlemektedir. Sistem, gerçek işlem verileriyle denenecek; elde edilen performans çıktıları doğrultusunda model iyileştirmeleri gerçekleştirilecektir.

Bu çalışmada, finansal işlemler üzerinde meydana gelebilecek dolandırıcılık eylemlerini tespit edebilen, esnek, veri odaklı ve öğrenebilir bir yapay zekâ sisteminin geliştirilmesi amaçlanmaktadır. Uygulanan yöntem ve teknikler, veri hazırlık sürecinden model değerlendirmesine kadar yapılandırılmış bir iş akışı içerisinde ele alınmıştır. Aşağıda, sistemin temel bileşenleri sistematik biçimde açıklanmaktadır.

2.2. Kullanılan Yöntem ve Teknikler

2.2.1. Veri Seti ve Hazırlık Süreci

Bu çalışmada kullanılan veri, **Université Libre de Bruxelles (ULB)** tarafından geliştirilen ve **Kaggle** platformunda yayımlanan açık erişimli **"Credit Card Fraud Detection"** veri setine dayanmaktadır [10]. Veri seti, 284.807 adet Avrupa merkezli kredi kartı işleminden oluşmakta olup, yalnızca 492 işlem dolandırıcılık vakası (etiket: 1) olarak işaretlenmiştir. Bu da verideki dolandırıcılık oranının yaklaşık **%0.172** olduğu anlamına gelmektedir. Bu tür veri setleri, ciddi bir **sınıf dengesizliği** (class imbalance) sorunu barındırır ve bu durum, denetimli öğrenme yöntemlerinin başarımını olumsuz yönde etkileyebilir.

Veri setinde yer alan değişkenler, gizlilik gereği anonimleştirilmiş olup, çoğu **Principal Component Analysis (PCA)** yöntemiyle boyut indirgemeye tabi tutulmuştur. Orijinal olarak yalnızca üç değişken (*Time*, *Amount* ve *Class*) ham hâlde bırakılmış, geri kalan 28 değişken ise **V1–V28** biçiminde ifade edilmiştir. Bu PCA dönüşümü, özellikle yüksek korelasyon içeren finansal verilerde gürültüyü azaltarak öğrenme sürecinin verimliliğini artırmak amacıyla tercih edilmiştir.

Bu veri seti, literatürde birçok dolandırıcılık tespit çalışmasına konu olmuş ve farklı makine öğrenmesi algoritmalarının değerlendirilmesinde yaygın olarak referans alınmıştır. Örneğin Bahnsen et al. (2016), bu veri seti üzerinde farklı sınıflandırma algoritmalarını karşılaştırmış ve sınıf dengesizliğinin model başarımı üzerinde doğrudan etkili olduğunu vurgulamıştır [11].

2.2.1.1. Veri Hazırlık Süreci ve Uygulanan Teknikler

Modelin başarıyla eğitilebilmesi için veri seti üzerinde aşağıdaki ön işleme adımları gerçekleştirilmiştir:

- **Zaman bazlı özellik mühendisliği:**

  Time değişkeni, trigonometrik dönüşümler yardımıyla *TimeSin* ve *TimeCos* gibi sürekli değişkenlere dönüştürülmüştür. Ayrıca, işlem gününe ve saatine göre *DayFeature* ve *HourFeature* gibi kategorik değişkenler oluşturulmuştur. Bu dönüşümler, işlem davranışlarının daha detaylı analiz edilmesine olanak tanımaktadır.

- **Sınıf dengesizliği çözümü:**

  Veri setindeki dengesiz yapı nedeniyle **random undersampling** yöntemi uygulanmıştır. Bu yöntem, çoğunluk sınıfındaki örneklerin bir kısmını kaldırarak, dolandırıcılık sınıfının model tarafından daha iyi öğrenilmesini sağlamaktadır.

- **Veri bölme:**

  Eğitim ve test kümeleri, sınıflar arası oranların korunmasını sağlayan **stratified sampling** yöntemiyle %70 eğitim ve %30 test olacak şekilde ayrılmıştır. Bu yöntem, özellikle dengesiz veri setlerinde test sonuçlarının güvenilirliğini artırmak için tercih edilmektedir.

2.2.2. Anomali Tespiti – PCA

Kredi kartı işlemlerinin büyük çoğunluğu yasal olsa da, az sayıda gerçekleşen dolandırıcılık vakaları genellikle istatistiksel olarak aykırı örüntüler sergilemektedir. Bu nedenle, veri seti içerisinde olağandışı işlem davranışlarının ön tespiti için **PCA** yöntemi kullanılmıştır [12].

PCA, yüksek boyutlu verileri daha az sayıda ana bileşenle temsil ederek, veri içerisindeki temel varyansları korur ve bilgi kaybını minimize eder. Aynı zamanda, aykırı veri noktalarının (anomalilerin) daha net biçimde ayrışmasına imkân tanır. Bu yöntem, dolandırıcılık gibi nadir olayların tespitinde ön eleme işlevi görerek sınıflandırma algoritmalarının başarımını artırmak amacıyla kullanılmıştır.

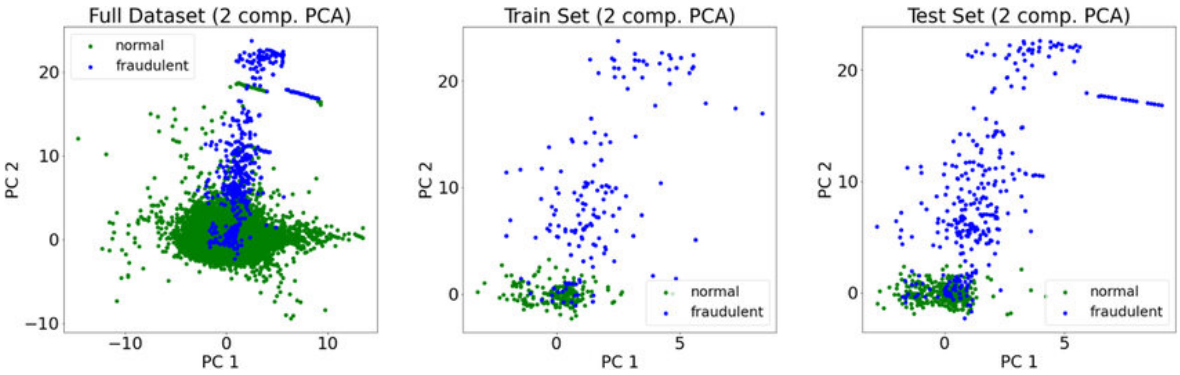
Yapılan güncel araştırmalar, PCA'nın özellikle dengesiz veri yapısına sahip finansal işlem setlerinde başarılı ön işlem adımı olarak kullanılabildiğini ve **anormal işlem örüntülerini daha görünür hâle getirdiğini** göstermektedir.

Bu çalışmada PCA aşağıdaki hedeflerle kullanılmıştır:

- **Amacı:** Aykırı işlem örüntülerinin belirlenmesi ve sınıflandırma öncesi ön eleme sağlanması
- **Doğruluk hedefi:** ≥ %60
- **Yanlış pozitif oranı (FP):** ≤ %20

PCA çıktıları, sınıflandırma modeline girdi sağlamak yerine, dolandırıcılık ihtimali taşıyan işlemleri önceden işaretleyerek modelin karar mekanizmasına yön vermiştir. Böylece sistemin genel doğruluğu korunurken, yanlış alarm oranları da düşürülmeye çalışılmıştır.

Aşağıda PCA ile elde edilen iki boyutlu bileşenler üzerinden normal ve dolandırıcılık işlemlerinin dağılımı örneklenmiştir (Şekil 2.2.2)[15].



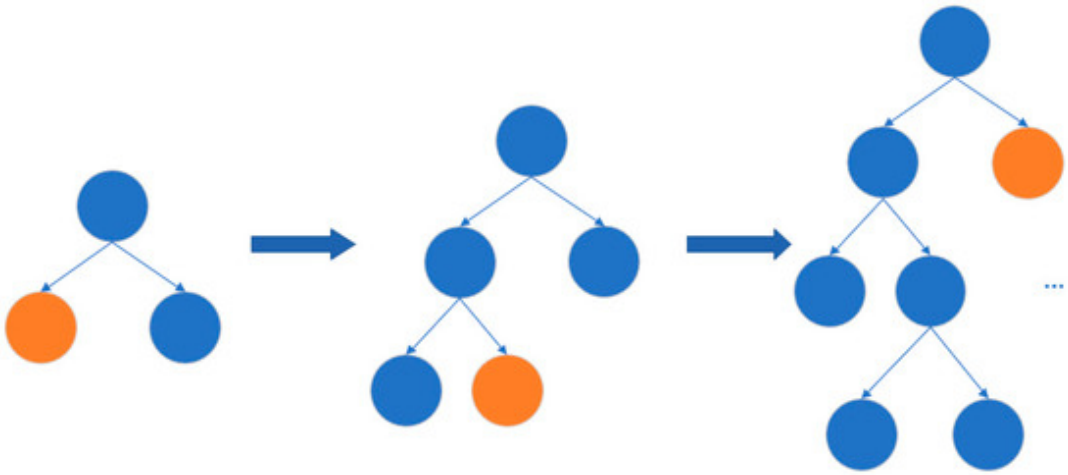
Şekil 2.2.2 PCA ile normal ve dolandırıcılık işlemlerinin ayrışımı

2.2.3. Risk Sınıflandırması – LightGBM

PCA ile ön değerlendirmeden geçen işlem verileri, dolandırıcılık riski açısından sınıflandırılmak üzere **Light Gradient Boosting Machine (LightGBM)** algoritması ile analiz edilmiştir. LightGBM, özellikle yüksek boyutlu ve sınıf dengesizliği içeren veri setlerinde hızlı ve etkili tahminler sunabilen bir **gradient boosting** yöntemidir. Ağaç tabanlı yapısı sayesinde karmaşık örüntüleri yüksek doğrulukla öğrenebilmekte ve düşük işlem süresiyle öne çıkmaktadır [13].

Güncel çalışmalar, LightGBM'in dolandırıcılık tespitinde geleneksel sınıflandırıcılara kıyasla daha iyi **precision–recall dengesi**, daha düşük **hata oranı** ve daha

yüksek **genel doğruluk** sağladığını ortaya koymaktadır. Aşağıda, LightGBM algoritmasının model oluşturma sürecine ilişkin genel yapısı görselleştirilmiştir (Şekil 2.2.3)[16]. Bu çalışmada LightGBM algoritması aşağıdaki teknik adımlarla yapılandırılmıştır:



Şekil 2.2.3. LightGBM algoritmasının karar ağacı evrimi

- **Hiperparametre optimizasyonu:**

  Modelin başarımını artırmak amacıyla **Grid Search** yöntemiyle çeşitli parametre kombinasyonları test edilmiştir. Bu kapsamda özellikle aşağıdaki parametreler optimize edilmiştir:

  - NumberOfLeaves = {64, 128, 256}
  - LearningRate = {0.005, 0.01}
  - NumberOfTrees = {500, 1000}

- **Sınıf ağırlıklandırma:**

  Dengesiz veri yapısına karşı önlem olarak, dolandırıcılık sınıfına (etiket: 1) **250.0 ağırlık** verilmiştir. Bu sayede modelin azınlık sınıfa (dolandırıcılık) duyarlılığı artırılmış, hatalı negatif sınıflandırmaların (FN) azaltılması hedeflenmiştir.

- **Özellik önemi analizi:**

  Eğitilen modelde kullanılan değişkenlerin **göreli önem değerleri (feature importance)** hesaplanmış ve dolandırıcılığı en iyi öngören faktörler belirlenmiştir. Bu analiz, modelin yorumlanabilirliğini artırmak ve açıklayıcı veri unsurlarını ön plana çıkarmak açısından önemli bir rol oynamaktadır.

LightGBM'in bu şekilde yapılandırılması, sistemin hem öğrenme başarısını hem de tahmin güvenilirliğini artırarak, dolandırıcılık vakalarının daha doğru ve zamanında tespit edilmesine katkı sağlamıştır.

2.2.4. Ensemble Modelleme Stratejisi

Model başarımını artırmak ve farklı algoritmaların güçlü yönlerini birleştirmek amacıyla bu çalışmada **ensemble modelleme** yaklaşımı benimsenmiştir. Özellikle **Principal Component Analysis (PCA)** ve **Light Gradient Boosting Machine (LightGBM)** algoritmalarının birlikte kullanımı ile farklı veri perspektiflerinden elde edilen çıktılar entegre edilerek daha güvenilir sınıflandırmalar yapılması hedeflenmiştir.

Bu bağlamda uygulanan adımlar aşağıda özetlenmiştir:

- **Ağırlıklı ortalama yöntemi:**

  İki algoritmadan elde edilen tahmin sonuçları, belirlenen ağırlık katsayıları ile birleştirilmiştir. Nihai risk skoru, **LightGBM tahminlerine %70**, **PCA temelli değerlendirmelere %30** ağırlık verilerek hesaplanmıştır. Bu yöntem, hem doğruluk hem de genellenebilirlik açısından dengeli bir çıktı üretmeyi amaçlamaktadır.

- **Güven eşiği (Confidence Threshold):**

  Sınıflandırma kararlarının daha isabetli olmasını sağlamak amacıyla, yalnızca **%80'in üzerinde güven skoru** elde edilen örnekler pozitif (dolandırıcılık riski taşıyan) olarak etiketlenmiştir.

- **Çapraz doğrulama:**

  Modelin istatistiksel sağlamlığı ve genellenebilirliğini test etmek amacıyla **5 katlı çapraz doğrulama (5-fold cross-validation)** uygulanmıştır. Bu sayede modelin farklı alt kümelerdeki performansı karşılaştırmalı olarak değerlendirilmiştir.

2.2.5. Normalizasyon ve İşlem Hattı (Pipeline) Yapısı

Modelleme sürecinin tutarlı, tekrar edilebilir ve sürdürülebilir olması için veri dönüşüm işlemleri ve model eğitimi, uçtan uca bir **işlem hattı (pipeline)** içerisinde yapılandırılmıştır. Bu yapı, hem veri ön işleme hem de tahmin sürecinde standartlaşmayı mümkün kılmaktadır.

- **Normalizasyon teknikleri:**
  - *V1–V28* değişkenleri için: **Z-score (mean-variance) normalizasyonu** uygulanmıştır.

- *Amount* değişkeni için: **Min-max normalizasyonu** tercih edilmiştir.

- **Pipeline mimarisi:**

  Özellik mühendisliği, normalizasyon, model eğitimi ve değerlendirme adımları, birbirine bağlı ancak modüler bir yapı içinde kurgulanmıştır. Bu yaklaşım sayesinde sistemin güncellenebilirliği ve test edilebilirliği artırılmıştır.

- **Model saklama ve tekrar kullanım:**

  Normalizasyon parametreleri, "CreditCard_Normalizer.zip" adlı dosyada kayıt altına alınmış ve bu sayede tahmin sürecinde dönüşümlerin tutarlılığı garanti altına alınmıştır. Bu uygulama, üretim ortamında modelin tutarlı ve güvenilir sonuçlar üretmesini desteklemektedir.

2.2.6. Performans Metrikleri ve Değerlendirme

Modelin etkinliği, çok yönlü performans metrikleriyle değerlendirilmiştir. Değerlendirme sürecinde hem genel başarım hem de sınıf bazlı duyarlılık göz önünde bulundurulmuştur:

- **Confusion matrix:**

  Modelin tahminleri, *True Positive (TP)*, *False Positive (FP)*, *True Negative (TN)* ve *False Negative (FN)* bileşenlerine ayrılarak değerlendirilmiştir. Bu metrikler üzerinden **doğruluk (accuracy)**, **duyarlılık (recall)** ve **özgüllük (specificity)** hesaplanmıştır.

- **ROC eğrisi ve AUC (Area Under Curve):**

  Modelin farklı eşik değerleri altındaki performansı, **ROC eğrisi** yardımıyla analiz edilmiş ve toplam başarıyı ölçen **AUC skoru** hesaplanmıştır. Bu analiz, sınıf dengesizliği olan veri setlerinde model performansını daha kapsamlı değerlendirmek için kritik öneme sahiptir.

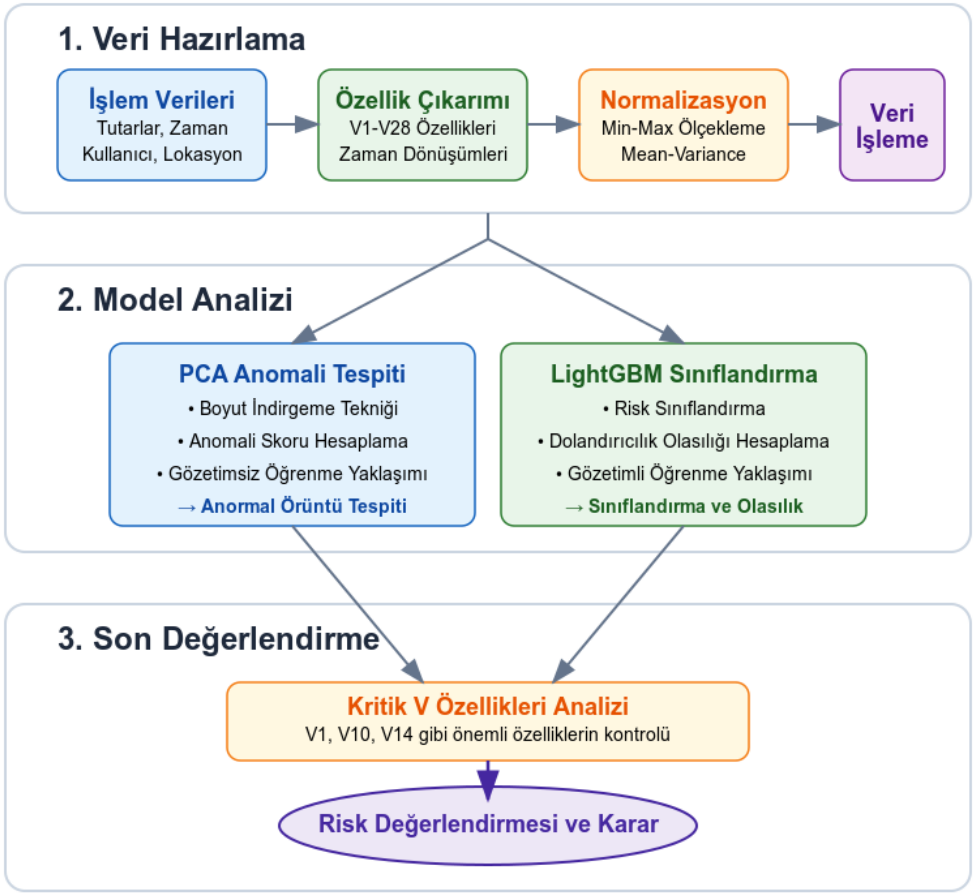
- **Model versiyon yönetimi:**

  Eğitim ve test süreçlerinde elde edilen farklı model versiyonları arasında karşılaştırmalar yapılmış; en yüksek başarıyı gösteren model üretim ortamına aktarılmıştır. Bu yaklaşım, sistemin güncel ve optimize edilmiş modellerle çalışmasını sağlamaktadır.

2.3. Analiz Teknikleri

Bu bölümde, geliştirilen modelin oluşturulmasında izlenen analiz süreçleri sistematik bir biçimde sunulmaktadır. Özellikle veri dengesizliği, işlem zamanına bağlı davranışsal örüntüler ve modelin genel başarımı gibi kritik zorluklara yönelik çözümler dikkate alınarak, analiz adımları yapılandırılmıştır. Süreç; özellik çıkarımı, model eğitimi ve test süreci, değerlendirme metrikleri, hiperparametre optimizasyonu ve eşik analizlerini kapsamaktadır.

Modelin genel analiz süreci Şekil 2.3.1'de özetlenmiş olup, veri hazırlama aşamasından son değerlendirmeye kadar olan adımlar bütüncül bir bakışla gösterilmiştir.



Şekil 2.3.1. Önerilen sistemin analiz süreci akış diyagramı

2.3.1. Özellik Çıkarımı (Feature Extraction)

Modelin doğruluğu ve genelleme yeteneği, büyük ölçüde veri setinden elde edilen özelliklerin temsil gücüne bağlıdır. Bu bağlamda, kapsamlı bir özellik mühendisliği süreci yürütülmüştür:

- **Zaman özelliklerinin dönüşümü:**

  İşlem zamanı bilgileri trigonometrik dönüşümler ile yeniden yapılandırılarak *TimeSin* ve *TimeCos* değişkenleri türetilmiştir. Ayrıca haftanın günü ve işlem saati bazında *DayFeature* ve *HourFeature* gibi kategorik değişkenler oluşturulmuştur.

- **İşlem tutarı dönüşümleri:**

  Amount değişkeni için hem **min-max normalizasyonu** (*Amount_normalized*) hem de **logaritmik dönüşüm** (*LogAmount*) uygulanmıştır.

- **PCA bileşenlerinin standardizasyonu:**

- Anonimleştirilmiş *V1–V28* değişkenleri, **z-score (mean-variance)** normalizasyonu ile ölçeklendirilmiştir. Böylece modelin tüm bileşenlere eşit duyarlılıkla yaklaşması sağlanmıştır.

- **Özellik birleştirme:**

  Elde edilen tüm dönüştürülmüş ve normalize edilmiş özellikler **concatenate** edilerek tek bir *Features* vektörü hâlinde model girişine sunulmuştur.

2.3.2. Model Eğitimi ve Test Süreci

Modelin güvenilir şekilde eğitilebilmesi ve test edilebilmesi için veri dağılımı ve sınıf dengesizliği gibi temel problemler göz önünde bulundurulmuştur:

- **Katmanlı örnekleme (Stratified Sampling):**

  Eğitim ve test veri kümeleri, sınıf oranlarını koruyacak şekilde ayrılmıştır.

- **Random undersampling:**

  Çoğunluk sınıfındaki (normal işlemler) veri sayısı azaltılarak, modelin dolandırıcılık sınıfını (azınlık sınıf) daha iyi öğrenmesi hedeflenmiştir.

- **5-katlı çapraz doğrulama:**

  Modelin genelleme yeteneğini test etmek ve aşırı öğrenmeyi önlemek için 5-fold **cross-validation** yöntemi kullanılmıştır.

2.3.3. Model Performans Ölçütleri

Modelin başarımı, hem genel doğruluk hem de pozitif/negatif sınıflar için ayrı ayrı hesaplanan istatistiksel ölçütlerle değerlendirilmiştir:

- **Confusion matrix:**

  Modelin çıktıları *True Positive (TP)*, *True Negative (TN)*, *False Positive (FP)* ve *False Negative (FN)* biçiminde ayrıştırılarak doğruluk, duyarlılık ve özgüllük gibi temel ölçütler hesaplanmıştır.

- **Area Under Precision-Recall Curve (AUPRC):**

  Özellikle sınıf dengesizliği durumlarında anlamlı kabul edilen precision-recall eğrisi altındaki alan hesaplanmıştır.

- **Sınıfa özgü metrikler:**

  Pozitif ve negatif sınıflar için ayrı ayrı **precision** ve **recall** değerleri hesaplanarak modelin sınıf bazlı başarımı ölçülmüştür.

2.3.4. Hiperparametre Optimizasyonu

Modelin doğruluğunu ve verimliliğini artırmak amacıyla, çeşitli hiperparametreler sistematik olarak test edilmiştir:

- **LightGBM parametreleri:**
  - NumberOfLeaves: {64, 128, 256}
  - LearningRate: {0.005, 0.01}
  - NumberOfTrees: {500, 1000}
- **Sınıf ağırlıkları:**

  Dolandırıcılık sınıfı için 250.0, normal sınıf için 1.0 ağırlık verilerek dengesizliğin etkisi azaltılmıştır.

- **Erken durdurma:**

  Aşırı öğrenmeyi engellemek amacıyla EarlyStoppingRound = 100 parametresi uygulanmıştır.

- **Çeşitlilik artırıcı ayarlar:**
  - BaggingFraction = 0.8

- FeatureFraction = 0.8
- BaggingFrequency = 5

- **Regularizasyon:**
  Aşırı öğrenmeye karşı **L1** ve **L2** düzenlemeleri için sırasıyla 0.01 değeri uygulanmıştır.

2.3.5. ROC ve Eşik Analizi

Modelin eşik değerlerine göre davranışı ve genel tahmin başarısı ROC eğrisi üzerinden analiz edilmiştir:

- **Manuel AUC hesaplama:**

  Veri dengesizliğinin yol açabileceği ölçüm sapmalarını engellemek amacıyla AUC değeri özel olarak hesaplanmıştır.

- **Eşik değeri analizi:**

  Farklı eşik değerleri altında **precision**, **recall** ve **F1-score** metrikleri gözlemlenmiş ve optimal eşik değeri belirlenmiştir.

2.3.6. Model Değerlendirme Senaryoları

Gerçek dünya uygulamalarında karşılaşılabilecek durumlara karşı modelin kararlılığını test edebilmek için aşağıdaki senaryolar uygulanmıştır:

- **Yetersiz pozitif örnek senaryoları:**

  Veri setinde dolandırıcılık sayısının çok az olduğu durumlar için özelleştirilmiş değerlendirme prosedürleri geliştirilmiştir.

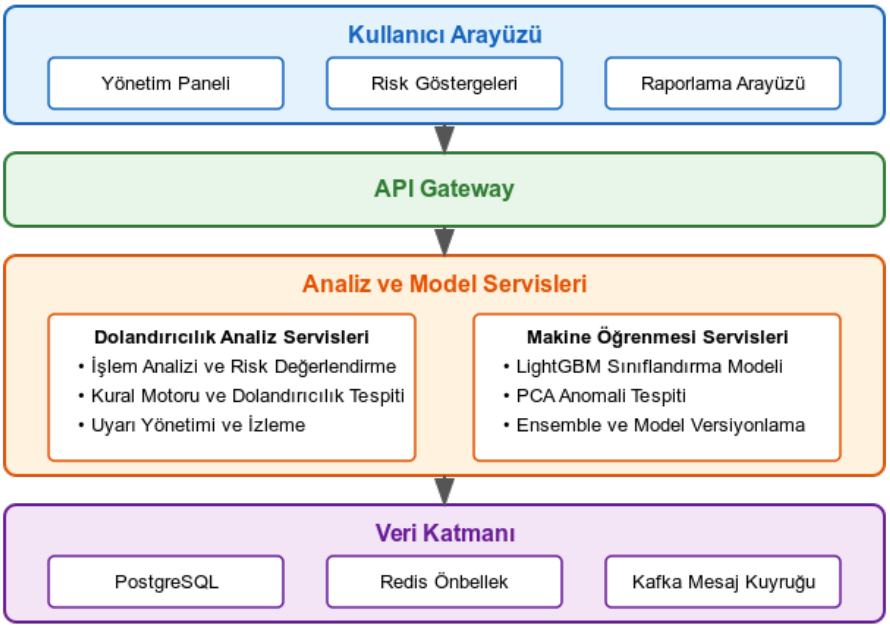
- **Performans raporu üretimi:**

  Modelin zaman içinde gösterdiği başarıyı izlemek amacıyla otomatik performans raporları oluşturulmuş ve versiyon kontrolü sağlanmıştır.

2.4. Sistem Altyapısı ve Uygulama Teknolojileri

Bu bölümde, FraudShield platformunun geliştirilmesinde kullanılan yazılım mimarisi ve uygulama teknolojilerine ilişkin detaylar sunulmaktadır. Sistem, gerçek zamanlı dolandırıcılık tespiti amacıyla geliştirilmiş olup, ölçeklenebilir, modüler ve kullanıcı odaklı bir yapıya sahiptir. Platform; veri işleme, analiz, görselleştirme ve yönetim süreçlerini birbirinden bağımsız ama entegre çalışan bileşenler aracılığıyla yürütmektedir.

Sistem mimarisine ilişkin genel yapı Şekil 2.4.1'de sunulmuştur.



Şekil 2.4.1. FraudShield sistem mimarisi bileşenleri

2.4.1. Arka Uç (Backend) Mimarisi

Sunucu taraflı uygulama yapısı, **.NET 9** tabanlı **mikroservis mimarisi** kullanılarak geliştirilmiştir. Bu mimari yaklaşım sayesinde; işlem analizi, risk değerlendirme, kullanıcı yönetimi ve bildirim gibi işlevler bağımsız servisler olarak yapılandırılmıştır. Her bir servis, kendi görev alanında çalışmakta ve bu durum sistemin yönetilebilirliğini ve hata izolasyonunu kolaylaştırmaktadır.

2.4.2. Ön Yüz (Frontend) Geliştirme

Kullanıcı arayüzü, modern web teknolojilerine dayalı olarak **React.js** kullanılarak geliştirilmiştir. Arayüz, **tek sayfa uygulama (SPA)** prensibine göre yapılandırılmış olup, kullanıcı etkileşimini hızlı ve kesintisiz hâle getirmektedir. **Bileşen tabanlı yapı**, kullanıcı

deneyimi göz önünde bulundurularak modüler ve sürdürülebilir bir arayüz tasarımına olanak tanımıştır.

2.4.3. Veritabanı ve Önbellekleme Katmanı

Veri saklama ve yönetim sürecinde **PostgreSQL** ilişkisel veritabanı tercih edilmiştir. Kullanıcı işlemleri, sistem konfigürasyonları ve model çıktıları bu veritabanında organize biçimde tutulmaktadır.

Performans artırımı amacıyla, sık erişilen verilere hızlı şekilde ulaşılabilmesi için **Redis** tabanlı önbellekleme mekanizması entegre edilmiştir. Bu yapı, sistemdeki işlem yoğunluğuna rağmen hızlı yanıt süreleri elde edilmesine katkı sağlamaktadır.

2.4.4. Geliştirme Süreci ve Sürüm Kontrolü

Yazılım geliştirme süreci, **Git** tabanlı versiyon kontrol sistemi ile yürütülmüştür. Ekip içi kod paylaşımı ve değişiklik takibi, **GitHub** üzerinden gerçekleştirilmiştir. Entegrasyon ve dağıtım adımları ise **CI/CD** (Continuous Integration / Continuous Deployment) prensiplerine uygun şekilde yapılandırılmış ve **GitHub Actions** aracılığıyla otomatikleştirilmiştir. Böylece sistem güncellemeleri düzenli, güvenli ve kesintisiz bir şekilde yönetilmiştir.

KAYNAKLAR

[1] SUN, Guanglin, et al. Digital finance and corporate financial fraud. International Review of Financial Analysis, 2023, 87: 102566. URL: https://www.sciencedirect.com/science/article/abs/pii/S1057521923000820 (Erişim tarihi: 23.03.2025)

[2] GOLD, Steve. The evolution of payment card fraud. Computer Fraud & Security, 2014, 2014.3: 12-17. URL: https://www.sciencedirect.com/science/article/abs/pii/S1361372314704713 (Erişim tarihi: 23.03.2025)

[3] RYMAN-TUBB, Nick F.; KRAUSE, Paul; GARN, Wolfgang. How Artificial Intelligence and machine learning research impacts payment card fraud detection: A survey and industry benchmark. Engineering Applications of Artificial Intelligence, 2018, 76: 130-157. URL: https://www.sciencedirect.com/science/article/abs/pii/S0952197618301520 (Erişim tarihi: 23.03.2025)

[4] BELLO, O. A., et al. Analysing the impact of advanced analytics on fraud detection: a machine learning perspective. European Journal of Computer Science and Information Technology, 2023, 11.6: 103-126. URL: https://www.researchgate.net/profile/Oluwabusayo-Bello/publication/381548526_Analysing_the_Impact_of_Advanced_Analytics_on_Fraud_Detection_A_Machine_Learning_Perspective/links/66736273d21e220d89c09836/Analysing-the-Impact-of-Advanced-Analytics-on-Fraud-Detection-A-Machine-Learning-Perspective.pdf (Erişim tarihi: 23.03.2025)

[5] HUANG, Dongxu, et al. CoDetect: Financial fraud detection with anomaly feature detection. Ieee Access, 2018, 6: 19161-19174. URL: https://ieeexplore.ieee.org/abstract/document/8325544 (Erişim tarihi: 23.03.2025)

[6] DU, Haichao, et al. AutoEncoder and LightGBM for credit card fraud detection problems. Symmetry, 2023, 15.4: 870. URL: https://www.mdpi.com/2073-8994/15/4/870 (Erişim tarihi: 23.03.2025)

[7] CHERGUI, Hamza, et al. Semi-supervised method to detect fraudulent transactions and identify fraud types while minimizing mounting costs. International journal of advanced computer science and applications (IJACSA), 2023, 14.2. URL: https://hal.science/hal-04209615/document (Erişim tarihi: 23.03.2025)

[8] SADDI, Venkata Ramana, et al. Fighting Insurance Fraud with Hybrid AI/ML Models: Discuss the Potential for Combining Approaches for Improved Insurance Fraud Detection. In: 2023 4th International Conference on Communication, Computing and Industry 6.0 (C216). IEEE, 2023. p. 01-06. URL: https://ieeexplore.ieee.org/abstract/document/10431155 (Erişim tarihi: 23.03.2025)

[9] KPMG Türkiye, "Üretken Yapay Zeka ile Finansın Yeni Normali", KPMG Türkiye, 2024. URL: https://assets.kpmg.com/content/dam/kpmg/tr/pdf/2024/01/uretken-yapay-zeka-ile-finansin-yeni-normali.pdf (Erişim tarihi: 22.03.2025)

[10] Kaggle, "Credit Card Fraud Detection Dataset", Kaggle, n.d. URL: https://www.kaggle.com/datasets/mlg-ulb/creditcardfraud (Erişim tarihi: 22.03.2025)

[11] BAHNSEN, Alejandro Correa, et al. Feature engineering strategies for credit card fraud detection. Expert Systems with Applications, 2016, 51: 134-142. URL: https://www.sciencedirect.com/science/article/abs/pii/S0957417415008386 (Erişim tarihi: 23.03.2025)

[12] HASAN, Basna Mohammed Salih; ABDULAZEEZ, Adnan Mohsin. A review of principal component analysis algorithm for dimensionality reduction. *Journal of Soft Computing and Data Mining*, 2021, 2.1: 20-30. URL: https://publisher.uthm.edu.my/ojs/index.php/jscdm/article/view/8032 (Erişim tarihi: 23.03.2025)

[13] KE, Guolin, et al. Lightgbm: A highly efficient gradient boosting decision tree. *Advances in neural information processing systems*, 2017, 30. URL: https://proceedings.neurips.cc/paper/2017/hash/6449f44a102fde848669bdd9eb6b76fa-Abstract.html

[14] Türkiye Yapay Zekâ İnisiyatifi (TRAI), "Bankacılıkta Yapay Zekâ ile Dolandırıcılık Tespiti", Türkiye.ai, 2024. URL: https://turkiye.ai/bankacilikta-yapay-zeka-dolandiricilik-tespiti/ (Erişim tarihi: 22.03.2025)

[15] ResearchGate, "Credit card fraud detection dataset (PCA ile türetilmiş)", 2023. URL: https://www.researchgate.net/figure/Credit-card-fraud-detection-dataset-The-dataset-is-obtained-by-doing-PCA-to-the-Kaggle_fig9_369404230 (Erişim tarihi: 22.03.2025)

[16] MDPI, "LightGBM Decision Tree Yapısı", Symmetry, 2023. URL: https://www.mdpi.com/2073-8994/15/4/870 (Erişim tarihi: 22.03.2025)