İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

KREDİ KARTI DOLANDIRICILIK TESPİTİNDE OPTUNA TABANLI HİPERPARAMETRE OPTİMİZASYONU İLE DERİN ÖĞRENME MODELLERİNİN KARŞILAŞTIRILMASI

YÜKSEK LİSANS TEZİ

Elvin MAMMADLİ

İşletme Anabilim Dalı

İşletme Yüksek Lisans Programı

Tez Danışmanı: Doç. Dr. Tuncay ÖZCAN

HAZİRAN 2026

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

KREDİ KARTI DOLANDIRICILIK TESPİTİNDE OPTUNA TABANLI HİPERPARAMETRE OPTİMİZASYONU İLE DERİN ÖĞRENME MODELLERİNİN KARŞILAŞTIRILMASI

YÜKSEK LİSANS TEZİ

Elvin MAMMADLİ
(403241015)

İşletme Anabilim Dalı

İşletme Yüksek Lisans Programı

Tez Danışmanı: Doç. Dr. Tuncay ÖZCAN

HAZİRAN 2026

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

COMPARISON OF DEEP LEARNING MODELS WITH OPTUNA-BASED HYPERPARAMETER OPTIMIZATION FOR CREDIT CARD FRAUD DETECTION

M.Sc. THESIS

Elvin MAMMADLI
(403241015)

Department of Management

Management Programme

Thesis Advisor: Assoc. Prof. Tuncay ÖZCAN

JUNE 2026

İTÜ, Lisansüstü Eğitim Enstitüsü'nün 403241015 numaralı Yüksek Lisans Öğrencisi Elvin MAMMADLİ, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı "KREDİ KARTI DOLANDIRICILIK TESPİTİNDE OPTUNA TABANLI HİPERPARAMETRE OPTİMİZASYONU İLE DERİN ÖĞRENME MODELLERİNİN KARŞILAŞTIRILMASI" başlıklı tezini aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

Tez Danışmanı :	Doç. Dr. Tuncay ÖZCAN İstanbul Teknik Üniversitesi	.....
Jüri Üyeleri :	Prof. Dr. Tolga KAYA İstanbul Teknik Üniversitesi	.....
	Prof. Dr. Ersin NAMLI İstanbul Üniversitesi-Cerrahpaşa	.....

Teslim Tarihi : 17 Haziran 2026
Savunma Tarihi : 29 Haziran 2026

Aileme,

ÖNSÖZ

Bu tez çalışmasının her aşamasında akademik bilgisi, geniş tecrübesi ve kıymetli yönlendirmeleri ile bana yol gösteren tez danışmanım Doç. Dr. Tuncay ÖZCAN'a en içten teşekkürlerimi sunarım. Kendisi, hem akademik yetkinliği hem de güçlü insani yönü sayesinde, yoğun emek gerektiren bu çalışmanın disiplinli bir biçimde tamamlanmasında belirleyici bir rol üstlenmiştir. Özellikle metodolojik tercihlerin gerekçelendirilmesi, deneysel kurguların eleştirel biçimde değerlendirilmesi ve elde edilen bulguların finansal bağlamda yorumlanması süreçlerinde sağladığı geri bildirimler benim için son derece değerli olmuştur.

İstanbul Teknik Üniversitesi İşletme Yüksek Lisans programı süresince, derslerde edindiğim bilgi ve perspektifi bu çalışmaya yansıtma fırsatı bulduğum tüm öğretim üyelerine teşekkür ederim. Ayrıca, lisansüstü eğitim sürecinde fikirlerini paylaşmaktan kaçınmayan, çalışmanın zorluklarını ve güzelliklerini birlikte deneyimlediğim arkadaşlarıma da şükranlarımı sunmak isterim.

Bu tezin teknik altyapısının kurulması sırasında, modellerin eğitimi için sağlanan hesaplama kaynaklarının ve açık kaynak topluluğunun katkısı yadsınamaz. TensorFlow, Keras, scikit-learn ve Optuna gibi kütüphanelerin geliştiricilerine ve açık kaynak ekosisteminin tüm katkıcılarına da teşekkür etmek isterim.

Son olarak, eğitim hayatım boyunca sabır ve fedakârlıkla yanımda olan, beni bugünlere getiren aileme ve sevdiklerime sonsuz teşekkür ederim. Onların koşulsuz desteği ve inancı olmasaydı bu çalışmanın tamamlanması mümkün olmazdı. Bu tez onlara ithaf edilmiştir.

Haziran 2026

Elvin MAMMADLİ
(Bilgisayar Mühendisi)

İÇİNDEKİLER

ÖNSÖZ ..... **ix**
İÇİNDEKİLER ..... **xi**
KISALTMALAR ..... **xiii**
SEMBOLLER ..... **xv**
ÇİZELGE LİSTESİ ..... **xvii**
ŞEKİL LİSTESİ ..... **xix**
ÖZET ..... **xxi**
SUMMARY ..... **xxiii**
1. GİRİŞ **1**
1.1 Arka Plan ve Motivasyon 1
1.2 Problem Tanımı 2
1.3 Çalışmanın Önemi 3
1.4 Araştırma Amaçları 3
2. LİTERATÜR ARAŞTIRMASI **5**
3. METODOLOJİ **17**
3.1 Finansal Dolandırıcılık Tespitinde Derin Öğrenme Yaklaşımı 17
3.1.1 Derin öğrenmenin teorik temelleri 18
3.1.2 Aktivasyon fonksiyonları 19
3.1.2.1 Sigmoid fonksiyonu 19
3.1.2.2 Hiperbolik tanjant (tanh) fonksiyonu 19
3.1.2.3 Rectified linear unit (ReLU) 19
3.2 Veri Ön İşleme Yöntemleri 19
3.2.1 Veri temizleme ve eksik veri analizi 20
3.2.2 Kategorik değişkenlerin kodlanması 20
3.2.3 Veri standartlaştırması 20
3.2.4 Sınıf dengesizliğinin giderilmesi (SMOTE Algoritması) 21
3.3 Evrişimli Sinir Ağları (Convolutional Neural Networks) 22
3.3.1 Evrişimli sinir ağları (Convolutional Neural Networks - CNN) 22
3.3.1.1 Evrişim operasyonu (convolution operation) 22
3.3.1.2 Aktivasyon fonksiyonu (ReLU) 22
3.3.1.3 Havuzlama (pooling) katmanı 23
3.4 Uzun-Kısa Süreli Bellek (Long Short-Term Memory - LSTM) 23
3.4.1 Tekrarlayan sinir ağları (RNN) ve kısıtları 23
3.4.2 LSTM hücre mimarisi ve matematiksel formülasyonu 23
3.4.2.1 Unutma kapısı (forget gate) 24
3.4.2.2 Giriş kapısı (input gate) 24
3.4.2.3 Hücre durumunun güncellenmesi 24
3.4.2.4 Çıkış kapısı (output gate) 25
3.5 Model Eğitimi ve Hiperparametre Optimizasyonu 25
3.5.1 Kayıp fonksiyonu (Loss Function) 25
3.5.2 Optimizasyon algoritması: Adam 25
3.5.3 Aşırı öğrenmeyi önleme (Regularization) 26

3.5.4 Bayes tabanlı hiperparametre optimizasyonu (Optuna) 26
3.6 Performans Değerlendirme Metrikleri 27
3.6.1 Karmaşıklık matrisi (Confusion Matrix) 27
3.6.2 Kesinlik (Precision) 28
3.6.3 Anma (Recall / Sensitivity) 28
3.6.4 F1-skoru (F1-Score) 28
3.6.5 ROC eğrisi ve AUC (Area Under Curve) 28
3.6.6 Doğruluk (Accuracy) 29
3.7 Model Doğrulama Yöntemi: Hold-out Doğrulama 29
3.8 Klasik Makine Öğrenmesi Algoritmaları 29
3.8.1 K-En yakın komşu (KNN) 30
3.8.2 Naive Bayes 30
3.8.3 Lojistik regresyon 30
3.8.4 Rastgele orman (Random Forest) 32
3.8.5 Aşırı gradyan artırma (XGBoost) 32
4. UYGULAMA 33
4.1 Çalışma Akışı 33
4.2 Veri Seti 34
4.3 Veri Ön İşleme 36
4.3.1 Eğitim ve test bölünmesi 36
4.3.2 Standartlaştırma 36
4.3.3 SMOTE ile sentetik örnekleme 36
4.3.4 Veri yeniden şekillendirme 37
4.4 Önerilen Modellerin Uygulanması 37
4.4.1 CNN mimarisi 38
4.4.2 LSTM mimarisi 38
4.4.3 CNN-LSTM hibrit mimarisi 38
4.4.4 Varsayılan parametrelerle eğitim sonuçları 38
4.5 Hiperparametre Optimizasyonu: Optuna 39
4.5.1 Optimizasyon stratejisi 39
4.5.2 Sınıf ağırlıklandırma (Class Weighting) 41
4.5.3 Karar eşiği optimizasyonu (eşik ayarlama) 41
4.6 Optimize Edilmiş Modellerin Sonuçları 42
4.7 Maliyet Analizi 46
4.8 Bulguların Değerlendirilmesi 49
4.9 Klasik Makine Öğrenmesi Modelleri ile Karşılaştırma 50
4.10 Ağaç Tabanlı Topluluk Modelleri ile Karşılaştırma 51
4.11 Genel Karşılaştırma: Klasik ML, Ağaç Toplulukları ve Derin Öğrenme 54
5. SONUÇLAR VE ÖNERİLER 57
5.1 Çalışmanın Akademik ve Pratik Katkıları 58
5.2 Yönetsel Çıkarımlar ve Politika Önerileri 59
5.2.1 Operasyonel verimlilik ve iş yükü yönetimi 59
5.2.2 Müşteri deneyimi ve sadakat 59
5.2.3 Stratejik bütçe planlaması ve risk yönetimi 60
5.3 Çalışmanın Sınırlılıkları 60
KAYNAKLAR 63
ÖZGEÇMİŞ 69

KISALTMALAR

ADASYN : Uyarlamalı Sentetik Örnekleme (Adaptive Synthetic Sampling)

AE : Otokodlayıcı (Autoencoder)

ANN : Yapay Sinir Ağı (Artificial Neural Network)

AUC : Eğri Altında Kalan Alan (Area Under Curve)

CNN : Evrişimli Sinir Ağı (Convolutional Neural Network)

DL : Derin Öğrenme (Deep Learning)

DM : Veri Madenciliği (Data Mining)

ELM : Aşırı Öğrenme Makinesi (Extreme Learning Machine)

FL : Birleşik Öğrenme (Federated Learning)

FN : Yanlış Negatif (False Negative)

FP : Yanlış Pozitif (False Positive)

GA : Genetik Algoritma (Genetic Algorithm)

GAN : Üretici Çekişmeli Ağ (Generative Adversarial Network)

GCN : Grafik Evrişimli Ağ (Graph Convolutional Network)

GNN : Grafik Sinir Ağı (Graph Neural Network)

GRU : Kapılı Yineleme Birimi (Gated Recurrent Unit)

KNN : K-En Yakın Komşu (K-Nearest Neighbors)

LIME : Yerel Yorumlanabilir Model-Agnostik Açıklamalar (Local Interpretable Model-Agnostic Explanations)

LogReg : Lojistik regresyon (Logistic Regression)

LSTM : Uzun-Kısa Süreli Bellek (Long Short-Term Memory)

ML : Makine Öğrenmesi (Machine Learning)

MLP : Çok Katmanlı Algılayıcı (Multi-Layer Perceptron)

NB : Naive Bayes

NLP : Doğal Dil İşleme (Natural Language Processing)

OCSVM : Tek Sınıflı Destek Vektör Makinesi (One-Class Support Vector Machine)

PCA : Temel Bileşen Analizi (Principal Component Analysis)

RF : Rastgele orman (Random Forest)

RNN : Yinelenen Sinir Ağı (Recurrent Neural Network)

ROC **:** Alıcı İşletim Karakteristiği (Receiver Operating Characteristic)

SHAP **:** SHapley Katkılı Açıklamalar (SHapley Additive exPlanations)

SMOTE **:** Sentetik Azınlık Aşırı Örnekleme Tekniği (Synthetic Minority Over-sampling Technique)

SVM **:** Destek Vektör Makinesi (Support Vector Machine)

TN **:** Doğru Negatif (True Negative)

TP **:** Doğru Pozitif (True Positive)

XAI **:** Açıklanabilir Yapay Zekâ (Explainable Artificial Intelligence)

XGBoost **:** Aşırı Gradyan Artırma (Extreme Gradient Boosting)

SEMBOLLER

b : Bias (eşik değeri) terimi

$\mathbf{C_t}$: Hücre durumu (cell state)

$\mathbf{\tilde{C}_t}$: Aday hücre durumu (candidate cell state)

e : Doğal logaritma tabanı (Euler sayısı)

f(·) : Aktivasyon fonksiyonu

$\mathbf{f_t}$: Unutma kapısı (forget gate)

$\mathbf{h_t}$: Gizli durum / hücre çıkışı (hidden state)

$\mathbf{i_t}$: Giriş kapısı (input gate)

$\mathbf{o_t}$: Çıkış kapısı (output gate)

t : Zaman adımı

W, w : Ağırlık matrisi / ağırlık katsayısı (weight)

x : Giriş özellik vektörü (input vector)

y : Nöron çıkışı (output)

η : Öğrenme hızı (learning rate)

σ : Sigmoid aktivasyon fonksiyonu

Σ : Ağırlıklı toplam operatörü

⊙ : Eleman bazında çarpım (Hadamard çarpımı)

ÇİZELGE LİSTESİ

Çizelge 2.1 : Literatürdeki Çalışmaların Karşılaştırılması **15**
Çizelge 4.1 : Veri Seti Genel İstatistikleri **35**
Çizelge 4.2 : Varsayılan Parametrelerle Model Performans Analizi..... **39**
Çizelge 4.3 : CNN Modeli için Optuna Hiperparametre Optimizasyonu..... **40**
Çizelge 4.4 : LSTM Modeli için Optuna Hiperparametre Optimizasyonu **40**
Çizelge 4.5 : CNN-LSTM Hibrit Modeli için Optuna Hiperparametre **40**
Optimizasyonu
Çizelge 4.6 : Optimize Edilmiş Modellerin Karşılaştırmalı Performans Analizi **42**
Çizelge 4.7 : Optimize Edilmiş Modellerin PR-AUC Değerleri **45**
Çizelge 4.8 : Modellerin Maliyet Analizi **46**
Çizelge 4.9 : Farklı Maliyet Senaryoları Altında Toplam Maliyetler (€) **48**
Çizelge 4.10 : Klasik Makine Öğrenmesi Modellerinin Performans Analizi..... **51**
Çizelge 4.11 : Ağaç Tabanlı Topluluk Modellerinin Performans Analizi..... **52**
Çizelge 4.12 : Tüm Modellerin Karşılaştırmalı Performans Analizi..... **54**

ŞEKİL LİSTESİ

Şekil 4.1 : Uygulamanın Akış Diyagramı **34**
Şekil 4.2 : Sınıf Dağılımı **35**
Şekil 4.3 : SMOTE Sonrası Sınıf Dağılımı **37**
Şekil 4.4 : CNN Modeli için Karar Eşiği Önem Analizi **42**
Şekil 4.5 : Eğitim ve Doğrulama Eğrileri **43**
Şekil 4.6 : Karmaşıklık Matrisi Karşılaştırması **43**
Şekil 4.7 : ROC Eğrileri Karşılaştırması **44**
Şekil 4.8 : Parametre Optimizasyonunun Kesinlik Ölçütüne Etkisi **44**
Şekil 4.9 : Kesinlik-Anma Eğrileri Karşılaştırması **45**
Şekil 4.10 : Maliyet Analizi Bar Grafiği **47**
Şekil 4.11 : Maliyet Duyarlılık Analizi **49**
Şekil 4.12 : XGBoost ve Rastgele Orman için Kesinlik-Anma Eğrileri **54**
Şekil 4.13 : XGBoost ve Rastgele Orman için Karmaşıklık Matrisleri **54**

KREDİ KARTI DOLANDIRICILIK TESPİTİNDE OPTUNA TABANLI HİPERPARAMETRE OPTİMİZASYONU İLE DERİN ÖĞRENME MODELLERİNİN KARŞILAŞTIRILMASI

ÖZET

Finansal sistemlerin dijitalleşmesi ve elektronik ödeme yöntemlerinin yaygınlaşmasıyla birlikte, kredi kartı dolandırıcılığı bireysel kullanıcılar, finans kurumları ve düzenleyici otoriteler için ciddi bir risk unsuru hâline gelmiştir. Geleneksel kural tabanlı sistemlerin değişen dolandırıcılık örüntüleri karşısında yetersiz kalması, veriden öğrenme yeteneğine sahip yapay zekâ tabanlı çözümleri zorunlu kılmıştır. Bu tez çalışmasında, Kaggle üzerinde yayımlanan ve yaygın olarak kullanılan Avrupa Kredi Kartı İşlemleri veri seti üzerinde, makine öğrenmesi ve derin öğrenme tabanlı dolandırıcılık tespit modelleri geliştirilmiş ve karşılaştırılmıştır.

Çalışma kapsamında ilk olarak K-En yakın komşu (KNN), Naive Bayes ve Lojistik regresyon gibi klasik makine öğrenmesi yöntemleri test edilmiş; bu yöntemlerin yüksek dengesizlik altında Kesinlik değerlerinin %6–45 aralığında kaldığı ve pratik olarak yetersiz olduğu gösterilmiştir. Ardından üç farklı derin öğrenme mimarisi ele alınmıştır: tek boyutlu Evrişimli Sinir Ağı (1D-CNN), Uzun-Kısa Süreli Bellek (LSTM) ve CNN-LSTM hibrit modeli. Veri setinin %0,17 oranında dolandırıcılık örneği içeren ileri derecede dengesiz yapısı, model eğitiminde önemli bir zorluk oluşturmuştur. Bu probleme yönelik olarak SMOTE algoritmasıyla sentetik veri üretimi, sınıf ağırlıklandırma ({0:1, 1:10}) ve karar eşiği optimizasyonu gibi tamamlayıcı stratejiler bütünleşik bir çerçeve içinde uygulanmıştır.

Modellerin performansını artırmak amacıyla Optuna çerçevesi kullanılarak Bayes tabanlı hiperparametre optimizasyonu gerçekleştirilmiştir. Her model için 50 deneme (trial) yürütülmüş ve hedef metrik olarak Kesinlik belirlenmiştir. Optimizasyon sürecinde filtre sayısı, LSTM birim sayısı, seyreltme oranı, öğrenme hızı ve yığın boyutu gibi kritik parametreler için en uygun kombinasyonlar tespit edilmiştir. eşik ayarlama aşamasında karar eşiği 0,50 yerine, üç modelde de toplam maliyeti minimize eden 0,98 değerine çekilmiş, böylece yanlış alarm sayısı önemli ölçüde azaltılmıştır.

Deneysel bulgular, Optuna tabanlı optimizasyonun özellikle LSTM ve CNN-LSTM modellerinde belirgin iyileşmeler sağladığını ortaya koymaktadır. LSTM modelinin Kesinlik değeri %44'ten %60'a, CNN-LSTM hibrit modelinin Kesinlik değeri ise %55'ten %76'ya yükselmiştir. Varsayılan parametrelerle dahi %69 Kesinlik değerine ulaşan CNN modelinde ise optimizasyon, performansı benzer seviyede korumuştur. F1-skoru değerlerinde de tutarlı iyileşmeler gözlemlenmiş; LSTM modelinde F1-skoru değeri %58'den %70'e, CNN-LSTM hibrit modelinde ise %68'den %80'e yükselmiştir. AUC-ROC değerleri üç modelde de 0,94 üzerinde gerçekleşmiştir.

Üç model karşılaştırıldığında, en yüksek Kesinlik (%76), en yüksek F1-skoru (%80), en yüksek PR-AUC (0,8506) ve en düşük tahmini maliyet (8.800 €) değerlerini CNN-LSTM hibrit modeli sağlamıştır. CNN modeli en yüksek AUC-ROC değerini (0,9575) sunarken, LSTM modeli zamansal bağımlılıkları yakalamada tamamlayıcı bir rol oynamıştır. Bulgular, dolandırıcılık tespiti gibi sınıf dengesizliğinin yoğun olduğu problemlerde, hiperparametre optimizasyonu, sınıf ağırlıklandırma ve karar eşiği ayarlamasının birlikte uygulanmasının; yalnızca sentetik örnekleme veya tek başına model seçiminden çok daha etkili sonuçlar verdiğini göstermektedir.

Çalışma kapsamında ayrıca, yapısal (tablo hâlindeki) verilerde güçlü referans yöntemler olarak kabul edilen iki ağaç tabanlı topluluk modeli olan Rastgele Orman ve XGBoost, aynı veri ve ön işleme koşulları altında Optuna ile optimize edilerek değerlendirilmiştir. XGBoost %96 Kesinlik ve %85 F1-skoru ile tüm modeller arasında en yüksek kesinliğe ulaşmış, yalnızca 3 yanlış alarm üretmiştir; Rastgele Orman ise 0,9821 ile en yüksek AUC-ROC değerini sağlamıştır. Bununla birlikte, her iki topluluk modeli de derin öğrenme modellerine kıyasla daha fazla dolandırıcılık vakası kaçırmış (sırasıyla 23 ve 20 yanlış negatif); benimsenen maliyet çerçevesinde en düşük toplam maliyeti (8.800 €) CNN-LSTM hibrit modeli korumuştur (XGBoost: 11.650 €, Rastgele Orman: 10.800 €). Bu bulgular, ağaç tabanlı toplulukların yapısal verilerdeki gücünü teyit ederken, kesinlik odaklı topluluklar ile duyarlılık odaklı derin mimariler arasındaki tercihin kurumun maliyet yapısına bağlı olduğunu göstermektedir.

Sonuç olarak, bu tez; klasik makine öğrenmesi algoritmalarının ileri düzey sınıf dengesizliği altındaki yetersizliğini deneysel olarak ortaya koyarak derin öğrenme yaklaşımlarının üstünlüğünü göstermiş; Optuna ile yürütülen sistematik hiperparametre optimizasyonunun bu modellerin performansını önemli ölçüde artırdığını kanıtlamıştır. Önerilen yaklaşım, bankacılık sistemlerinde uygulanabilir bir çerçeve sunarak hem akademik literatüre hem de finans sektörüne pratik katkı sağlamaktadır.

COMPARISON OF DEEP LEARNING MODELS WITH OPTUNA-BASED HYPERPARAMETER OPTIMIZATION FOR CREDIT CARD FRAUD DETECTION

SUMMARY

With the rapid digitalization of financial systems and the widespread adoption of electronic payment instruments, credit card transactions have become a fundamental component of global economic activity. According to the 2023 Nilson Report, worldwide losses caused by credit card fraud are expected to exceed 49 billion USD by 2027. This sharp upward trajectory poses a significant threat to consumers and financial institutions alike, and underscores the growing need for robust and adaptive fraud detection systems. Traditional rule-based mechanisms, which rely on fixed thresholds, blacklists, and expert-defined heuristics, struggle to keep pace with the constantly evolving tactics employed by fraudsters. As a result, data-driven artificial intelligence approaches that learn directly from transaction histories have emerged as the dominant paradigm in modern fraud detection research.

Machine learning (ML) and, more recently, deep learning (DL) techniques have offered substantial improvements over traditional methods by automatically discovering complex patterns in large and high-dimensional financial datasets. Convolutional Neural Networks (CNN) are capable of capturing local temporal patterns in transaction sequences, while Long Short-Term Memory (LSTM) networks excel at modelling long-term dependencies in user behaviour. Hybrid CNN-LSTM architectures combine these complementary strengths and have been shown to outperform either component when used alone. Nevertheless, the performance of deep learning models is highly sensitive to hyperparameter selection. Manually tuning parameters such as the number of filters, dropout rate, learning rate, and batch size is both labour-intensive and error-prone, motivating the use of systematic optimisation frameworks such as Optuna, which leverages Bayesian optimisation through the Tree-structured Parzen Estimator algorithm.

A central challenge in credit card fraud detection is the extreme class imbalance inherent in real-world transaction data, where fraudulent cases typically constitute less than 1% of total observations. While many existing studies attempt to address this issue through synthetic oversampling techniques such as SMOTE, these methods, when used in isolation, tend to inflate Recall while leaving Precision values disappointingly low. In practical banking environments, low Precision translates directly into a high volume of false alarms, increased operational cost, and degraded customer experience. The literature gap therefore lies in the absence of integrated frameworks that simultaneously address hyperparameter optimisation, class weighting, and decision-threshold tuning under a unified evaluation methodology. The present thesis was designed to address this gap by experimentally comparing three deep learning architectures within a coherent optimisation pipeline.

The experimental study was carried out using the widely adopted European Credit Card Transactions dataset published on Kaggle. This dataset contains 284,807 transactions performed by European cardholders in September 2013, of which only 492 (0.17%) are labelled as fraudulent. The dataset comprises 30 features: 28 anonymised components (V1–V28) obtained via Principal Component Analysis to preserve privacy, the Time feature representing seconds elapsed since the first transaction, the Amount feature denoting the transaction value, and a binary Class label distinguishing legitimate (0) from fraudulent (1) transactions. The severe imbalance and the limited number of available features make this dataset a particularly challenging benchmark for fraud detection research.

The preprocessing pipeline began with the stratified partitioning of the dataset into 80% training and 20% test subsets, preserving the original class distribution in both partitions. The training set contained 227,845 transactions including 394 fraudulent cases, while the test set contained 56,962 transactions including 98 fraudulent cases. All numeric features were standardised using Z-score normalisation through the scikit-learn StandardScaler, fit exclusively on the training set to prevent data leakage. SMOTE was subsequently applied only to the training set, with k=5 nearest neighbours, balancing the fraud class to match the majority class size (227,451 samples each). To accommodate the input requirements of CNN and LSTM architectures, each feature vector was reshaped into a three-dimensional tensor of size (30, 1), treating each transaction as a 30-step signal with a single channel.

Three deep learning architectures were implemented in TensorFlow and Keras. The 1D-CNN model consisted of two convolution-pooling blocks (with 64 and 128 filters, kernel size 3, max-pooling 2), followed by a flatten layer, a dense layer with 64 neurons, and a sigmoid output. The LSTM model used two stacked LSTM layers (64 and 32 units) with dropout regularisation. The CNN-LSTM hybrid combined the convolutional feature-extraction blocks with an LSTM layer (64 units) to model the temporal dependencies within the locally-extracted features. All models employed Binary Cross-Entropy loss, the Adam optimiser, EarlyStopping callbacks, and a class-weight ratio of {0:1, 1:10} during training to further reinforce the importance of the minority class.

To systematically improve model performance, Bayesian hyperparameter optimisation was performed using the Optuna framework. For each architecture, 50 trials were executed with Precision as the objective metric. The choice of Precision rather than F1-score or Accuracy reflects the operational priority of banking institutions to minimise false-alarm cost. Search spaces were defined for the number of filters [32, 64, 128, 256], LSTM units [32, 64, 128, 256], dense units [32, 64, 128], dropout rate [0.1, 0.5], learning rate [1e-4, 1e-2] on a logarithmic scale, and batch size [128, 256, 512]. To prevent over-aggressive Precision-Recall trade-offs, a minimum-Recall constraint of 0.75 was enforced; trials yielding Recall below this threshold were assigned a fitness of zero. In parallel, decision-threshold tuning was applied, raising the default 0.50 cutoff to a cost-minimising value of 0.98 for all three models, thereby filtering out low-confidence positive predictions.

The experimental findings clearly demonstrate the value of the proposed integrated optimisation framework. With default parameters and a 0.50 decision threshold, the CNN, LSTM, and CNN-LSTM models achieved Precision values of 69%, 44%, and 55% respectively, against Recall values consistently around 85–88%. While the CNN model performed relatively well even without tuning, the lower Precision

values of the LSTM and hybrid models translated to elevated volumes of false alarms, leaving substantial room for systematic improvement. After applying the integrated optimisation pipeline, Precision rose markedly for these models: from 44% to 60% for LSTM and from 55% to 76% for CNN-LSTM, corresponding to relative improvements of roughly 1.4-fold in both cases, while the already strong CNN model retained its 69% Precision. F1-scores improved correspondingly, from 58% to 70% for LSTM and from 68% to 80% for CNN-LSTM. AUC-ROC values exceeded 0.94 across all three optimised models, confirming the robustness of the underlying classification capability.

To complement the classical statistical metrics, a financial cost analysis was conducted. Each missed fraud (false negative) was assigned a cost of 500 EUR, while each false alarm (false positive) was assigned an operational cost of 50 EUR, in line with prior literature estimates. Under this framework, the optimised CNN-LSTM hybrid produced both the lowest false-alarm count (26 false positives) and the lowest total cost at 8,800 EUR, making it the most financially attractive option. The CNN and LSTM models followed with total costs of 10,300 EUR and 10,250 EUR respectively. A complementary scenario-based sensitivity analysis showed that the hybrid model remains the least costly option under most cost assumptions, although in scenarios where the unit cost of a missed fraud becomes dominant, high-Recall alternatives such as KNN can become competitive. These findings illustrate that model selection in real banking contexts must be informed not only by statistical metrics but also by the institution's operational cost structure.

For comparative purposes, three classical machine learning algorithms were also evaluated under identical preprocessing conditions: K-Nearest Neighbors (k=5), Gaussian Naive Bayes, and Logistic Regression (max_iter=1000). KNN produced the strongest classical-ML performance with 45% Precision, while Naive Bayes and Logistic Regression both stagnated at only 6% Precision, indicating that for every true fraud captured these models generated roughly seventeen false alarms. These results experimentally substantiate the literature claim that classical ML algorithms are insufficient for highly imbalanced fraud-detection tasks, and provide concrete justification for the adoption of deep learning approaches augmented by systematic optimisation.

The principal academic contributions of this thesis are fourfold. First, it experimentally demonstrates the inadequacy of classical machine learning algorithms in highly imbalanced fraud-detection tasks, thereby providing an empirical foundation for deep learning adoption. Second, it offers a head-to-head comparison of three deep learning architectures within a unified optimisation pipeline, removing many of the confounding variables that plague cross-paper comparisons. Third, it shows that the combined use of Optuna-based hyperparameter optimisation, class weighting, and decision-threshold tuning yields substantially better outcomes than synthetic oversampling alone, an integration rarely examined as a coherent framework in prior literature. Fourth, by introducing a financial cost-analysis layer, the work extends evaluation beyond classical statistical metrics and provides decision-makers with a more transparent and actionable picture of model behaviour.

In addition to the deep learning architectures, two tree-based ensemble models widely regarded as strong baselines for tabular data, Random Forest and XGBoost, were evaluated under identical data and preprocessing conditions, with their hyperparameters likewise optimised via Optuna. XGBoost achieved the highest

Precision (96%) and F1-score (85%) among all models, producing only three false alarms, while Random Forest attained the highest AUC-ROC value (0.9821). Nevertheless, both ensembles missed more fraudulent cases than the deep models (23 and 20 false negatives, respectively), and under the adopted cost framework the CNN-LSTM hybrid retained the lowest estimated total cost (8,800 EUR), compared with 11,650 EUR for XGBoost and 10,800 EUR for Random Forest. These findings confirm that tree-based ensembles remain highly competitive on structured tabular data, while highlighting that the choice between precision-oriented ensembles and recall-oriented deep architectures ultimately depends on the institution's cost structure.

Several limitations and avenues for future research deserve mention. The dataset employed dates from 2013 and the fraud patterns it captures may have evolved since then, raising the issue of concept drift in real deployments. The PCA-anonymised features also constrain the scope of feature-engineering experimentation. The unit-cost values used in the financial analysis (500 EUR and 50 EUR) represent typical figures drawn from the literature and may vary across institutions. Future work could explore Focal Loss and other imbalance-aware loss functions, advanced sampling schemes such as SMOTE-ENN, ADASYN, or Borderline-SMOTE, ensemble strategies including stacking and voting, Transformer-based architectures with attention mechanisms, explainable AI techniques such as SHAP and LIME, Federated Learning for privacy-preserving deployment, larger Optuna trial budgets (e.g., 200–500 trials), and Graph Neural Networks for modelling relational links between transactions. Overall, the proposed framework offers a strong methodological foundation that can be readily extended in these directions, while already providing an immediately applicable solution for banking and financial-services institutions.

1. GİRİŞ

Bu bölümde, tez çalışmasının arka planı, ele alınan problem, çalışmanın önemi, araştırma amaçları ve tezin genel yapısı sunulmaktadır.

1.1 Arka Plan ve Motivasyon

21. yüzyılda dijital dönüşüm ve finansal teknolojilerin (FinTech) hızla yaygınlaşmasıyla birlikte, kredi kartı işlemleri küresel ekonomik faaliyetlerin temel taşlarından biri haline gelmiştir. Nilson Report'un 2023 yılı verilerine göre dünya genelinde kredi kartı dolandırıcılığından kaynaklanan kayıpların 2027 yılına kadar 49 milyar doları aşması beklenmektedir. Bu hızlı büyüme, hem tüketiciler hem de finans kurumları için ciddi bir tehdit oluşturmakta; aynı zamanda etkili dolandırıcılık tespit sistemlerine olan ihtiyacı her geçen gün daha belirgin hâle getirmektedir.

Geleneksel olarak finansal kurumlar, dolandırıcılık tespiti için kural tabanlı (rule-based) sistemler kullanmıştır. Bu sistemler; sabit eşik değerleri, kara listeler ve uzman kuralları üzerinden çalışmakta; ancak değişen ve sürekli evrilen dolandırıcılık taktikleri karşısında esneklik ve uyum kabiliyetinden yoksundur. Saldırganlar yeni desenler keşfettikçe, sabit kuralların güncellenmesi gerekmekte; bu da hem maliyetli hem de gecikmeli bir süreç olarak ortaya çıkmaktadır. Bu durum, veriden öğrenme yeteneğine sahip yapay zekâ tabanlı yaklaşımların önemini artırmıştır.

Makine öğrenmesi (Machine Learning – ML) ve özellikle son yıllarda öne çıkan derin öğrenme (Deep Learning – DL) teknikleri, büyük ve karmaşık veri kümeleri üzerinde otomatik desen tanıma yetenekleriyle dolandırıcılık tespitinde önemli avantajlar sunmaktadır. Evrişimli Sinir Ağları (Convolutional Neural Networks – CNN), zaman serisi verilerindeki yerel örüntüleri yakalayabilirken; Uzun-Kısa Süreli Bellek (Long Short-Term Memory – LSTM) ağları, kullanıcı davranışlarındaki uzun vadeli zamansal bağımlılıkları modellemede etkin sonuçlar üretmektedir. Hibrit mimarilerin (örneğin CNN-LSTM) ise her iki yaklaşımı bir araya getirerek tek

başlarına elde edilen performansların ötesinde sonuçlar verebildiği literatürde gösterilmiştir.

Bununla birlikte, derin öğrenme modellerinin performansı büyük ölçüde hiperparametre seçimine bağlıdır. Filtre sayısı, seyreltme (dropout) oranı, öğrenme hızı ve yığın (batch) boyutu gibi hiperparametrelerin optimal değerlerinin bulunması; manuel deneme yanılma yerine sistematik bir optimizasyon süreciyle gerçekleştirildiğinde model başarısında belirgin iyileşmeler sağlanabilmektedir. Bu noktada, Bayes optimizasyonu temelli Optuna çerçevesi, akıllı arama stratejileri ve erken durdurma yetenekleri sayesinde geleneksel grid search ve random search yöntemlerine göre daha verimli sonuçlar üretmektedir.

1.2 Problem Tanımı

Kredi kartı dolandırıcılık tespitinin temel zorluklarından biri, veri setlerindeki ileri düzey sınıf dengesizliğidir. Gerçek dünyadan toplanan finansal işlem verilerinde dolandırıcılık örnekleri tipik olarak %1'in altında kalmakta; bu çalışmada kullanılan Avrupa Kredi Kartı Veri Seti'nde de bu oran yalnızca %0,17 düzeyindedir. Sınıf dengesizliği, modellerin çoğunluk sınıfına aşırı uyum sağlamasına ve dolandırıcılığı kaçırmasına yol açabilmektedir. Yalnızca Doğruluk metriğine dayalı değerlendirmelerin yanıltıcı olmasının başlıca nedeni de budur.

Literatürdeki birçok çalışma, sınıf dengesizliğini sentetik örnekleme (örneğin SMOTE) gibi tekniklerle gidermeye odaklanmış olsa da; tek başına bu yaklaşımlar genellikle yüksek Anma (Recall) değerleri üretirken, Kesinlik (Precision) değerlerini düşük seviyelerde bırakmaktadır. Bu durum, gerçek bankacılık sistemlerinde yüksek sayıda yanlış alarmın üretilmesine ve operasyonel maliyetlerin artmasına neden olmaktadır. Dolayısıyla, hem dolandırıcılık vakalarını yakalamada (Anma) hem de gereksiz alarmı önlemede (Kesinlik) yüksek başarı sağlayabilen modellerin geliştirilmesi araştırmacılar için açık bir problem alanı oluşturmaktadır.

Bu bağlamda tezde ele alınan ana problem şu şekilde tanımlanmaktadır: ileri düzey sınıf dengesizliğine sahip kredi kartı işlem verisi üzerinde, derin öğrenme tabanlı modellerin (CNN, LSTM, CNN-LSTM hibrit) Kesinlik ve Anma metrikleri arasında dengeli bir başarı sergilemesini sağlayacak bir optimizasyon ve değerlendirme çerçevesinin geliştirilmesi. Bu çerçeve; veri ön işleme, sentetik örnekleme,

hiperparametre optimizasyonu, sınıf ağırlıklandırma ve karar eşiği ayarlamasını bütünleşik biçimde ele almalıdır.

1.3 Çalışmanın Önemi

Bu tez çalışmasının önemi birkaç açıdan ele alınabilir. Akademik açıdan; literatürdeki birçok çalışma yalnızca tek bir derin öğrenme modeline veya tek bir optimizasyon stratejisine odaklanmaktayken, bu çalışma üç farklı mimariyi (CNN, LSTM, CNN-LSTM hibrit) eşit koşullarda ve bütünleşik bir optimizasyon çerçevesi içinde karşılaştırmaktadır. Optuna tabanlı hiperparametre optimizasyonunun, eşik ayarlama (threshold tuning) ve sınıf ağırlıklandırma ile birlikte uygulanması, literatürde yaygın olarak ele alınmamış bir kombinasyondur.

Pratik açıdan; çalışmada elde edilen modeller yalnızca yüksek doğruluk değil, aynı zamanda finansal anlamda yorumlanabilir maliyet metrikleri (kaçan dolandırıcılık başına 500 €, yanlış alarm başına 50 €) üzerinden değerlendirilmektedir. Bu yaklaşım, modellerin gerçek bankacılık operasyonlarındaki etkisini somutlaştırmakta ve karar vericilere daha şeffaf bilgi sunmaktadır. Önerilen çerçevenin gerçek zamanlı dolandırıcılık tespit sistemlerine uyarlanması mümkündür.

Toplumsal açıdan ise dolandırıcılık tespitinin etkin biçimde yapılması; tüketicilerin finansal güvenliğini korumakta, finans kurumlarının itibar ve kayıplarını azaltmakta ve genel olarak ödeme ekosistemine güveni artırmaktadır. Bu nedenle çalışmanın bulguları yalnızca akademik değil, doğrudan toplumsal fayda da üreten bir nitelik taşımaktadır.

1.4 Araştırma Amaçları

Bu tez çalışmasının temel amaçları aşağıdaki gibi özetlenebilir:

- İleri düzey sınıf dengesizliğine sahip kredi kartı işlem verisi üzerinde, üç farklı derin öğrenme mimarisinin (CNN, LSTM ve CNN-LSTM hibrit) performansını eşit koşullarda karşılaştırmak.
- SMOTE algoritması, sınıf ağırlıklandırma ve karar eşiği optimizasyonu gibi farklı dengesizlik giderme stratejilerini bütünleşik biçimde uygulamak ve bu stratejilerin model performansına etkisini analiz etmek.

- Optuna çerçevesi aracılığıyla Bayes tabanlı hiperparametre optimizasyonu uygulayarak, varsayılan parametrelere kıyasla elde edilen iyileşme oranlarını sayısal olarak ortaya koymak.
- Modellerin yalnızca klasik metrikler (Doğruluk, Kesinlik, Anma, F1-skoru, AUC-ROC) üzerinden değil; kaçan dolandırıcılık ve yanlış alarm maliyetlerini de içeren maliyet odaklı bir çerçeveyle değerlendirilmesini sağlamak.
- Elde edilen sonuçları literatürdeki benzer çalışmalarla karşılaştırmak ve önerilen yaklaşımın katkılarını ortaya koymaktır.

Bu çalışmada ayrıca, klasik makine öğrenmesi algoritmalarının (KNN, Naive Bayes, Lojistik regresyon) aynı veri seti ve ön işleme koşulları altında test edilerek derin öğrenme yaklaşımlarıyla karşılaştırılması da gerçekleştirilmiştir. Bu karşılaştırma, derin öğrenme modellerinin neden tercih edildiğini deneysel kanıtlarla desteklemektedir. Buna ek olarak, yapısal verilerde güçlü referans yöntemler olarak kabul edilen ağaç tabanlı topluluk modelleri (Rastgele orman (Random Forest) ve XGBoost) de Optuna ile optimize edilerek karşılaştırmaya dâhil edilmiş; böylece derin öğrenme modellerinin göreli konumu hem zayıf hem de güçlü geleneksel yöntemlere karşı sınanmıştır.

Bu çalışmanın diğer bölümleri aşağıdaki gibi yapılandırılmıştır. İkinci bölümde, finansal dolandırıcılık tespiti alanında yapılmış güncel akademik çalışmalar; klasik makine öğrenmesi yöntemleri, derin öğrenme yaklaşımları, hibrit mimariler ve ileri teknikler bağlamında detaylı şekilde incelenmektedir. Üçüncü bölümde, veri ön işleme yöntemleri, derin öğrenme modellerinin ve klasik makine öğrenmesi algoritmalarının teorik altyapısı, hiperparametre optimizasyonu yaklaşımları ve performans değerlendirme ölçütleri sunulmaktadır. Dördüncü bölümde, kullanılan veri seti tanıtılmakta; ön işleme adımları, model eğitim süreçleri, Optuna optimizasyonu ve elde edilen deneysel sonuçlar detaylı olarak aktarılmaktadır. Beşinci bölümde ise bulgular özetlenmekte, çalışmanın katkıları değerlendirilmekte ve gelecek çalışmalar için öneriler sunulmaktadır.

2. LİTERATÜR ARAŞTIRMASI

Finansal dolandırıcılık, günümüzün dijitalleşen ekonomilerinde hem bireyler hem de kurumlar için ciddi bir tehdit oluşturmaktadır. Artan işlem hacmi ve çeşitlenen saldırı yöntemleri karşısında, geleneksel denetim teknikleri yetersiz kalmakta; bu da otomatik, veri odaklı yaklaşımların önemini artırmaktadır. Bu bağlamda, makine öğrenmesi ve veri madenciliği teknikleri, büyük ve karmaşık veri setlerinden dolandırıcılık örüntülerini tespit etme potansiyeli sunmaktadır. Bu literatür taraması, son yıllarda bu alanda gerçekleştirilen çalışmaları derinlemesine inceleyerek, kullanılan yöntemleri, veri kaynaklarını ve elde edilen bulguları sistematik bir biçimde sunmayı amaçlamaktadır. İncelenen çalışmalar kronolojik sıra ile sunulmuştur.

Fu ve diğ. (2016) kredi kartı sahtekârlığını tespit etmek için derin öğrenmeye dayalı bir yaklaşımla Convolutional Neural Network (CNN) tabanlı bir model önermiştir. Araştırmacılar, karmaşık harcama alışkanlıklarını daha iyi yansıtabilmek için "trading entropy" adını verdikleri yeni bir özellik geliştirmiştir. Bu özellik, kullanıcıların alışveriş davranışlarındaki ani değişimleri tespit ederek dolandırıcılık sinyallerini daha hassas şekilde ortaya koymaktadır. Çalışmada kullanılan veri seti, Çin'deki büyük bir ticari bankadan alınan 260 milyondan fazla işlem içeren gerçek kredi kartı verilerinden oluşmaktadır. Bu işlemlerden yalnızca 4.000 kadarı sahte olarak etiketlenmiştir. Veri dengesizliğini dengelemek için cost-based sampling yöntemiyle sahte işlemler sentetik olarak çoğaltılmıştır. Deneysel sonuçlar, CNN modelinin SVM, Rastgele Orman ve klasik NN'ye göre daha yüksek F1-skoru elde ettiğini göstermektedir. Özellikle trading entropy özelliği eklendiğinde modelin başarımının anlamlı şekilde arttığı raporlanmıştır.

Raghavan ve El Gayar (2019) tarafından gerçekleştirilen çalışmada, kredi kartı dolandırıcılığını tespit etmek amacıyla hem makine öğrenimi hem de derin öğrenme algoritmaları karşılaştırmalı olarak analiz edilmiştir. Araştırmacılar, Avrupa, Avustralya ve Almanya'ya ait üç farklı veri seti üzerinde deneyler gerçekleştirerek, çeşitli modellerin farklı veri büyüklüklerinde ve dengesizlik oranlarında nasıl

performans gösterdiğini incelemiştir. Kullanılan algoritmalar arasında KNN, Rastgele Orman, SVM gibi geleneksel makine öğrenmesi yöntemlerinin yanı sıra Autoencoder, RBM, CNN ve DBN gibi derin öğrenme tekniklerine de yer verilmiştir. Özellikle CNN'in hem doğruluk hem de işlem süresi açısından öne çıktığı; SVM'in ise genel başarı ve düşük maliyetle dikkat çektiği görülmüştür. Toplu modeller bireysel algoritmalara göre daha iyi sonuçlar vermiştir. Çalışmada değerlendirme ölçütü olarak MCC, AUC ve sahtekârlık maliyeti kullanılmıştır.

Sánchez-Aguayo ve diğ. (2021) finansal dolandırıcılığın sadece teknik değil aynı zamanda davranışsal yönünü de inceleyen bütüncül bir literatür taraması sunmaktadır. Çalışma, Fraud Triangle teorisini (ihtiyaç, fırsat, gerekçelendirme) çerçeve olarak benimseyerek bireylerin dolandırıcılık eğilimlerinin veri madenciliği teknikleriyle nasıl modellenebileceğini sistematik biçimde değerlendirmiştir. Karar ağaçları, lojistik regresyon ve kümeleme algoritmaları gibi klasik makine öğrenmesi yöntemlerinin Fraud Triangle çerçevesi ile birlikte kullanılabileceği gösterilmiş; teorik kriminoloji altyapısının teknik tespit modelleriyle entegre edilmesinin daha açıklanabilir ve sektörel olarak yorumlanabilir sonuçlar ürettiği vurgulanmıştır.

Badjatiya ve diğ. (2021) kredi kartı dolandırıcılığının tespiti için kullanılan derin öğrenme yaklaşımlarını sistematik şekilde incelemiştir. Sınıf dengesizliği, veri gizliliği, açıklanabilirlik ve gerçek zamanlı işlem gerekliliği gibi zorlukların altı çizilmiş; CNN, RNN, LSTM, Autoencoder ve GAN gibi popüler modellerin performansları karşılaştırmalı biçimde analiz edilmiştir. Kaggle Avrupa Kredi Kartı Dolandırıcılığı ve IEEE-CIS Fraud Detection Dataset gibi dengesiz veri setleri kullanılmıştır. SMOTE, ADASYN ve GAN tabanlı veri üretimi gibi tekniklerle veri dengeleme yapılmıştır. Hibrit mimariler (CNN+LSTM, AE+RNN) ve ensemble modellerin daha güçlü sonuçlar verdiği ortaya konmuştur. Ayrıca açıklanabilir yapay zekâ (XAI) tekniklerinin finans sektöründe vazgeçilmez olduğu vurgulanmıştır.

Chen ve Lai (2021)'nin çalışmasında, kredi kartı dolandırıcılığına karşı etkili bir savunma mekanizması olarak Derin Konvolüsyonel Sinir Ağı (DCNN) tabanlı bir model geliştirmiştir. Model, çok katmanlı bir CNN mimarisi ile büyük ölçekli işlem verileri üzerinde eğitilmiş ve sahte işlemleri gerçek zamanlı olarak tespit etmeyi hedeflemiştir. Kullanılan veri setinde, 5 milyon işlem arasında yalnızca %0,12 oranında sahte işlem bulunmasına rağmen, model %99'a varan doğruluk oranı ile oldukça başarılı sonuçlar elde etmiştir. Özellikle Adam optimizasyon algoritması,

farklı yöntemlerle (RMSprop, Adagrad) karşılaştırıldığında en iyi performansı göstermiştir.

Shen (2021) kredi kartı dolandırıcılığını tespit etmek amacıyla autoencoder tabanlı derin sinir ağı modeli geliştirmiştir. Model, normal işlemler ve dolandırıcılık işlemleri üzerinde ayrı ayrı eğitilmiş; böylece her iki sınıfın örüntülerini daha iyi öğrenmesi sağlanmıştır. Autoencoder'ın sıkıştırma ve yeniden oluşturma mekanizması, dolandırıcılık sinyallerini belirlemede etkili bir araç olarak kullanılmıştır. Model, Kesinlik, Anma, F1-skoru ve ROC-AUC gibi metriklerle değerlendirilmiş ve klasik makine öğrenimi yöntemlerine göre üstün performans sergilemiştir.

Esenogho ve diğ. (2022) kredi kartı dolandırıcılığı tespiti için sinir ağı ensemble yaklaşımını özellik mühendisliğiyle birleştiren bir model önermiştir. Çalışmada LSTM tabanlı bir sinir ağı temel öğrenici olarak kullanılmış; AdaBoost ile birleştirilerek ensemble yapısı oluşturulmuştur. Veri ön işleme aşamasında SMOTE-ENN hibrit yeniden örnekleme tekniği uygulanmış ve sınıf dengesizliği etkin biçimde giderilmiştir. Önerilen LSTM-AdaBoost ensemble modeli, klasik makine öğrenmesi algoritmalarına ve yalnızca LSTM tabanlı modellere kıyasla daha yüksek MCC ve sensitivite değerleri elde etmiş; özellik mühendisliği ile birleştirilmiş ensemble yaklaşımlarının dolandırıcılık tespitindeki üstünlüğünü ortaya koymuştur.

Ali, M. N. Y. ve diğ. (2022)'nin çalışmasında, dengesiz veri problemine odaklanarak SMOTE yöntemini kullanmış ve farklı derin öğrenme modellerini (CNN, ANN ve LSTM) karşılaştırmıştır. CNN modeli hem doğruluk hem de F1-skoru açısından üstün performans göstermiş; doğruluk %99,97, F1-skoru ise %99,96 olarak raporlanmıştır. CNN modelinin karmaşık desenleri yüksek başarıyla öğrenebilmesi ve özellikle SMOTE ile desteklendiğinde klasik modellere göre çok daha iyi sonuçlar vermesi çalışmanın öne çıkan yanıdır.

Almalki ve Masud (2022) tarafından yapılan çalışmada, finansal dolandırıcılık tespitinde yüksek doğruluk oranı elde etmenin yanı sıra modelin kararlarını şeffaf bir şekilde açıklayabilmek hedeflenmiştir. CatBoost, XGBoost ve LightGBM'i bir araya getirerek stacking ensemble tekniğiyle birleşik bir sınıflandırıcı model oluşturmuşlardır. IEEE-CIS Fraud Detection Dataset kullanılarak yapılan deneylerde, geliştirilen model %99'un üzerinde doğruluk ve AUC-ROC = 0,992 sonuçlarına

ulaşmıştır. Modelin karar mekanizmasını analiz etmek amacıyla SHAP, LIME, PDP ve ELI5 gibi açıklayıcı teknikler kullanılmıştır.

Bao ve diğ. (2022) yapay zekâ ve makine öğrenmesinin dolandırıcılık tespiti alanındaki fırsatlarını ve zorluklarını hem uygulamalı hem de akademik bir bakış açısıyla ele almıştır. Yazarlar özellikle finansal tablo sahtekârlığı üzerine yoğunlaşarak, dolandırıcılığın doğası gereği nadir, değişken ve düşmanca olması nedeniyle makine öğrenmesi tekniklerinin geleneksel kullanım alanlarından çok daha zor koşullarda çalıştığını vurgulamıştır. Etiketli veri eksikliği, dengesiz sınıflar ve zaman içinde değişen sahtekârlık motifleri gibi faktörlerin modelleme sürecini zorlaştırdığı belirtilmiştir. Unsupervised learning, semi-supervised, positive-unlabeled learning (PU), transfer learning ve meta-learning gibi alternatif yaklaşımlar önerilmiştir. Değerlendirme metrikleri olarak AUC ve sıralama bazlı NDCG@k gibi metriklerin daha anlamlı olduğu ifade edilmiştir.

Gambo ve diğ. (2022) 1D Convolutional Neural Network (Conv1D) yapısı ve ADASYN örnekleme ile dengelenmiş veri kullanarak kredi kartı dolandırıcılığı tespit modeli sunmuştur. Dengelenmiş veri ile eğitilen model, %99,82 doğruluk, %99,65 kesinlik ve %99,99 anma skorlarına ulaşarak oldukça yüksek bir performans sergilemiştir. Önerilen modelin performansı önceki literatürdeki CNN+SMOTE, CNN+LSTM ve GAN+NN tabanlı modellerle karşılaştırılmış ve bu yeni yaklaşımın daha iyi sonuçlar verdiği gösterilmiştir.

Güven ve Aras (2022) Kentkart akıllı toplu taşıma sistemine ait gerçek bir mobil ödeme (mobile ticketing) uygulaması verisi üzerinde dolandırıcılık tespit modeli geliştirilmiştir. Çalışmada Rastgele Orman, Destek Vektör Makineleri (SVM), Lojistik regresyon, K-En yakın komşu (KNN) ve Naive Bayes algoritmaları karşılaştırılmış; tek tek sınıflandırıcıların çıktıları basit topluluk öğrenmesi (ensemble learning) yöntemleriyle birleştirilerek tahmin başarısı artırılmıştır. Dolandırıcı olarak sınıflandırılan kullanıcılar otomatik olarak kara listeye alınmış ve böylece şirketin dolandırıcılık kaynaklı kayıplarının azaltılması hedeflenmiştir. Çalışma, mobil ödeme sistemlerinde dolandırıcılık tespitine yönelik uygulanabilir bir çerçeve sunmaktadır.

Ileberi ve diğ. (2022) kredi kartı dolandırıcılığı tespitinde özellik seçimi için Genetik Algoritma (GA) kullanılan bir makine öğrenmesi çerçevesi sunmuştur. Sınıf

dengesizliği problemini ele almak için SMOTE uygulanmış; Rastgele Orman, Decision Tree, KNN, Logistic Regression ve Naive Bayes algoritmaları karşılaştırılmıştır. GA tabanlı özellik seçimi ile birlikte uygulanan Rastgele Orman modeli, doğruluk ve AUC açısından diğer kombinasyonlardan üstün performans göstermiş; Avrupa Kredi Kartı Veri Seti üzerinde yüksek sınıflandırma başarısı sağlanmıştır.

Ali ve diğ. (2022) 2010–2021 yılları arasında finansal dolandırıcılık tespitinde makine öğrenmesi temelli yaklaşımları inceleyen sistematik bir literatür taraması sunmuştur. Kitchenham metodolojisi kullanılarak 287 makale toplanmış; filtreleme ve kalite değerlendirmesi sonrası 93 makale analize dâhil edilmiştir. En sık kullanılan veri setlerinden biri Kaggle'daki Avrupa kredi kartı işlem verisidir. Makine öğrenmesi teknikleri arasında en çok tercih edilenler sırasıyla Naive Bayes, SVM, ANN, Lojistik regresyon ve Karar Ağaçları olarak belirtilmiştir. Araştırma; finansal dolandırıcılık verilerinin büyük ölçüde dengesiz olması, veri boyutunun sınırlı olması ve kavram kayması (concept drift) problemini önemli metodolojik boşluklar olarak tanımlamaktadır.

Bu alanda, Yang ve diğ. (2022) finansal dolandırıcılığı tespit etmede derin öğrenme algoritmalarının performansını kapsamlı bir şekilde analiz etmiştir. CNN, LSTM ve Autoencoder gibi modeller kullanılarak farklı örüntü yapılarının öğrenilmesi hedeflenmiştir. Veriler genellikle dengesiz yapıdadır; bu nedenle Random Oversampling ve SMOTE teknikleriyle dengeleme yapılmıştır. Derin ağların, karmaşık ilişkileri klasik algoritmalardan daha iyi modelleyebildiği sonucuna ulaşılmıştır. Hibrit mimariler (örneğin CNN+LSTM) ile daha yüksek başarı oranlarına ulaşılmış; derin öğrenme modelleri gerçek zamanlı finansal gözetim sistemlerinde önerilmiştir.

Mizher ve Nassif (2023) kredi kartı dolandırıcılığı için derin CNN tabanlı bir yaklaşım önermiştir. Kaggle Avrupa kredi kartı dolandırıcılığı veri seti üzerinde yapılan deneyde, dengesiz veri yapısı yeniden örnekleme stratejileriyle ele alınmıştır. CNN tabanlı sınıflandırıcı, Rastgele Orman ile karşılaştırıldığında dengesiz veriler üzerinde rekabetçi sonuçlar üretmiş; ileri derecede dengesiz veri seti üzerinde %99,78 doğruluk seviyesine ulaşılmıştır. Çalışma, dengesizliğin uygun yöntemlerle ele alınması durumunda CNN modelleriyle yüksek doğruluk elde edilebileceğini ve

geleneksel sınıflandırıcılarla rekabet edilebilir bir performans sağlanabileceğini ortaya koymaktadır.

Al Balawi ve Aljohani (2023) tarafından yapılan çalışmada, ANN ve CNN modellerini kullanarak dolandırıcılık vakalarının doğrulukla sınıflandırılması hedeflenmiştir. 284.807 işlemden oluşan ve sadece %0,17'si dolandırıcılık içeren veri seti kullanılmış; modeller Google Colab ortamında TensorFlow ve Keras ile eğitilmiştir. F1-skoru skorları açısından ANN %71, pooling'li CNN %81 ve pooling'siz CNN %84 başarı göstermiştir. CNN'in, karmaşık desenleri tanıma konusunda ANN'e göre daha güçlü olduğu ortaya konmuştur.

Alzahrani ve Aljabri (2023) çevrimiçi reklamcılık alanında ciddi bir tehdit hâline gelen tıklama dolandırıcılığı (ad click fraud) tespitine yönelik son on yılda yayımlanmış yapay zekâ tabanlı çalışmaları sistematik biçimde incelemiştir. İncelemede, makine öğrenmesi (ML) ve derin öğrenme (DL) yaklaşımları kategorize edilmiş; modelin eğitiminde kullanılan öznitelikler, kamuya açık veri setleri ve tıklama davranışının dolandırıcılık göstergesi olarak değerlendirilmesinde belirleyici olan zamansal özellikler ayrıntılı biçimde tartışılmıştır. Çalışma, finansal dolandırıcılığa kıyasla görece az incelenen ancak reklamveren ve yayıncılar açısından milyarlarca dolarlık kayba neden olan bu alanda gelecekteki araştırmalar için somut yönlendirmeler sunmaktadır.

Bir başka çalışmada Aurna ve diğ. (2023) kredi kartı dolandırıcılığının tespitinde hem veri gizliliğini koruyan hem de yüksek doğruluk sağlayan bir sistem geliştirmiştir. Çalışmanın temel yaklaşımı, verileri merkezi bir sunucuya göndermeden cihazlar üzerinde yerel olarak işleyen Federated Learning yöntemidir. SMOTE ile desteklenen CNN modeli %99,51 doğruluk ve %99,11 F1-skoru ile en başarılı sonuçları vermiştir. FL yaklaşımı sayesinde modelin başarısı merkezi yöntemlerle karşılaştırılabilir düzeyde kalmış, buna karşın kullanıcı verilerinin gizliliği korunmuştur.

Barman ve diğ. (2023) veri madenciliği tekniklerinin finansal dolandırıcılık tespitindeki rollerini kapsamlı şekilde ortaya koymuştur. Classification, clustering ve association rule mining gibi temel DM yöntemleri detaylandırılmış ve uygulama örnekleriyle desteklenmiştir. Sektörel bazda karşılaştırmalar yapılarak, bankacılık, sigortacılık ve e-ticaret gibi farklı alanlardaki veri kullanımı analiz edilmiştir.

Özellikle karar ağaçları, k-en yakın komşuluk ve destek vektör makineleri gibi algoritmaların performansları detaylı şekilde ele alınmıştır. Veri ön işleme, eksik veri yönetimi ve özellik seçimi gibi süreçlerin nihai model başarısındaki etkisi vurgulanmıştır.

Berhane (2023) ise kredi kartı dolandırıcılığını tespit etmek için özellik çıkarımı için CNN, sınıflandırma için SVM kullanarak hibrit bir model önermiştir. SMOTE yöntemiyle veri dengelenmiştir. Kaggle Avrupa kredi kartı dolandırıcılığı veri setinde doğruluk oranı %99'un üzerine çıkmış; Kesinlik %96,43, Anma %92,55, F1-skoru %94,45 olarak raporlanmıştır. Hem derin öğrenmenin güçlü özellik çıkarım kapasitesinden hem de SVM'in sınıflandırma başarısından yararlanılmıştır.

Cherif ve diğ. (2023) tarafından gerçekleştirilen çalışmada, 2015–2021 yılları arasında yayımlanan 40 akademik çalışmayı inceleyerek kredi kartı dolandırıcılığı tespitinde kullanılan geleneksel ve derin öğrenme tekniklerinin evrimi kapsamlı biçimde ele alınmıştır. Büyük veri, bulut bilişim, nesnelerin interneti gibi yıkıcı teknolojilerin etkisi, veri dengesizliği, özellik mühendisliği ve gerçek zamanlı tespit konularına odaklanılmıştır. İncelenen modeller arasında SVM, Rastgele Orman, Decision Tree, Naive Bayes, KNN gibi klasik ML yöntemlerinin yanı sıra LSTM, GRU, CNN, Autoencoder gibi derin öğrenme mimarileri de yer almıştır. Gerçek zamanlı tespit ihtiyacı, federated learning gibi gizlilik odaklı yaklaşımlar ve ölçeklenebilir dağıtık sistem tasarımları üzerine de odaklanılmıştır.

Soltani ve diğ. (2023) 20 yıllık süreçte finansal tablo dolandırıcılığı tespitine yönelik çalışmaların evrimini kapsamıştır. Konu modelleme ve bibliyometrik analiz tekniklerini birleştirerek literatürdeki eğilimleri ve odak noktalarını belirlemiştir. Denetimli öğrenme teknikleri ile metin madenciliği yaklaşımları birlikte incelenmiş; SEC gibi kamuya açık finansal veri setleriyle yapılan uygulamalar örnek olarak sunulmuştur. Yöntem seçiminin tespit başarısı üzerinde belirleyici olduğu vurgulanmıştır.

Marazqah Btoush ve diğ. (2023) kredi kartı dolandırıcılığına odaklanan siber güvenlik tehditlerini ve bunların makine öğrenmesiyle tespitini ele almıştır. Logistic regression, XGBoost ve decision tree gibi algoritmalar karşılaştırmalı olarak değerlendirilmiştir. European Credit Card Fraud Dataset yaygın olarak kullanılan kaynaklardan biri olarak belirtilmiştir. Model performansları sadece doğruluk değil,

hatalı pozitiflik açısından da analiz edilmiş; Kesinlik ve Anma metriklerinin kritik önemde olduğu vurgulanmıştır.

Mienye ve Sun (2023) kredi kartı dolandırıcılığı tespiti için derin öğrenme tabanlı bir ensemble yaklaşımı önermiştir. SMOTE-ENN hibrit yeniden örnekleme tekniği ile sınıf dengesizliği giderilmiş; akabinde LSTM, GRU ve MLP gibi derin öğrenme bileşenleri birleştirilerek bir ensemble model oluşturulmuştur. Avrupa Kredi Kartı veri seti üzerinde yapılan deneylerde önerilen ensemble model, %99'un üzerinde AUC değerine ulaşmış; tekil derin öğrenme modellerine kıyasla daha yüksek dengeli doğruluk ve MCC değerleri elde etmiştir. Çalışma, derin öğrenme modellerinin ensemble stratejileriyle birleştirilmesinin, ileri düzey dengesizlik altında dahi etkin bir dolandırıcılık tespit performansı sağladığını ortaya koymuştur.

Alarfaj ve Shahzadi (2024) bankacılık dolandırıcılığı tespiti için Grafik Sinir Ağları (GNN) ve Otokodlayıcı (Autoencoder) mimarilerini birleştiren gerçek zamanlı bir derin öğrenme çerçevesi sunmuştur. GNN bileşeni Lambda mimarisi ile entegre edilerek kullanıcılar ve işlemler arasındaki ilişkisel desenleri modellemiş; Autoencoder bileşeni ise kredi kartı işlemlerindeki anormallikleri yakalamak için kullanılmıştır. İki gerçek bankadan elde edilen vaka çalışmalarıyla doğrulanan model, Kesinlik ve Anma arasında dengeli bir performans göstermiş; bankacılık sistemlerinin gerçek zamanlı dolandırıcılık önleme verimliliğini artırma potansiyelini ortaya koymuştur.

Hernandez Aros ve diğ. (2024) makine öğrenmesi algoritmalarının dolandırıcılık tespitindeki performanslarını karşılaştırmalı olarak ele almıştır. Rastgele Orman, AdaBoost, Gradient Boosting gibi ensemble yöntemlerin başarı oranları analiz edilmiş; ayrıca derin öğrenme modelleriyle kıyaslamalar yapılmıştır. Kaggle bankacılık işlemleri veri seti kullanılmış; model açıklanabilirliği için SHAP ve LIME gibi araçlar kullanılmıştır. Hibrit yöntemlerin daha istikrarlı sonuçlar verdiği sonucuna ulaşılmıştır.

Jain ve Shah (2024) anomali tespiti için CNN mimarisinin başarısını detaylı şekilde incelemiştir. MNIST, KDD Cup 1999, CIFAR-10, Yahoo S5 ve NAB gibi farklı alanlardan gelen veri kümeleriyle deneyler yapılmış; CNN modellerinin yüksek doğrulukla anormal verileri ayırt edebildiği gösterilmiştir. KDD Cup 1999 veri setinde %99,2 doğruluk, MNIST üzerinde %98,75 doğruluk elde edilmiştir. CNN

mimarisinin çeşitli sektörlerde — ağ güvenliği, dolandırıcılık tespiti, üretim hataları, sağlık — başarıyla kullanılabileceği ortaya konmuştur.

Papasavva ve diğ. (2024) çevrimiçi sosyal medya platformları üzerinden dolandırıcılık tespitini amaçlamıştır. BERT ve Transformer tabanlı NLP modelleri, dijital metinlerdeki dolandırıcılık izlerini analiz etmek üzere eğitilmiştir. Twitter ve Reddit gibi platformlardan elde edilen veri kümeleri kullanılmış; sahte hesaplar, aldatıcı içerikler ve dolandırıcılık teşvikleri metin madenciliği ile tespit edilmiştir. Metin içerisindeki semantik ilişkileri yakalamanın dolandırıcılığı önlemede kritik rol oynadığı belirtilmiştir.

Sinap (2024) kredi kartı dolandırıcılığı tespitinde yedi farklı denetimli sınıflandırma algoritmasının (Lojistik regresyon, Karar Ağaçları, Rastgele Orman, XGBoost, Naive Bayes, KNN ve SVM) performansı karşılaştırmalı olarak değerlendirilmiştir. Çalışma özellikle veri setindeki sınıf dengesizliğinin model performansı üzerindeki etkisini ele alarak; Doğruluk, Kesinlik, Anma, F-skoru, AUC ve AUPRC gibi metrikler üzerinden algoritmaların karşılaştırılmasını sağlamıştır. ROC eğrileri ve karmaşıklık matrisleri kullanılarak yapılan değerlendirme sonucunda, dengesiz veri kümelerinde Kesinlik-Anma dengesinin tek bir doğruluk metriğinden çok daha bilgilendirici olduğu ve XGBoost ile Rastgele Orman tabanlı yaklaşımların pratik kullanımda en başarılı sonuçları verdiği ortaya konulmuştur.

Banu ve diğ. (2024) banka ödemelerinde dolandırıcılığı tespit etmek için CNN ve RNN mimarilerinin birleşimini kullanan hibrit bir model önermiştir. Kaggle'daki Banksim veri seti kullanılarak gerçekleştirilen çalışmada, Min-Max normalizasyonuyla veriler ölçeklendirilmiş; CNN ile yapılandırılmamış verilerden özellik çıkarımı yapılmış ve RNN ile bu özelliklerin zaman sırasına göre analizi gerçekleştirilmiştir. Modelin başarısı, %99,2 doğruluk ve %99'un üzerinde kesinlik, anma ve F1-skoru ile desteklenmiştir.

Yu ve diğ. (2024) tarafından gerçekleştirilen çalışmada, kredi kartı dolandırıcılığı tespitinde son dönem Transformer modellerinin yenilikçi uygulamalarına odaklanılmıştır. Veri kaynakları titizlikle işlenerek dengesizlik sorunu giderilmiş ve yüksek korelasyonlu vektörler seçilerek eğitim güçlendirilmiştir. SVM, Rastgele Orman, Neural Network ve Logistic Regression gibi yaygın modellerle Kesinlik, Anma ve F1-skoru metrikleri üzerinden karşılaştırma yapılmıştır. Transformer

modelinin geleneksel uygulamaların ötesinde, dolandırıcılık tespiti gibi niş alanlarda da büyük potansiyel taşıdığı gösterilmiştir.

Chen ve diğ. (2025) derin öğrenme yaklaşımlarının yıllar içerisindeki evrimini detaylı bir biçimde incelemiştir. LSTM, CNN ve Autoencoder gibi modeller, zaman serisi ve sekans veriler üzerinde test edilmiştir. Finansal işlemlerin zamansal desenlerini tanımada LSTM gibi tekrarlayan sinir ağlarının etkinliği vurgulanmıştır. Model eğitim süresi, parametre sayısı ve işlemci gereksinimi gibi konular performans analizine dahil edilmiş; DL modellerinin overfitting'e karşı aldığı önlemler seyreltme (dropout) ve yığın normalizasyonu (batch normalization) teknikleriyle açıklanmıştır. Hibrit sistemlerin daha verimli olabileceği önerilmiştir.

Karim ve diğ. (2025) tarafından yapılan çalışmada, kredi kartı dolandırıcılık tespitinde hem yüksek doğruluk hem de açıklanabilirlik sağlamayı amaçlayan bir ensemble öğrenme yaklaşımı önerilmiştir. Gelişmiş veri dengeleme ve özellik seçimi teknikleri kullanılarak, farklı sınıflandırıcıların bir araya getirildiği bir model geliştirilmiştir. Modelin değerlendirilmesinde kesinlik, anma, F1-skoru ve AUC gibi metrikler kullanılmış; ensemble yaklaşımının tekil modellerden daha üstün performans gösterdiği ortaya konmuştur.

Zhou ve diğ. (2025) kredi kartı dolandırıcılığını tespit etmek amacıyla RNN ve LSTM mimarilerini birleştiren hibrit bir derin öğrenme modeli geliştirmiştir. Model, CNN ve GRU bileşenlerine benzer şekilde hem mekânsal hem de zamansal örüntüleri yakalama kapasitesine sahiptir. Kaggle kredi kartı dolandırıcılığı veri seti üzerinde çeşitli dengeleme teknikleri uygulanmış; Doğruluk, Kesinlik, Anma, F1-skoru ve ROC-AUC gibi yaygın ölçütlerle değerlendirilmiş ve klasik modellerin çoğuna kıyasla daha başarılı olduğu gösterilmiştir.

Bu alanda, Alamri ve Ykhlef (2026) kredi kartı dolandırıcılık tespitinde anomali tespiti ve optimize edilmiş sınıflandırmayı birleştiren yeni bir hibrit makine öğrenmesi modeli sunmuştur. One-Class SVM (OCSVM) ile anomali tespiti, Rastgele Orman (RF) ile sınıflandırma ve Genetik Algoritma (GA) ile hiperparametre optimizasyonu entegre edilmiştir. PaySim veri seti üzerinde değerlendirilen model, F1-skoru, doğruluk, kesinlik, anma ve AUC-ROC gibi metriklerle temel yöntemlere kıyasla üstün sonuçlar elde etmiştir.

Yukarıda incelenen çalışmalar; kullanılan metodoloji, veri seti, performans ölçütleri ve ilgili teknik özellikler açısından Çizelge 2.1'de özetlenmiştir. Bu özet, literatürdeki yaklaşımların çeşitliliğini ve bu çalışmada önerilen yöntemin literatürdeki konumunu netleştirmek amacıyla sunulmuştur.

Çizelge 2.1 : Literatürdeki Çalışmaların Karşılaştırılması.

Çalışma	Metodoloji	Veri Seti	Performans Ölçütü
Fu ve diğ. (2016)	CNN + Trading Entropy	Çin Bankası (260M+)	F1-skoru
Raghavan ve El Gayar (2019)	KNN, RF, SVM, AE, CNN	3 farklı veri seti	MCC, AUC, Maliyet
Sánchez-Aguayo ve diğ. (2021)	SLR + Fraud Triangle	Literatür	Derleme
Badjatiya ve diğ. (2021)	CNN, RNN, LSTM, AE, GAN	Kaggle, IEEE-CIS	Doğruluk, F1-skoru, AUC
Chen ve Lai (2021)	Derin CNN	Banka (5M)	Doğruluk
Shen (2021)	Autoencoder DNN	Kaggle	Kesinlik, Anma, F1-skoru, AUC
Esenogho ve diğ. (2022)	LSTM + AdaBoost + SMOTE-ENN	Kaggle	MCC, Duyarlılık
Ali, M.N.Y. ve diğ. (2022)	CNN + SMOTE	Kaggle	Doğruluk
Almalki ve Masud (2022)	Stacking Ensemble + XAI	IEEE-CIS	AUC
Bao ve diğ. (2022)	Unsupervised, semi-supervised, transfer learning	Çeşitli	AUC, NDCG@k
Gambo ve diğ. (2022)	1D CNN + ADASYN	Kaggle	Doğruluk
Güven ve Aras (2022)	RF, SVM, LR, KNN, NB + Ensemble	Kentkart Mobile Ticketing	Sınıflandırma ölçütleri
Ileberi ve diğ. (2022)	RF + GA + SMOTE	Kaggle	Doğruluk, AUC
Ali, A. ve diğ. (2022)	SLR (93 makale)	Çeşitli	Derleme
Yang ve diğ. (2022)	CNN, LSTM, AE	Çeşitli	Doğruluk, F1-skoru, AUC
Mizher ve Nassif (2023)	CNN + Resampling	Kaggle	Doğruluk
Al Balawi ve Aljohani (2023)	ANN, CNN	Kaggle	F1-skoru
Alzahrani ve Aljabri (2023)	ML/DL Review	Click fraud literatürü	Derleme
Aurna ve diğ. (2023)	FL + CNN, LSTM	Kaggle	Doğruluk
Barman ve diğ. (2023)	DM Derleme	Çeşitli	Derleme
Berhane (2023)	CNN + SVM + SMOTE	Kaggle	Kesinlik
Cherif ve diğ. (2023)	Sistematik Derleme	40 çalışma	Derleme
Soltani ve diğ. (2023)	Bibliyometrik	Literatür	Derleme

Çizelge 2.1 (devamı) : Literatürdeki Çalışmaların Karşılaştırılması.

Çalışma	Metodoloji	Veri Seti	Performans Ölçütü
Marazqah Btoush ve diğ. (2023)	Sistematik Derleme	181 makale	Derleme
Mienye ve Sun (2023)	LSTM + GRU + MLP Ensemble	Kaggle	AUC
Alarfaj ve Shahzadi (2024)	GNN + Autoencoder	İki banka verisi	Kesinlik, Anma
Hernandez Aros ve diğ. (2024)	RF, AdaBoost + SHAP	Kaggle	Doğruluk, F1-skoru, AUC
Jain ve Shah (2024)	CNN	MNIST, KDD	Doğruluk
Papasavva ve diğ. (2024)	BERT, Transformer	Twitter, Reddit	Kesinlik, Anma, F1-skoru
Sinap (2024)	LR, DT, RF, XGBoost, NB, KNN, SVM	Kaggle	Kesinlik, Anma, F1-skoru, AUC
Banu ve diğ. (2024)	CNN + RNN	Banksim	Doğruluk
Yu ve diğ. (2024)	Transformer	Kaggle	Kesinlik, Anma, F1-skoru, AUC
Chen ve diğ. (2025)	LSTM, CNN, AE (SLR)	Çeşitli	Derleme
Karim ve diğ. (2025)	Explainable Ensemble	Kaggle	Kesinlik, Anma, F1-skoru, AUC
Zhou ve diğ. (2025)	Hibrit RNN-LSTM	Kaggle	Doğruluk, Kesinlik, F1-skoru, AUC
Alamri ve Ykhlef (2026)	OCSVM + RF + GA	PaySim	F1-skoru, Doğruluk, AUC

3. METODOLOJİ

Bu bölümde, tez çalışmasının temel araştırma problemini oluşturan finansal dolandırıcılık tespiti için geliştirilen metodolojik çerçeve, kullanılan veri işleme teknikleri, seçilen derin öğrenme mimarileri ve modelin başarısını değerlendirmek için kullanılan istatistiksel metrikler kapsamlı bir şekilde ele alınmaktadır. Çalışma, ham finansal verilerin işlenmesinden başlayarak, uçtan uca bir derin öğrenme çerçevesinin inşasını ve optimizasyonunu kapsamaktadır.

3.1 Finansal Dolandırıcılık Tespitinde Derin Öğrenme Yaklaşımı

Finansal sistemlerin dijitalleşmesiyle birlikte, işlem hacimlerinin artması ve dolandırıcılık yöntemlerinin karmaşıklaşması, geleneksel kural tabanlı sistemlerin yetersiz kalmasına neden olmuştur. Bu bağlamda, veriden öğrenme yeteneğine sahip yapay zeka modelleri, özellikle Derin Öğrenme (Deep Learning), stratejik bir önem kazanmıştır.

Literatürde derin öğrenme yaklaşımlarının finansal dolandırıcılık tespitindeki başarısı, geleneksel makine öğrenimi yöntemleriyle (örneğin Lojistik regresyon, Karar Ağaçları) sıkça karşılaştırılmıştır. Goodfellow ve diğ. (2016), derin öğrenmenin özellikle "özellik mühendisliği" gerektirmeyen yapısının, büyük veri setlerinde insan müdahalesini azalttığını ve modelin ham veriden karmaşık desenleri kendiliğinden öğrendiğini vurgulamaktadır. Finansal verilerin yüksek boyutlu ve doğrusal olmayan yapısı göz önüne alındığında, Roy ve diğ. (2018) tarafından yapılan çalışmada, derin sinir ağlarının (DNN), kredi kartı dolandırıcılık tespitinde geleneksel yöntemlere göre %90'ın üzerinde doğruluk sağladığı, ancak eğitim süresinin daha uzun olduğu belirtilmiştir.

Buna ek olarak, LeCun ve diğ. (2015), derin öğrenmenin çok katmanlı yapısının, verideki hiyerarşik soyutlamaları öğrenmede kritik rol oynadığını belirtir. Dolandırıcılık tespiti bağlamında bu, ilk katmanların basit işlem tutarlarını öğrenirken, daha derin katmanların kullanıcı davranışlarındaki karmaşık ve zamana

yayılan anormallikleri (örneğin; gece yarısı yapılan ardışık yurt dışı harcamaları) modelleyebilmesi anlamına gelir. Bu tezde kullanılan mimari, bu teorik temeller üzerine inşa edilerek, hem yerel (CNN ile) hem de zamansal (LSTM ile) özelliklerin çıkarılmasını hedeflemektedir.

Önerilen metodoloji dört ana fazdan oluşmaktadır:

- Veri Ön İşleme ve Dengesizlik Yönetimi: Ham verinin temizlenmesi, normalizasyonu ve SMOTE algoritması ile sınıf dengesizliğinin giderilmesi.
- Özellik Çıkarımı (Feature Extraction): Evrişimli Sinir Ağları (CNN) kullanılarak işlem dizileri üzerindeki yerel ve mekansal örüntülerin otomatik olarak çıkarılması.
- Sırasal Öğrenme ve Sınıflandırma: Uzun-Kısa Süreli Bellek (LSTM) ağları ile işlemlerin zamansal bağımlılıklarının modellenmesi.
- Model Optimizasyonu ve Değerlendirme: Hiperparametrelerin ayarlanması ve modelin genelleme yeteneğinin test edilmesi.

3.1.1 Derin öğrenmenin teorik temelleri

Derin öğrenme, insan beynindeki biyolojik nöron ağlarından esinlenerek geliştirilen Yapay Sinir Ağlarının, çok katmanlı ve karmaşık yapılarla (hiyerarşik özellik çıkarımı) güçlendirilmiş halidir. Temel bir yapay sinir ağı, giriş katmanı, gizli katmanlar ve çıkış katmanından oluşur.

Bir nöronun, kendisinden önceki katmandaki nöronlardan aldığı sinyalleri işlemesi matematiksel olarak Eşitlik (3.1) ile ifade edilir:

$$y_j = f(\Sigma\, w_{ij} \cdot x_i + b_j) \tag{3.1}$$

Burada x_i giriş sinyallerini, w_{ij} i. giriş ile j. nöron arasındaki ağırlık katsayısını, b_j nöronun eşik değerini (bias), $f(.)$ aktivasyon fonksiyonunu ve y_j nöronun çıkışını ifade etmektedir. Bu tezde kullanılan derin öğrenme modelleri, bu temel işlemin milyonlarca kez tekrarlandığı ve ağırlıkların "Geri Yayılım Algoritması" (Backpropagation) ile optimize edildiği yapılardır.

3.1.2 Aktivasyon fonksiyonları

Modelin doğrusal olmayan karmaşık finansal örüntüleri öğrenebilmesi için aktivasyon fonksiyonları kritik öneme sahiptir. Bu çalışmada kullanılan temel fonksiyonlar şunlardır:

3.1.2.1 Sigmoid fonksiyonu

Sigmoid fonksiyonu, Eşitlik (3.2) ile tanımlanır:

$$\sigma(x) = 1 / (1 + e^{-x}) \tag{3.2}$$

Sigmoid fonksiyonu, özellikle LSTM kapılarında (Unutma, Giriş, Çıkış) bilginin ne kadarının geçeceğine karar vermek için kullanılır; çıktısı 0 ile 1 arasında bir olasılık değeri üretir.

3.1.2.2 Hiperbolik tanjant (tanh) fonksiyonu

Hiperbolik tanjant fonksiyonu, Eşitlik (3.3) ile tanımlanır:

$$tanh(x) = (e^{x} - e^{-x}) / (e^{x} + e^{-x}) \tag{3.3}$$

Hiperbolik tanjant fonksiyonu, veriyi -1 ile 1 arasına sıkıştırarak ağın negatif değerleri de işleyebilmesini sağlar. LSTM'in hücre durumu güncellemelerinde kullanılmaktadır.

3.1.2.3 Rectified linear unit (ReLU)

ReLU fonksiyonu, Eşitlik (3.4) ile tanımlanır:

$$ReLU(x) = max(0, x) \tag{3.4}$$

ReLU fonksiyonu, CNN katmanlarında negatif değerleri sıfırlayarak işlem hızını artırmakta ve "vanishing gradient" problemini hafifletmektedir.

3.2 Veri Ön İşleme Yöntemleri

Derin öğrenme modellerinin başarısı, doğrudan modele beslenen verinin kalitesine bağlıdır. Finansal işlem verileri genellikle gürültülü, eksik ve farklı ölçeklerde değerler içerdiğinden, model eğitimine geçilmeden önce kapsamlı bir ön işleme süreci uygulanması gerekmektedir. Bu bölümde, finansal veri setlerine uygulanan temel ön işleme yöntemleri teorik olarak ele alınmaktadır.

3.2.1 Veri temizleme ve eksik veri analizi

Veri setinde bulunan hatalı veya eksik gözlemler (Missing Values), modelin öğrenme sürecinde sapmalara (bias) neden olabilmektedir. Bu nedenle model eğitimine geçilmeden önce veri setinin eksik değer açısından incelenmesi gerekmektedir. Eksik değerlerin bulunduğu durumlarda, finansal verilerde sıkça rastlanan aşırı uç değerlere (outliers) karşı ortalamadan daha dirençli olması nedeniyle "Medyan" ile doldurma tercih edilebilecek uygun bir yöntemdir.

3.2.2 Kategorik değişkenlerin kodlanması

Finansal veri setleri genel olarak "İşlem Türü", "Şehir" veya "Cihaz Bilgisi" gibi sayısal olmayan kategorik değişkenler içerebilmektedir. Derin öğrenme modelleri matematiksel işlemler üzerine kurulu olduğundan, böyle durumlarda bu verilerin sayısal formatlara dönüştürülmesi gerekir. Bu amaçla yaygın olarak One-Hot Encoding yöntemi kullanılır; bu yöntemde K adet kategoriye sahip bir değişken, K uzunluğunda ve yalnızca ilgili kategorinin indeksinde 1, diğerlerinde 0 değerini alan bir vektör ile temsil edilir (örneğin [Nakit, EFT, Kredi Kartı] için "EFT" işlemi [0, 1, 0] vektörü ile ifade edilir).

3.2.3 Veri standartlaştırması

Finansal işlemlerdeki "Tutar" (Amount) değişkeni çok geniş bir aralıkta (örneğin 1 € ile 1.000.000 € arasında) değişebilirken, "Saat" gibi değişkenler daha dar bir aralıktadır. Bu ölçek farkı, gradyan tabanlı optimizasyon algoritmalarının yakınsama süresini uzatmakta ve modelin büyük değerli değişkenlere aşırı ağırlık vermesine neden olmaktadır.

Bu sorunu aşmak için yaygın olarak Z-Score standartlaştırma tekniği kullanılmaktadır. Bu dönüşüm, her bir özelliğin ortalamasını 0, standart sapmasını 1 olacak biçimde ölçeklendirmektedir. Min-Max normalizasyonuna kıyasla, Z-Score yöntemi özellikle finansal verilerdeki aykırı değerlere (outliers) karşı daha dirençli sonuçlar üretmekte ve gradyan tabanlı optimizasyon algoritmalarının daha hızlı yakınsamasına olanak tanımaktadır. Standartlaştırma, Eşitlik (3.5) ile ifade edilir:

$$z = (x - \mu) / \sigma \tag{3.5}$$

Burada x orijinal değeri, μ eğitim setindeki özelliğin ortalamasını, σ ise standart sapmasını ve z ölçeklendirilmiş değeri ifade etmektedir.

3.2.4 Sınıf dengesizliğinin giderilmesi (SMOTE Algoritması)

Dolandırıcılık tespiti çalışmalarındaki en büyük zorluklardan biri Sınıf Dengesizliğidir (Class Imbalance). Gerçek hayat verilerinde meşru işlemler (Majority Class) milyonlarca iken, dolandırıcılık işlemleri (Minority Class) sadece binlerle ifade edilmektedir. Bu durum, modelin her işlemi "meşru" olarak tahmin ederek yüksek doğruluk elde etmeye çalışmasına, ancak dolandırıcılıkları kaçırmasına yol açar.

Finansal dolandırıcılık veri setlerinin en belirgin karakteristiği olan bu problem, model performansını doğrudan etkileyen bir faktör olarak literatürde geniş yer bulmuştur. Chawla ve diğ. (2002), azınlık sınıfının (dolandırıcılık) yetersiz temsil edilmesinin, modelin çoğunluk sınıfına aşırı uyum sağlamasına (overfitting) neden olduğunu ve bu durumun "Yanlış Negatif" oranlarını artırdığını ispatlamıştır.

Bu sorunu aşmak için literatürde önerilen yöntemler arasında Rastgele Örnekleme (Random Undersampling) ve Sentetik Örnekleme (Oversampling) öne çıkmaktadır. He ve Garcia (2009), veri seviyesinde yapılan müdahalelerin, algoritma seviyesinde yapılan müdahalelere göre daha genel geçer sonuçlar verdiğini belirtmiştir. SMOTE (Synthetic Minority Over-sampling Technique) algoritması, sadece veriyi kopyalamak yerine, özellik uzayında yeni sentetik veriler türeterek modelin genelleme yeteneğini artırmaktadır. Makki ve diğ. (2019), finansal veri setlerinde SMOTE kullanımının, Kesinlik ve Duyarlılık dengesini (F1-skoru) önemli ölçüde iyileştirdiğini deneysel olarak göstermiştir.

SMOTE algoritmasının matematiksel çalışma prensibi şöyledir:

- Azınlık sınıfından rastgele bir x örneği seçilir.
- Bu örneğin özellik uzayındaki k en yakın komşusu (k-Nearest Neighbors) bulunur.
- Komşulardan biri (x_{nn}) rastgele seçilir.
- İki nokta arasındaki fark vektörü, 0 ile 1 arasında rastgele bir δ sayısı ile çarpılarak orijinal örneğe eklenir: $x_{new} = x + \delta \cdot (x_{nn} - x)$

Bu işlem sayesinde, dolandırıcılık sınıfının karar sınırları (decision boundaries) genişletilerek modelin bu sınıfı öğrenme kapasitesi artırılabilmektedir.

3.3 Evrişimli Sinir Ağları (Convolutional Neural Networks)

Bu çalışmada, finansal işlemlerin hem mekansal (spatial) hem de zamansal (temporal) özelliklerini yakalayabilmek amacıyla hibrit ve katmanlı derin öğrenme mimarileri tercih edilmiştir. Aşağıda, kullanılan modellerin matematiksel temelleri detaylandırılmıştır.

3.3.1 Evrişimli sinir ağları (Convolutional Neural Networks - CNN)

CNN mimarisi, kökenini LeCun ve diğ. (1998) tarafından geliştirilen ve el yazısı tanıma problemlerinde çığır açan LeNet-5 modelinden almaktadır. Başlangıçta 2 boyutlu görüntü verileri için tasarlanan bu yapı, zamanla 1 boyutlu zaman serisi verilerine de uyarlanmıştır. Kiranyaz ve diğ. (2021), 1D-CNN'lerin, zaman serisi verilerindeki "yerel özellikleri" yakalamada, tam bağlı ağlara göre çok daha az parametre ile daha yüksek performans gösterdiğini ortaya koymuştur.

Finansal dolandırıcılık tespiti özelinde, Fu ve diğ. (2016), işlem dizileri üzerinde CNN kullanarak, kullanıcıların harcama alışkanlıklarındaki ani değişimlerin (yerel sapmaların) tespit edilebileceğini göstermiştir. CNN'in "Konumsal Değişmezlik" (Translation Invariance) özelliği, bir dolandırıcılık paterninin zaman serisinin başında, ortasında veya sonunda gerçekleşmesinden bağımsız olarak model tarafından tanınmasını sağlar.

3.3.1.1 Evrişim operasyonu (convolution operation)

Zaman serisi verisi X ve öğrenilebilir filtre (kernel) K olmak üzere, evrişim işlemi matematiksel olarak Eşitlik (3.6) ile tanımlanır:

$$S_t = (X * K)(t) = \Sigma_i X(t - i) \cdot K(i) \quad (3.6)$$

Burada S_t, işlem dizisi üzerindeki yerel özellikleri (örneğin; kısa süreli, ardışık, küçük tutarlı deneme işlemleri) temsil eden özellik haritasıdır (feature map).

3.3.1.2 Aktivasyon fonksiyonu (ReLU)

Evrişim katmanından çıkan değerler, doğrusal olmayan ilişkileri modellemek ve negatif değerleri filtrelemek için ReLU fonksiyonundan geçirilir: ReLU(x) = max(0, x). ReLU, "Gradient Kaybolması" problemini azalttığı ve işlem yükü hafif olduğu için tercih edilmiştir.

3.3.1.3 Havuzlama (pooling) katmanı

Özellik haritasının boyutunu küçültmek ve en belirgin özellikleri öne çıkarmak için Maksimum Havuzlama (Max Pooling) uygulanır. Belirli bir pencere içindeki en yüksek değer seçilir. Bu işlem, modelin gürültüye karşı dayanıklılığını artırır ve hesaplama maliyetini düşürür.

3.4 Uzun-Kısa Süreli Bellek (Long Short-Term Memory - LSTM)

Finansal işlem verileri, doğası gereği zamana bağlı (time-dependent) ve sırasal (sequential) bir yapı sergilemektedir. Bir işlemin dolandırıcılık olup olmadığının tespiti, sadece o anki işlemin özelliklerine değil, aynı zamanda kullanıcının geçmiş işlem tarihçesindeki davranış kalıplarına da sıkı sıkıya bağlıdır. Bu nedenle, bu çalışmada zaman serisi verilerini modellemek için geliştirilmiş en güçlü derin öğrenme mimarilerinden biri olan LSTM ağları kullanılmıştır.

3.4.1 Tekrarlayan sinir ağları (RNN) ve kısıtları

Geleneksel İleri Beslemeli Sinir Ağları, giriş verilerinin birbirinden bağımsız olduğunu varsayar. Ancak finansal işlemlerde t anındaki bir işlem, t-1, t-2 ve hatta t-n anındaki işlemlerle ilişkilidir. Bu ilişkiyi modellemek için Tekrarlayan Sinir Ağları (RNN) geliştirilmiştir.

RNN'ler, bilgiyi bir döngü içinde taşıyarak "hafıza" özelliği kazanır. Ancak, klasik RNN mimarileri, geriye doğru yayılım sırasında "Gradyan Kaybolması" (Vanishing Gradient) veya "Gradyan Patlaması" (Exploding Gradient) problemlerinden muzdariptir. Bu durum, ağın çok eski (uzun vadeli) bilgileri hatırlamasını ve öğrenmesini engeller. Finansal dolandırıcılıkta, bir kullanıcının aylar önceki profil değişikliğinin bugünkü işlemle ilişkisini kurmak hayati olduğundan, klasik RNN'ler yetersiz kalmaktadır.

3.4.2 LSTM hücre mimarisi ve matematiksel formülasyonu

Hochreiter ve Schmidhuber (1997) tarafından önerilen LSTM mimarisi, klasik RNN'lerin yaşadığı uzun vadeli bağımlılık sorununu çözmek için özel bir hücre durumu (cell state) ve kapı (gate) mekanizması kullanır.

LSTM ağlarının başarısının arkasındaki temel mekanizma, daha sonra Gers ve diğ. (2000) tarafından "Unutma Kapısı" (Forget Gate) eklenerek geliştirilen bu hücre

yapısıdır. Klasik RNN'lerin yaşadığı "Gradyan Kaybolması" problemi, Bengio ve diğ. (1994) tarafından teorik olarak kanıtlanmıştır.

LSTM'in Unutma Kapısı, bu soruna dinamik bir çözüm sunar. Greff ve diğ. (2017), binlerce LSTM varyasyonu üzerinde yaptıkları kapsamlı çalışmada, Unutma Kapısının modelin performansını etkileyen en kritik bileşen olduğunu saptamıştır. Finansal verilerde bu kapı, kullanıcının değişen davranışlarına (konsept kayması - concept drift) adapte olmayı sağlar.

LSTM hücresi, bilginin akışını denetleyen üç ana kapıdan oluşur: Unutma Kapısı (Forget Gate), Giriş Kapısı (Input Gate) ve Çıkış Kapısı (Output Gate). Aşağıda, bir LSTM hücresinin t zaman adımındaki matematiksel çalışma prensibi adım adım türetilmiştir.

3.4.2.1 Unutma kapısı (forget gate)

İlk adım, hücre durumundan hangi bilginin atılacağına karar vermektir. Bu karar, bir önceki gizli durum (h_{t-1}) ve mevcut girdi (x_t) kullanılarak bir sigmoid (σ) katmanı tarafından verilir. Sigmoid fonksiyonu 0 ile 1 arasında bir değer üretir; 0 "tamamen unut", 1 "tamamen hatırla/sakla" anlamına gelir (Eşitlik (3.7)).

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{3.7}$$

3.4.2.2 Giriş kapısı (input gate)

İkinci adım, hücre durumuna hangi yeni bilginin ekleneceğine karar vermektir. Bu süreç iki aşamadan oluşur. İlk olarak, "giriş kapısı katmanı" (i_t) hangi değerlerin güncelleneceğine karar verir. İkinci olarak, bir "tanh" katmanı, hücre durumuna eklenebilecek yeni aday değerler vektörünü ($\tilde{C}_t$) oluşturur (Eşitlik (3.8) ve Eşitlik (3.9)).

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3.8}$$

$$\tilde{C_t} = tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{3.9}$$

3.4.2.3 Hücre durumunun güncellenmesi

Eski hücre durumu (C_{t-1}), yeni hücre durumuna (C_t) güncellenir. Bu işlem, eski durumun unutma faktörü (f_t) ile çarpılması ve giriş kapısı (i_t) ile aday değerlerin ($\tilde{C}_t$) çarpımının eklenmesiyle gerçekleşir. Bu adım, LSTM'in uzun süreli hafıza yeteneğinin merkezidir (Eşitlik (3.10)).

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (3.10)$$

3.4.2.4 Çıkış kapısı (output gate)

Son adım, hücrenin çıktısının (h_t) ne olacağına karar vermektir. Çıktı, hücre durumuna (C_t) dayanır ancak filtrelenmiş bir versiyonudur. Önce bir sigmoid katmanı (o_t) hücre durumunun hangi parçasının çıkışa verileceğini belirler. Sonra hücre durumu tanh fonksiyonundan geçirilir ve çıkış kapısının kararı ile çarpılır (Eşitlik (3.11) ve Eşitlik (3.12)).

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (3.11)$$

$$h_t = o_t \odot tanh(C_t) \quad (3.12)$$

Bu mekanizma sayesinde LSTM, finansal işlem dizileri boyunca dolandırıcılık sinyallerini uzun süre hafızasında tutabilir ve gereksiz gürültüleri filtreleyebilir.

3.5 Model Eğitimi ve Hiperparametre Optimizasyonu

Geliştirilen Derin Öğrenme modelinin başarısı, sadece mimari seçimine değil, aynı zamanda modelin nasıl eğitildiğine ve hiperparametrelerin nasıl ayarlandığına da bağlıdır.

3.5.1 Kayıp fonksiyonu (Loss Function)

Bu tezde ele alınan problem bir "İkili Sınıflandırma" (Binary Classification) problemi olduğundan, modelin tahmin ettiği olasılık dağılımı ile gerçek etiketler arasındaki farkı ölçmek için İkili Çapraz Entropi (Binary Cross-Entropy) kayıp fonksiyonu kullanılmıştır. Amaç, Eşitlik (3.13) ile verilen bu fonksiyonu minimize etmektir.

$$L = -(1/N)\,\Sigma_i\,[\,y_i \cdot log(\hat{y}_i) + (1 - y_i) \cdot log(1 - \hat{y}_i)\,] \quad (3.13)$$

Burada y_i gerçek sınıf etiketi (0: Meşru, 1: Dolandırıcı), $\hat{y}_i$ model tarafından tahmin edilen dolandırıcılık olasılığı, N örneklem sayısıdır.

3.5.2 Optimizasyon algoritması: Adam

Model ağırlıklarının güncellenmesinde, Stokastik Gradyan İnişi (SGD) yönteminin daha gelişmiş ve adaptif bir versiyonu olan Adam (Adaptive Moment Estimation) optimizasyon algoritması (Kingma ve Ba, 2015) tercih edilmiştir. Adam, her bir

parametre için öğrenme oranını ayrı ayrı adapte ederek, özellikle gürültülü ve seyrek gradyanlara sahip finansal verilerde daha hızlı ve kararlı bir yakınsama sağlamaktadır.

Adam algoritması, gradyanların birinci momenti (m_t, ortalama) ve ikinci momenti (v_t, varyans) tahminlerine dayanarak parametre güncellemesi yapar (Eşitlik (3.14)–(3.16)):

$$m_t = \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t \tag{3.14}$$

$$v_t = \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t{}^2 \tag{3.15}$$

$$\theta_t = \theta_{t-1} - \eta \cdot \hat{m}_t / (\sqrt{\hat{v}_t} + \varepsilon) \tag{3.16}$$

Burada η öğrenme oranı, β_1 ve β_2 bozunma oranlarıdır.

3.5.3 Aşırı öğrenmeyi önleme (Regularization)

Modelin eğitim verisini ezberleyip (Overfitting) yeni verilerde başarısız olmasını engellemek için iki temel strateji uygulanmıştır:

- Seyreltme (Dropout): Srivastava ve diğ. (2014) tarafından önerilen bu yöntemde, eğitim sırasında nöronların rastgele bir kısmının (örneğin %20'si) geçici olarak devre dışı bırakılmaktadır. Bu işlem, nöronların birbirine aşırı bağımlı hale gelmesini engeller ve modelin genelleme yeteneğini artırır.
- Erken Durdurma (Early Stopping): Eğitim sırasında doğrulama hatası belirli bir süre boyunca düşmezse, eğitim otomatik olarak durdurulur. Bu, modelin aşırı öğrenmeye başlamadan önceki en iyi halinin korunmasını sağlar.

3.5.4 Bayes tabanlı hiperparametre optimizasyonu (Optuna)

Hiperparametreler, model eğitimi başlamadan önce belirlenen ve eğitim sürecinde sabit kalan parametrelerdir (filtre sayısı, seyreltme oranı, öğrenme hızı, yığın boyutu vb.). Bu parametrelerin optimal değerlerinin bulunması, model başarısını doğrudan etkilemektedir.

Geleneksel yaklaşımlar arasında Grid Search (tüm kombinasyonların denenmesi) ve Random Search (rastgele kombinasyonlar) yer almaktadır. Ancak bu yöntemler, arama uzayı büyüdükçe hesaplama maliyetinin üstel olarak artmasına neden olmaktadır. Bu çalışmada, Bayes Optimizasyonu temelli Optuna çerçevesi tercih

edilmiştir. Optuna, Tree-structured Parzen Estimator (TPE) algoritması ile önceki denemelerden öğrenerek arama uzayını akıllıca daraltmakta ve umut vermeyen denemeler için erken durdurma (pruning) uygulamaktadır.

Akiba ve diğ. (2019), Optuna'nın hem hesaplama açısından verimli hem de derin öğrenme modelleriyle entegrasyonu kolay olduğunu göstermiştir. Bu nedenle çalışmada Optuna; CNN, LSTM ve CNN-LSTM hibrit modellerinin hiperparametre optimizasyonunda kullanılmıştır.

3.6 Performans Değerlendirme Metrikleri

Finansal dolandırıcılık veri setlerinde görülen yüksek Sınıf Dengesizliği nedeniyle, modelin başarısını değerlendirmek için sadece "Doğruluk" (Doğruluk) metriğine güvenmek yanıltıcıdır. Model değerlendirme aşamasında metrik seçimi, problemin doğasına göre yapılmalıdır. Provost ve Fawcett (1997), dengesiz veri setlerinde "Doğruluk" metriğinin yanıltıcı olduğunu, çünkü modelin çoğunluk sınıfını tahmin ederek yapay bir başarı elde edebileceğini vurgulamıştır. Bu duruma literatürde "Doğruluk Paradoksu" (Doğruluk Paradox) adı verilir.

Bu nedenle, bu tezde modelin ayırt etme gücünü ölçmek için ROC Eğrisi Altındaki Alan (AUC-ROC) ve F1-skoru temel alınmıştır. Bradley (1997), AUC metriğinin sınıf dağılımından ve sınıflandırma eşik değerinden bağımsız olması nedeniyle, farklı modellerin karşılaştırılmasında en objektif ölçüt olduğunu belirtmiştir. Ayrıca, Saito ve Rehmsmeier (2015), özellikle pozitif sınıfın (dolandırıcılık) nadir olduğu durumlarda, Kesinlik-Anma eğrisinin ve F1-Skorunun, modelin pratik başarısını yansıtmada ROC eğrisinden daha bilgilendirici olabileceğini savunmuştur.

3.6.1 Karmaşıklık matrisi (Confusion Matrix)

Modelin tahminleri ile gerçek değerlerin karşılaştırıldığı 2x2'lik bir matristir.

- Doğru Pozitif (True Positive - TP): Gerçekte dolandırıcı olan işlemin, model tarafından da "Dolandırıcı" olarak tespit edilmesi (Arzulanan durum).
- Doğru Negatif (True Negative - TN): Gerçekte meşru olan işlemin, model tarafından "Meşru" olarak tespit edilmesi.
- Yanlış Pozitif (False Positive - FP - Tip 1 Hata): Meşru bir işlemin yanlışlıkla "Dolandırıcı" olarak etiketlenmesi (Müşteri memnuniyetsizliğine yol açar).

- Yanlış Negatif (False Negative - FN - Tip 2 Hata): Dolandırıcı işlemin "Meşru" sanılarak kaçırılması (Finansal zarara yol açar).

3.6.2 Kesinlik (Precision)

Modelin "Dolandırıcı" olarak etiketlediği işlemlerin ne kadarının gerçekten dolandırıcılık olduğunu gösterir. Yanlış alarmların (FP) maliyetini ölçmek için önemlidir (Eşitlik (3.17)).

$$Kesinlik = TP / (TP + FP) \quad (3.17)$$

3.6.3 Anma (Recall / Sensitivity)

Gerçek dolandırıcılık vakalarının ne kadarının sistem tarafından yakalanabildiğini gösterir. Finansal dolandırıcılık tespitinde en kritik metrik budur, çünkü bir dolandırıcılığı kaçırmanın (FN) maliyeti çok yüksektir (Eşitlik (3.18)).

$$Anma = TP / (TP + FN) \quad (3.18)$$

3.6.4 F1-skoru (F1-Score)

Kesinlik ve Duyarlılık metriklerinin harmonik ortalamasıdır. Dengesiz veri setlerinde modelin genel başarısını tek bir skora indirgemek için kullanılır. Hem yanlış alarmları hem de kaçırılan dolandırıcılıkları dengeli bir şekilde cezalandırır (Eşitlik (3.19)).

$$F1\text{-}skoru = 2 \cdot (Kesinlik \cdot Anma) / (Kesinlik + Anma) \quad (3.19)$$

3.6.5 ROC eğrisi ve AUC (Area Under Curve)

ROC (Receiver Operating Characteristic) Eğrisi, farklı eşik değerlerinde (thresholds) modelin "Duyarlılık" (True Positive Rate) ile "Yanlış Pozitif Oranı" (False Positive Rate) arasındaki takas (trade-off) ilişkisini gösterir.

AUC (Eğri Altındaki Alan) değeri, modelin sınıflandırma yeteneğinin genel bir özetidir:

- AUC = 0.5: Rastgele tahmin (Yazı-tura atmakla eşdeğer).
- $0.7 < AUC < 0.8$: Kabul edilebilir ayrım.
- $0.8 \leq AUC < 0.9$: Çok iyi ayrım.
- $0.9 \leq AUC < 1.0$: Mükemmel ayrım.

- AUC = 1.0: Kusursuz sınıflandırma.

3.6.6 Doğruluk (Accuracy)

Doğruluk metriği, modelin doğru sınıflandırdığı örneklerin toplam örnek sayısına oranı olarak tanımlanmaktadır (Eşitlik (3.20)):

$$Doğruluk = (TP + TN) / (TP + TN + FP + FN) \quad (3.20)$$

Ancak, sınıf dengesizliğinin ileri düzeyde olduğu problemlerde (bu çalışmada %0,17 dolandırıcılık oranı) doğruluk metriği yanıltıcı olabilmektedir. Model hiçbir dolandırıcılık vakasını tespit etmese dahi yaklaşık %99,83 doğruluk elde edebilir. Bu nedenle bu çalışmada doğruluk metriği ana değerlendirme ölçütü olarak kullanılmamış; bunun yerine Kesinlik, Anma, F1-skoru ve AUC-ROC gibi sınıf dengesizliğine duyarlı metrikler tercih edilmiştir.

3.7 Model Doğrulama Yöntemi: Hold-out Doğrulama

Modelin genelleme yeteneğini değerlendirmek ve aşırı uyum (overfitting) riskini denetlemek amacıyla bu çalışmada Hold-out Doğrulama (Hold-out Validation) yöntemi tercih edilmiştir. Hold-out yöntemi, hesaplama maliyeti düşük, uygulaması kolay ve büyük veri setlerinde kararlı sonuçlar üreten bir doğrulama tekniğidir.

Bu yöntemde veri seti rastgele iki alt kümeye ayrılır: eğitim seti (training set) ve test seti (test set). Eğitim seti modelin parametrelerinin öğrenilmesi için, test seti ise modelin daha önce görmediği veriler üzerindeki performansının değerlendirilmesi için kullanılır. Bu çalışmada veri seti %80 eğitim ve %20 test olacak biçimde stratified (katmanlı) örnekleme ile bölünmüş; sınıf dağılımının her iki alt kümede de korunması sağlanmıştır. Eğitim seti içinde ayrıca %20'lik bir doğrulama (validation) bölümü ayrılarak, model eğitiminde erken durdurma (early stopping) için kullanılmıştır. Bu yapı, sınıf dengesizliğinin yoğun olduğu büyük ölçekli veri setlerinde, K-Katlı Çapraz Doğrulamaya kıyasla daha uygun hesaplama maliyeti sunmaktadır.

3.8 Klasik Makine Öğrenmesi Algoritmaları

Bu çalışmada, derin öğrenme modellerinin üstünlüğünü deneysel olarak gösterebilmek amacıyla, üç klasik makine öğrenmesi algoritması da referans

(baseline) model olarak değerlendirilmiştir. Bu algoritmalar, literatürde dolandırıcılık tespiti alanında yaygın olarak kullanılan temel yöntemlerdir.

3.8.1 K-En yakın komşu (KNN)

K-En Yakın Komşu algoritması, örnek tabanlı (instance-based) öğrenme yöntemlerinden biridir. Sınıflandırma aşamasında, test örneğinin özellik uzayındaki en yakın k komşusu belirlenir ve çoğunluk oylaması ile sınıf ataması yapılır. İki örnek arasındaki uzaklık Öklid mesafesi ile hesaplanmaktadır (Eşitlik (3.21)):

$$d(x, y) = \sqrt{(\Sigma_i (x_i - y_i)^2)} \quad (3.21)$$

Burada x ve y iki örneğin özellik vektörlerini, i ise özellik indeksini temsil etmektedir. KNN'in avantajları basitlik ve parametresiz doğasıdır; dezavantajları ise yüksek boyutlu veride performans kaybı ve hesaplama maliyetinin yüksekliğidir.

3.8.2 Naive Bayes

Naive Bayes sınıflandırıcısı, Bayes teoremine dayanan olasılıksal bir modeldir. Özellikler arasında koşullu bağımsızlık varsayımı yaparak posterior olasılıkları hesaplar ve en yüksek olasılıklı sınıfı atar. Sınıflandırma kararı Eşitlik (3.22)'de verilen Bayes formülüne dayanmaktadır:

$$P(C \mid X) = (P(X \mid C) \cdot P(C)) / P(X) \quad (3.22)$$

Burada $P(C|X)$ posterior olasılığı, $P(X|C)$ olabilirliği (likelihood), $P(C)$ sınıfın önsel (prior) olasılığını ve $P(X)$ kanıt (evidence) değerini ifade etmektedir. Koşullu bağımsızlık varsayımı altında olabilirlik Eşitlik (3.23) ile ayrıştırılır:

$$P(X \mid C) = \Pi_i P(x_i \mid C) \quad (3.23)$$

Gaussian Naive Bayes varyantında, sürekli değişkenlerin normal dağılıma uyduğu varsayılmaktadır. Naive Bayes'in avantajı hızlı eğitim süresi ve küçük veri setlerinde iyi performansıdır; ancak özellikler arası bağımsızlık varsayımı gerçek dünya verilerinde nadiren sağlanmaktadır.

3.8.3 Lojistik regresyon

Lojistik regresyon, ikili sınıflandırma problemleri için yaygın olarak kullanılan doğrusal bir modeldir. Modelin çıktısı, sigmoid fonksiyonu aracılığıyla 0 ile 1 arasında bir olasılık değerine dönüştürülmektedir (Eşitlik (3.24)):

$$P(y = 1 \mid x) = \sigma(w \cdot x + b) = 1 / (1 + e^{\wedge}(-(w \cdot x + b))) \quad (3.24)$$

Burada w ağırlık vektörünü, b bias terimini, x giriş özellik vektörünü ve σ sigmoid fonksiyonunu temsil etmektedir. Tahmin edilen olasılıklar, bir karar eşiği (varsayılan 0,5) ile ikili çıktıya dönüştürülmektedir. Modelin eğitimi Binary Cross-Entropy kayıp fonksiyonunun minimize edilmesiyle gerçekleştirilmektedir.

3.8.4 Rastgele orman (Random Forest)

Rastgele Orman, Breiman (2001) tarafından önerilen ve çok sayıda karar ağacının torbalama (bagging) yöntemiyle birleştirilmesine dayanan bir topluluk öğrenme algoritmasıdır. Her ağaç, eğitim verisinden yerine koymalı örnekleme (bootstrap) ile oluşturulan farklı bir alt küme üzerinde eğitilir ve her düğümde bölünme, rastgele seçilen bir özellik alt kümesi içinden yapılır. Nihai sınıflandırma kararı, ağaçların çoğunluk oylamasıyla verilir. Bu iki kademeli rastgelelik, ağaçlar arasındaki korelasyonu azaltarak modelin varyansını düşürür ve tekil karar ağaçlarına kıyasla aşırı öğrenmeye karşı belirgin bir direnç sağlar.

Rastgele Orman, özellik ölçeklerine duyarsız olması, gürültülü ve yüksek boyutlu verilerde kararlı sonuçlar üretmesi ve özellik önem düzeylerinin (feature importance) doğal olarak hesaplanabilmesi nedeniyle, dolandırıcılık tespiti gibi dengesiz sınıflandırma problemlerinde güçlü bir referans yöntem olarak yaygın biçimde kullanılmaktadır. Sınıf dengesizliği, sınıf ağırlıklandırma (class_weight) parametresi aracılığıyla doğrudan modele yansıtılabilmektedir.

3.8.5 Aşırı gradyan artırma (XGBoost)

XGBoost, Chen ve Guestrin (2016) tarafından geliştirilen, gradyan artırma (gradient boosting) çerçevesinin ölçeklenebilir ve düzenlileştirilmiş bir uygulamasıdır. Yöntem, karar ağaçlarını ardışık biçimde eğitir; her yeni ağaç, önceki topluluğun kayıp fonksiyonu üzerindeki hatalarını (artıklarını) azaltacak yönde, kayıp fonksiyonunun birinci ve ikinci dereceden gradyan bilgisi kullanılarak inşa edilir. Amaç fonksiyonuna eklenen L1 ve L2 düzenlileştirme terimleri ile ağaç karmaşıklığını cezalandıran bileşenler, modelin aşırı öğrenme eğilimini kontrol altında tutar.

XGBoost; öğrenme hızı (learning rate), ağaç derinliği, alt örnekleme oranı (subsample) ve sınıf dengesizliğini telafi eden scale_pos_weight gibi zengin bir hiperparametre kümesi sunar. Bu yapı, algoritmayı Optuna gibi sistematik optimizasyon çerçeveleriyle uyumlu kılmakta; literatürde yapısal (tablo hâlindeki) veriler üzerinde en güçlü sınıflandırıcılardan biri olarak kabul edilmektedir (Chen ve Guestrin, 2016).

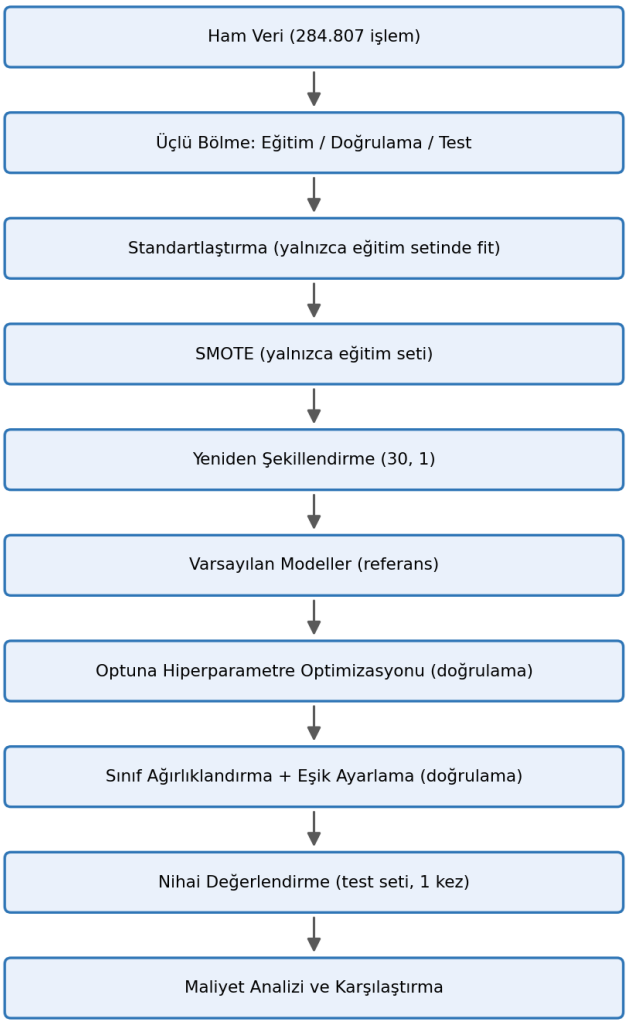
4. UYGULAMA

Bu bölümde, tezin Metodoloji kısmında ele alınan kuramsal çerçevenin gerçek bir veri seti üzerindeki uygulaması detaylı bir biçimde sunulmaktadır. Sırasıyla; kullanılan veri seti tanıtılmakta, uygulanan veri ön işleme adımları açıklanmakta, önerilen derin öğrenme modellerinin (CNN, LSTM ve CNN-LSTM hibrit) varsayılan parametrelerle elde edilen sonuçları paylaşılmakta, Optuna çerçevesi aracılığıyla yürütülen hiperparametre optimizasyon süreci detaylandırılmakta ve son olarak optimize edilmiş modellerin karşılaştırmalı performansı sunulmaktadır.

4.1 Çalışma Akışı

Tez kapsamında uygulanan çalışma akışı, bütüncül bir derin öğrenme çerçevesi (pipeline) olarak tasarlanmıştır. Akış sırasıyla şu aşamalardan oluşmaktadır: ham verinin yüklenmesi ve betimsel istatistiklerle incelenmesi; eksik verilerin yönetilmesi ve aykırı değer analizi; verinin eğitim ve test alt kümelerine ayrılması; standartlaştırma (StandardScaler); SMOTE algoritması ile sentetik örnekleme; modellerin varsayılan parametrelerle eğitilmesi; Optuna ile hiperparametre optimizasyonunun yürütülmesi; eşik ayarlama ile karar eşiğinin optimize edilmesi; sınıf ağırlıklandırma uygulaması; modellerin nihai performansının değerlendirilmesi ve maliyet analizi. Bu sistematik akış, sonuçların yeniden üretilebilirliğini sağlamakta ve yöntemler arasındaki adil karşılaştırmayı mümkün kılmaktadır.

Tez kapsamında izlenen uygulama adımlarının genel görünümü Şekil 4.1'de sunulmaktadır.



Şekil 4.1 : Uygulamanın Akış Diyagramı.

4.2 Veri Seti

Bu çalışmada, dolandırıcılık tespiti araştırmalarında en yaygın kullanılan açık kaynaklı veri setlerinden biri olan Kaggle Avrupa Kredi Kartı İşlemleri Veri Seti (Credit Card Fraud Detection Dataset) tercih edilmiştir. Veri seti, Eylül 2013 tarihinde Avrupa'daki kart sahipleri tarafından gerçekleştirilen 284.807 kredi kartı işlemini içermektedir. Bu işlemlerden yalnızca 492 tanesi dolandırıcılık olarak etiketlenmiş olup, dolandırıcılık oranı %0,17 düzeyindedir. Bu durum, ileri düzey sınıf dengesizliği nedeniyle veri setini özellikle zorlu hâle getirmektedir.

Veri setinde toplam 30 özellik bulunmaktadır. Bunlardan 28 tanesi (V1, V2, ..., V28), gizlilik nedenleriyle PCA dönüşümü uygulanarak elde edilen anonimleştirilmiş

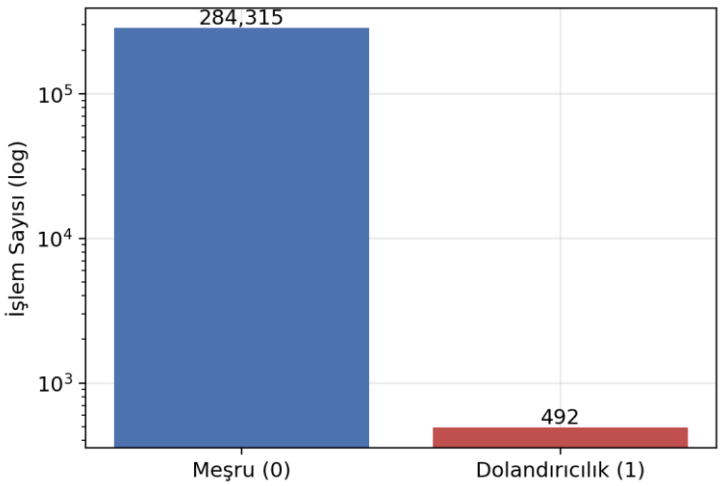
öznitelikleri temsil etmektedir. "Time" özelliği, ilk işlemden itibaren geçen saniye sayısını; "Amount" özelliği ise işlem tutarını ifade etmektedir. Hedef değişken olan "Class" sütunu ikili değerler almaktadır: 0 değeri meşru işlemi, 1 değeri ise dolandırıcılık işlemini temsil etmektedir.

Veri setine ilişkin genel istatistikler Çizelge 4.1'de özetlenmektedir.

Çizelge 4.1 : Veri Seti Genel İstatistikleri.

Özellik	Değer
Toplam İşlem Sayısı	284.807
Meşru İşlem Sayısı	284.315
Dolandırıcılık İşlem Sayısı	492
Dolandırıcılık Oranı	%0,17
Toplam Özellik Sayısı	30 (V1–V28 + Time + Amount)
Hedef Değişken	Class (0: Meşru, 1: Dolandırıcılık)
Eğitim Seti Boyutu	227.845 (394 dolandırıcılık)
Test Seti Boyutu	56.962 (98 dolandırıcılık)
SMOTE Sonrası Eğitim Boyutu	454.902 (227.451 her sınıf)
Veri Kaynağı	Kaggle Credit Card Fraud Detection
Toplama Dönemi	Eylül 2013, Avrupa

Sınıf etiketlerinin veri setindeki dağılımı Şekil 4.2'de gösterilmektedir.



Şekil 4.2 : Sınıf Dağılımı.

Şekil 4.2'de görüldüğü üzere, veri setindeki sınıf dağılımı son derece dengesizdir. Bu durum, klasik sınıflandırma modellerinin çoğunluk sınıfına aşırı uyum sağlamasına neden olmakta; modellerin hiçbir öğrenme yapmadan tüm işlemleri "meşru" olarak tahmin etmesi durumunda dahi yaklaşık %99,83 doğruluk oranına ulaşılmasına yol

açmaktadır. Bu nedenle bu çalışmada, sadece doğruluk metriği yerine Kesinlik, Anma, F1-skoru ve AUC-ROC gibi sınıf dengesizliğine duyarlı metrikler kullanılmıştır.

4.3 Veri Ön İşleme

Veri ön işleme aşaması, modellerin başarısı için kritik öneme sahiptir. Bu çalışmada uygulanan veri ön işleme adımları aşağıdaki alt başlıklarda detaylı şekilde sunulmaktadır. Veri seti öncelikle eksik değer açısından incelenmiş; kullanılan Avrupa Kredi Kartı İşlemleri veri seti tamamı sayısal ve önceden temizlenmiş bir yapıya sahip olduğundan herhangi bir eksik gözlem içermediği görülmüş, dolayısıyla imputasyon (yerine koyma) işlemine gerek duyulmamıştır. Ayrıca veri setindeki tüm öznitelikler (V1–V28, Time ve Amount) sayısal değerlerden oluştuğundan, kategorik kodlama işlemine de gerek duyulmamıştır.

4.3.1 Eğitim ve test bölünmesi

Veri seti, modellerin genelleme yeteneğini değerlendirmek amacıyla %80 eğitim ve %20 test olacak biçimde bölünmüştür. Bölünme işleminde sınıf dağılımının her iki alt kümede de korunmasını sağlamak için stratified (katmanlı) örnekleme kullanılmıştır. Sonuç olarak eğitim setinde 227.845, test setinde ise 56.962 işlem yer almaktadır. Test setindeki dolandırıcılık örneği sayısı 98 olup, bu değer modelin genelleme performansını ölçmek için referans alınmaktadır.

4.3.2 Standartlaştırma

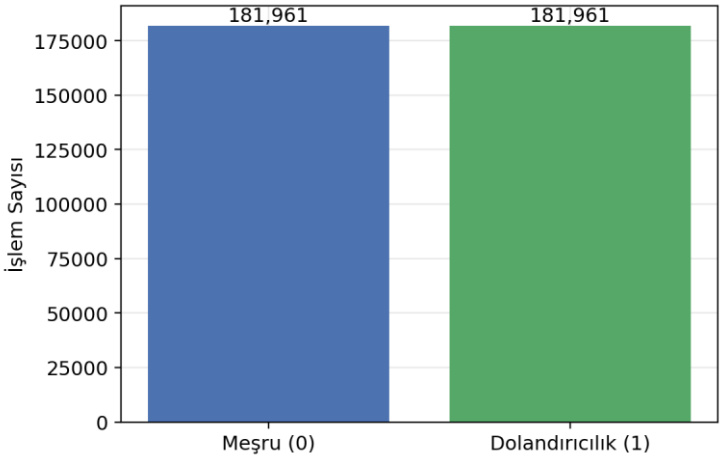
Veri setindeki özelliklerin (özellikle Amount ve Time) farklı ölçeklerde değerler içermesi, gradyan tabanlı optimizasyon algoritmalarının yakınsama süresini olumsuz etkilemektedir. Bu sorunu gidermek amacıyla scikit-learn kütüphanesindeki StandardScaler kullanılarak tüm özellikler standart normal dağılıma (ortalama 0, standart sapma 1) çekilmiştir. Standartlaştırma işlemi, veri sızıntısını önlemek amacıyla yalnızca eğitim seti üzerinde uydurulmuş (fit), test setine ise sadece dönüşüm (transform) uygulanmıştır.

4.3.3 SMOTE ile sentetik örnekleme

Veri setindeki ileri düzey sınıf dengesizliğini gidermek amacıyla SMOTE (Synthetic Minority Over-sampling Technique) algoritması uygulanmıştır. SMOTE, azınlık

sınıfından (dolandırıcılık) rastgele bir örnek seçer, bu örneğin k en yakın komşusunu bulur ve komşu örneklerden biriyle arasında yeni sentetik örnekler üretir. Bu çalışmada k=5 olarak belirlenmiş ve SMOTE yalnızca eğitim setine uygulanmıştır; test seti orijinal dağılımıyla bırakılarak modelin gerçek dünya performansının doğru değerlendirilmesi amaçlanmıştır. SMOTE sonrasında eğitim setindeki sınıflar dengelenerek her iki sınıftan da 227.451 örnek elde edilmiştir.

SMOTE uygulaması sonrasında elde edilen dengeli sınıf dağılımı Şekil 4.3'te gösterilmektedir.



Şekil 4.3 : SMOTE Sonrası Sınıf Dağılımı.

4.3.4 Veri yeniden şekillendirme

CNN ve LSTM mimarileri, üç boyutlu giriş tensörleri (samples, time, features) beklemektedir. Bu nedenle her örnek (30 özellikli vektör), (30, 1) boyutunda yeniden şekillendirilmiştir. Bu yapı, 30 zaman adımında tek bir özellik değeri olan bir sinyal olarak yorumlanmaktadır ve hem 1D-CNN hem de LSTM mimarileri bu girdi formatını işleyebilmektedir.

4.4 Önerilen Modellerin Uygulanması

Bu çalışmada üç farklı derin öğrenme mimarisi karşılaştırılmıştır: tek boyutlu Evrişimli Sinir Ağı (1D-CNN), Uzun-Kısa Süreli Bellek (LSTM) ağı ve CNN-LSTM hibrit modeli. Modellerin tamamı TensorFlow/Keras kütüphanesi kullanılarak Python programlama dilinde uygulanmıştır. Eğitim süreçleri Google Colab ortamında, NVIDIA T4 GPU üzerinde gerçekleştirilmiştir. Tüm modellerde Binary

Cross-Entropy kayıp fonksiyonu, Adam optimizatörü ve EarlyStopping geri çağırma fonksiyonu (callback) kullanılmıştır.

4.4.1 CNN mimarisi

Önerilen 1D-CNN mimarisi; iki adet konvolüsyon-havuzlama bloğu ve bunları takip eden tam bağlı (dense) katmanlardan oluşmaktadır. İlk evrişim katmanında 64 filtre, ikinci evrişim katmanında 128 filtre kullanılmıştır. Her iki katmanda da 3 boyutlu çekirdek (kernel) ve ReLU aktivasyon fonksiyonu tercih edilmiştir. Maksimum havuzlama (max pooling) işlemi 2 boyutunda uygulanmıştır. Düzleştirme (Flatten) işleminin ardından 64 nöronlu tam bağlı katman ve son olarak sigmoid aktivasyonlu çıkış katmanı yer almaktadır. Aşırı öğrenmeyi önlemek için seyreltme oranı 0,2 olarak belirlenmiştir.

4.4.2 LSTM mimarisi

LSTM mimarisi; iki yığılmış (stacked) LSTM katmanından ve bunları takip eden bir tam bağlı katmandan oluşmaktadır. İlk LSTM katmanı 64 birimle ve return_sequences=True parametresiyle yapılandırılmış; böylece zamansal bilginin bir sonraki LSTM katmanına aktarımı sağlanmıştır. İkinci LSTM katmanı 32 birimle yapılandırılmıştır. Her iki LSTM katmanından sonra seyreltme (0,2) ile düzenlileştirme uygulanmıştır. Son katman tek nöronlu, sigmoid aktivasyonlu çıkış katmanıdır.

4.4.3 CNN-LSTM hibrit mimarisi

Hibrit mimari; CNN'in yerel özellik çıkarma ve LSTM'in zamansal modelleme yeteneklerini birleştirmektedir. Mimari öncelikle iki konvolüsyon-havuzlama bloğundan geçer (64 ve 128 filtre); ardından LSTM katmanına (64 birim) iletilir. LSTM katmanını seyreltme (0,2) ve 32 nöronlu tam bağlı katman izlemektedir. Çıkış katmanı, diğer mimarilerle aynı biçimde sigmoid aktivasyonlu tek nörondan oluşmaktadır. Bu yapı, ham işlem sinyalindeki yerel örüntüleri CNN ile çıkarıp, LSTM ile zamansal bağımlılıkları modellemeyi hedeflemektedir.

4.4.4 Varsayılan parametrelerle eğitim sonuçları

Üç model de varsayılan parametrelerle (epochs=20, batch_size=256, learning_rate=0,001) eğitilmiş ve test setindeki performansları değerlendirilmiştir.

Karar eşiği için klasik 0,5 değeri kullanılmıştır. Çizelge 4.2'de varsayılan parametrelerle elde edilen sonuçlar sunulmaktadır.

Çizelge 4.2 : Varsayılan Parametrelerle Model Performans Analizi.

Model	TP	FP	FN	Kesinlik	Anma	F1-skoru	AUC-ROC
CNN	83	37	15	%69,2	%84,7	%76,1	0,9668
LSTM	85	110	13	%43,6	%86,7	%58,0	0,9633
CNN-LSTM	86	69	12	%55,5	%87,8	%68,0	0,9439

Çizelge 4.2'den de görüleceği üzere, varsayılan parametrelerle eğitilen modellerin Anma değerleri yüksektir (%85-88). Kesinlik tarafında ise modeller arasında belirgin farklılıklar bulunmaktadır: CNN modeli %69 ile görece güçlü bir başlangıç performansı sergilerken, LSTM (%44) ve CNN-LSTM (%55) modelleri çok sayıda yanlış alarm üretmektedir. AUC-ROC değerleri ise oldukça yüksek olup (0,94 üzerinde), modellerin sınıflandırma yeteneğinin temel düzeyde başarılı olduğunu göstermektedir. Bu bulgular, özellikle LSTM ve CNN-LSTM modellerinin daha sistematik bir hiperparametre optimizasyon süreciyle iyileştirilebileceğini ortaya koymaktadır.

4.5 Hiperparametre Optimizasyonu: Optuna

Modellerin performansını sistematik biçimde iyileştirmek amacıyla Optuna çerçevesi kullanılarak Bayes tabanlı hiperparametre optimizasyonu uygulanmıştır. Optuna; akıllı arama algoritması olan Tree-structured Parzen Estimator (TPE) ve erken durdurma (pruning) yetenekleriyle, geleneksel grid search ve random search yöntemlerine kıyasla çok daha verimli sonuçlar üretmektedir.

4.5.1 Optimizasyon stratejisi

Bu çalışmada her bir model için 50 deneme (trial) yürütülmüştür. Önceki literatür çalışmalarında F1-skoru veya doğruluk metrikleri hedef alınmasına rağmen, bu tezde Kesinlik metriği hedef olarak belirlenmiştir. Bu seçimin nedeni, finansal kurumların yanlış alarm maliyetlerini düşürme yönündeki operasyonel önceliğidir. Yüksek Kesinlik, modelin "dolandırıcılık" olarak işaretlediği işlemlerin gerçekten dolandırıcılık olma olasılığını artırmaktadır.

Optimizasyon sürecinde, her model için arama uzayı (search space) tanımlanmış; filtre sayısı, LSTM birim sayısı, seyreltme oranı, öğrenme hızı ve yığın boyutu gibi

kritik parametreler için makul aralıklar belirlenmiştir. Çizelge 4.3'te CNN modeli için tanımlanan arama uzayı ve elde edilen optimum değerler sunulmaktadır.

Çizelge 4.3 : CNN Modeli için Optuna Hiperparametre Optimizasyonu.

Hiperparametre	Arama Aralığı	Veri Tipi	Optimum Değer
filters_1	[32, 64, 128, 256]	Kategorik	256
filters_2	[32, 64, 128, 256]	Kategorik	128
dense_units	[32, 64, 128]	Kategorik	128
dropout_rate	[0,1 – 0,5]	Sürekli	0,319
learning_rate	[1e-4 – 1e-2]	Sürekli (log)	0,00023
batch_size	[128, 256, 512]	Kategorik	128
Deneme sayısı	50	Sabit	—
Hedef metrik	Kesinlik (min_recall ≥ 0,75)	Sabit	—

LSTM modeli için yürütülen optimizasyon sürecine ait arama uzayı ve elde edilen en iyi hiperparametre değerleri Çizelge 4.4'te sunulmaktadır.

Çizelge 4.4 : LSTM Modeli için Optuna Hiperparametre Optimizasyonu.

Hiperparametre	Arama Aralığı	Veri Tipi	Optimum Değer
units (LSTM birim)	[32, 64, 128, 256]	Kategorik	128
dense_units	[32, 64, 128]	Kategorik	64
dropout_rate	[0,1 – 0,5]	Sürekli	0,282
learning_rate	[1e-4 – 1e-2]	Sürekli (log)	0,00372
batch_size	[128, 256, 512]	Kategorik	512
Deneme sayısı	50	Sabit	—
Hedef metrik	Kesinlik (min_recall ≥ 0,75)	Sabit	—

CNN-LSTM hibrit modeline ait arama uzayı ve en iyi hiperparametre değerleri ise Çizelge 4.5'te verilmektedir.

Çizelge 4.5 : CNN-LSTM Hibrit Modeli için Optuna Hiperparametre Optimizasyonu.

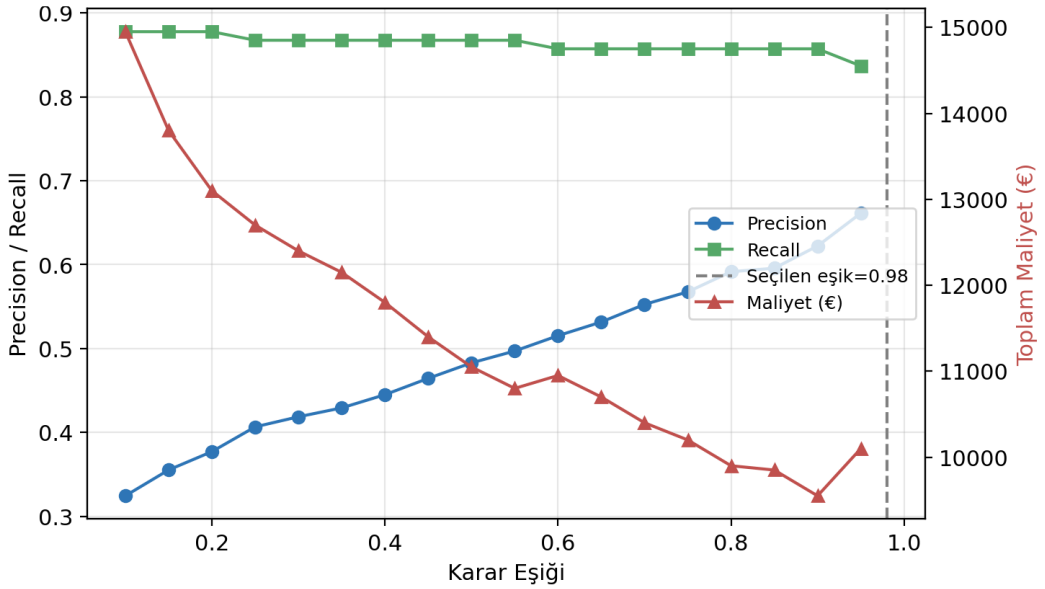
Hiperparametre	Arama Aralığı	Veri Tipi	Optimum Değer
filters (CNN)	[32, 64, 128, 256]	Kategorik	256
lstm_units	[32, 64, 128, 256]	Kategorik	128
dense_units	[32, 64, 128]	Kategorik	32
dropout_rate	[0,1 – 0,5]	Sürekli	0,267
learning_rate	[1e-4 – 1e-2]	Sürekli (log)	0,00051
batch_size	[128, 256, 512]	Kategorik	128
Deneme sayısı	50	Sabit	—
Hedef metrik	Kesinlik (min_recall ≥ 0,75)	Sabit	—

4.5.2 Sınıf ağırlıklandırma (Class Weighting)

Optimizasyon sürecinde, sınıf dengesizliğine ek bir önlem olarak sınıf ağırlıklandırma tekniği uygulanmıştır. Eğitim sırasında dolandırıcılık sınıfına 10 kat daha fazla ağırlık verilerek (class_weight = {0: 1, 1: 10}), modelin azınlık sınıfına daha fazla odaklanması sağlanmıştır. Bu yaklaşım, kayıp fonksiyonunda dolandırıcılık örneklerinin yanlış sınıflandırılmasının daha yüksek bir cezayla karşılanmasını sağlamaktadır.

4.5.3 Karar eşiği optimizasyonu (eşik ayarlama)

Klasik ikili sınıflandırma problemlerinde varsayılan karar eşiği 0,5 olarak alınmaktadır; ancak sınıf dengesizliğinin yoğun olduğu problemlerde bu değer optimal sonuçlar üretmemektedir. Bu çalışmada, her bir model için Kesinlik-Anma eğrisi üzerinde F1-skorunu maksimize eden eşik değeri sistematik biçimde aranmıştır. Optimizasyon sürecinde her üç model için de toplam maliyeti minimize eden eşik değeri 0,98 olarak belirlenmiştir. Bu yaklaşım, modelin yalnızca yüksek olasılıkla dolandırıcılık olarak değerlendirdiği işlemleri alarm olarak işaretlemesini sağlayarak yanlış alarm sayısını önemli ölçüde azaltmıştır. Karar eşiğinin model performansı üzerindeki etkisini bütüncül biçimde görselleştirmek amacıyla CNN modeli için 0,10 ile 0,95 aralığında 0,05'lik adımlarla bir eşik taraması (threshold sweep) gerçekleştirilmiştir. Şekil 4.4'te sunulan analiz, Kesinlik-Anma ödünleşiminin (trade-off) ve buna bağlı toplam finansal maliyetin eşik değerine göre nasıl değiştiğini ortaya koymaktadır. Düşük eşik değerlerinde (0,10-0,40 aralığı) modelin neredeyse her şüpheli işlemi dolandırıcılık olarak işaretlemesi, yüksek Anma sağlamakla birlikte aşırı sayıda yanlış alarma ve dolayısıyla yüksek operasyonel maliyete yol açmaktadır. Eşik yükseldikçe Kesinlik istikrarlı biçimde artmakta, Anma ise görece dar bir bantta (0,83-0,88) kalmakta; toplam maliyet de eşik 0,90 ve üzerine çıktığında en düşük seviyelerine inmektedir. Optuna çerçevesi, bu trade-off eğrisi üzerinde toplam finansal maliyeti minimize eden optimal noktayı (0,98) matematiksel olarak saptamıştır. Bu bulgu, bir işletme yöneticisinin operasyonel risk ile müşteri deneyimi arasında yapması gereken dengeyi sayısal bir çerçeveye oturtmaktadır.



Şekil 4.4 : CNN Modeli için Karar Eşiği Önem Analizi.

4.6 Optimize Edilmiş Modellerin Sonuçları

Optuna tabanlı hiperparametre optimizasyonu, sınıf ağırlıklandırma ve karar eşiği ayarlamasının birlikte uygulanmasının ardından elde edilen sonuçlar Çizelge 4.6'da sunulmaktadır.

Çizelge 4.6 : Optimize Edilmiş Modellerin Karşılaştırmalı Performans Analizi.

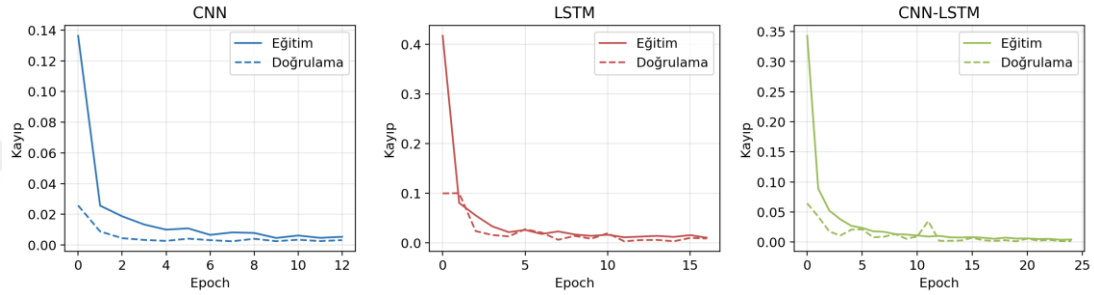
Model	Kesinlik	Anma	F1-skoru	AUC-ROC	Eşik (Threshold)
CNN (Optuna)	%69,2	%82,7	%75,3	0,9575	0,98
LSTM (Optuna)	%60,1	%84,7	%70,3	0,9492	0,98
CNN-LSTM (Optuna)	%76,1	%84,7	%80,2	0,9527	0,98

Çizelge 4.6'dan açıkça görüleceği üzere, Optuna tabanlı optimizasyon sürecinin ardından özellikle LSTM ve CNN-LSTM modellerinin Kesinlik değerlerinde belirgin iyileşmeler elde edilmiştir. LSTM modelinin Kesinlik değeri %44'ten %60'a, CNN-LSTM hibrit modelinin Kesinlik değeri ise %55'ten %76'ya yükselerek yaklaşık 1,4 kat artış göstermiştir. Varsayılan parametrelerle dahi %69 Kesinlik değerine ulaşan CNN modeli ise bu seviyesini korumuştur. F1-skorlarında da en belirgin iyileşme hibrit modelde gözlemlenmiş ve CNN-LSTM modeli %80 F1-skoru ile en yüksek değere ulaşmıştır.

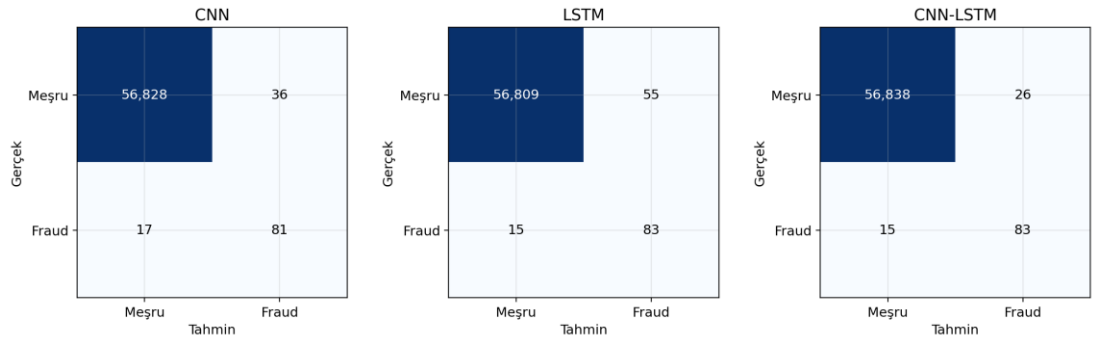
Anma değerlerinde bir miktar düşüş gözlemlenmesi (%85-88'den %83-85'e), beklenen bir sonuçtur; çünkü daha yüksek karar eşiği değerlerinin kullanılması,

modelin yalnızca yüksek olasılıklı vakaları işaretlemesine yol açmaktadır. Buna karşın AUC-ROC değerleri yüksek seviyelerini korumuş, modellerin temel sınıflandırma yeteneğinin etkilenmediğini göstermiştir.

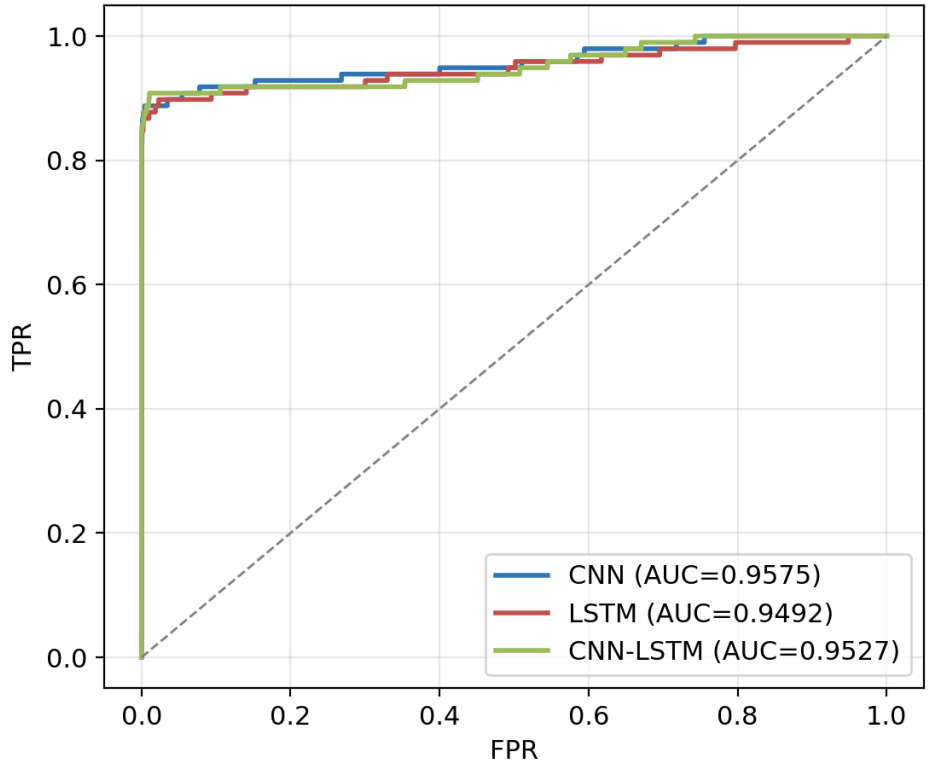
Optimize edilmiş modellere ait eğitim ve doğrulama eğrileri Şekil 4.5'te, karmaşıklık matrislerinin karşılaştırması Şekil 4.6'da, ROC eğrileri Şekil 4.7'de ve Optuna optimizasyonu ile Kesinlik değerlerinde elde edilen iyileşme Şekil 4.8'de sunulmaktadır.



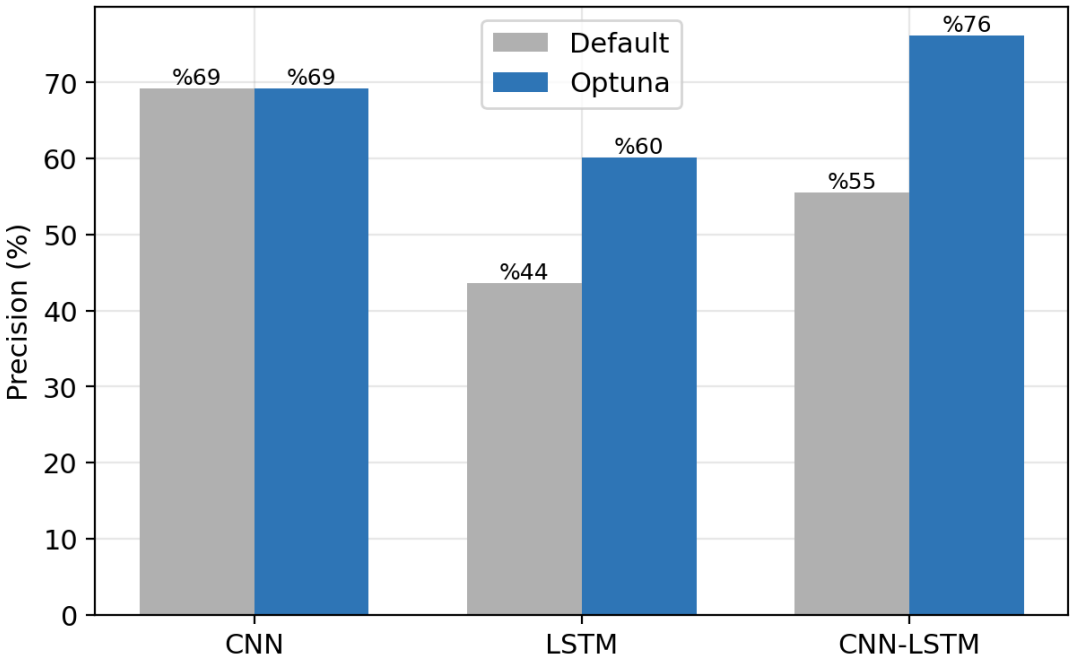
Şekil 4.5 : Eğitim ve Doğrulama Eğrileri.



Şekil 4.6 : Karmaşıklık Matrisi Karşılaştırması.



Şekil 4.7 : ROC Eğrileri Karşılaştırması.



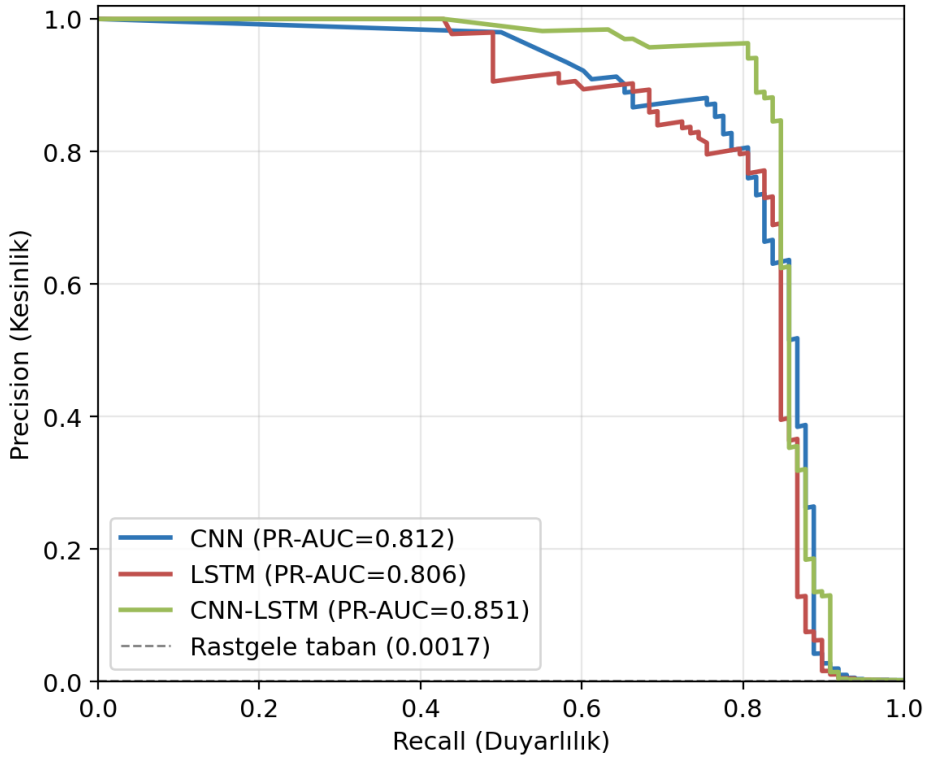
Şekil 4.8 : Parametre Optimizasyonunun Kesinlik Ölçütüne Etkisi.

ROC eğrisinin ileri düzey sınıf dengesizliği altında iyimser bir tablo çizebilmesi nedeniyle (Saito ve Rehmsmeier, 2015), optimize edilmiş modellerin performansı

Kesinlik-Anma (PR) eğrileri üzerinden de değerlendirilmiştir. Şekil 4.9'da sunulan PR eğrileri ve Çizelge 4.7'deki PR-AUC (Ortalama Kesinlik) değerleri, modellerin azınlık sınıfı üzerindeki ayırt etme gücünü çok daha gerçekçi biçimde yansıtmaktadır. CNN-LSTM hibrit modeli 0,8506 ile en yüksek PR-AUC değerini elde etmiş; CNN (0,8117) ve LSTM (0,8060) modelleri birbirine yakın seviyelerde kalmıştır. Rastgele bir sınıflandırıcının PR-AUC tabanının yalnızca 0,0017 (pozitif sınıf oranı) olduğu dikkate alındığında, üç modelin de güçlü bir ayırt etme kabiliyeti sergilediği; bununla birlikte hibrit mimarinin yüksek Anma bölgesinde Kesinlik değerini korumada belirgin bir üstünlük taşıdığı görülmektedir.

Çizelge 4.7 : Optimize Edilmiş Modellerin PR-AUC Değerleri.

Model	PR-AUC (Ortalama Kesinlik)
CNN (Optuna)	0,8117
LSTM (Optuna)	0,8060
CNN-LSTM (Optuna)	0,8506



Şekil 4.9 : Kesinlik-Anma Eğrileri Karşılaştırması.

4.7 Maliyet Analizi

Modellerin pratik performansını daha iyi değerlendirebilmek için, sadece klasik metriklerle yetinmeyip finansal anlamda yorumlanabilir bir maliyet analizi de uygulanmıştır. Bu analizde iki temel hata türü göz önünde bulundurulmuştur: kaçırılan dolandırıcılık (False Negative) ve yanlış alarm (False Positive). Birim maliyet değerleri, dolandırıcılık tespitinde maliyet duyarlı (cost-sensitive) öğrenme yaklaşımlarını ele alan literatür çalışmalarından esinlenerek belirlenmiştir. Bhattacharyya ve diğ. (2011) ile Bahnsen ve diğ. (2014), gerçek bankacılık verileri üzerinde kaçan dolandırıcılığın yanlış alarmdan yaklaşık 10 kat daha maliyetli olduğunu raporlamışlardır. Dal Pozzolo ve diğ. (2018) ise Avrupa kredi kartı verisi üzerinde benzer bir maliyet oranını (FN ≈ 10 × FP) doğrulamıştır. Bu literatür bulgularına paralel olarak, çalışmada kaçan her dolandırıcılık vakasının tahmini maliyeti 500 €, her yanlış alarmın operasyonel maliyeti ise 50 € olarak kabul edilmiştir. Bu oran, gerçek bankacılık operasyonlarındaki ortalama dolandırıcılık tutarı ile yanlış alarm sonucu yapılan müşteri doğrulama çağrısı arasındaki maliyet farkını temsil etmektedir.

Toplam maliyet, denklem 4.1 ile hesaplanmaktadır:

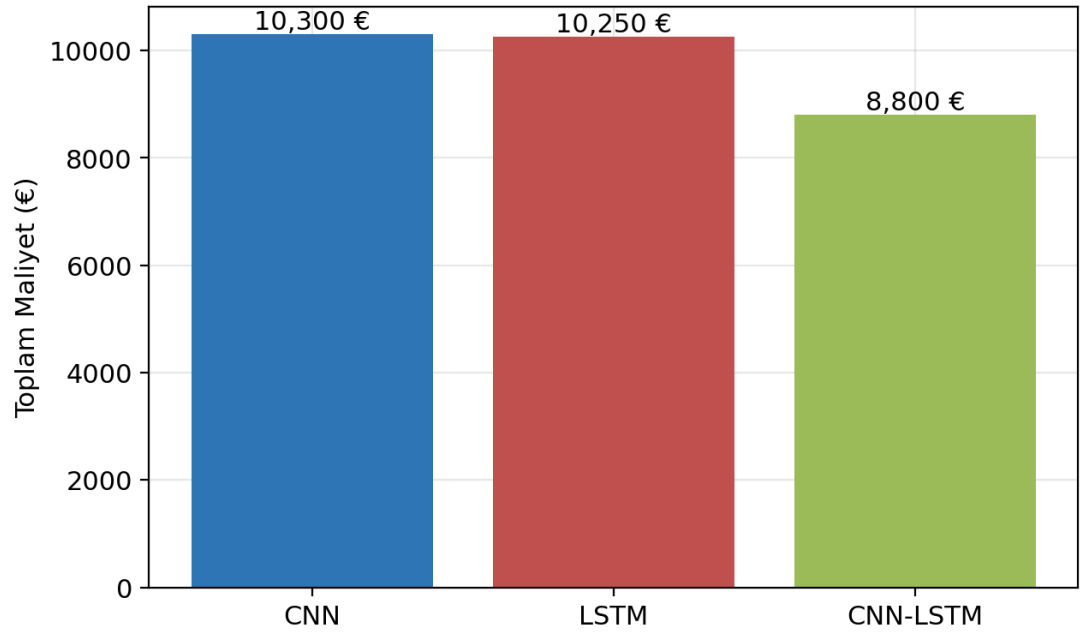
$$Toplam\ Maliyet = (FN \times 500\ €) + (FP \times 50\ €) \qquad (4.1)$$

Bu çerçevede üç modelin hata sayıları ve tahmini toplam maliyetleri Çizelge 4.8'de sunulmaktadır.

Çizelge 4.8 : Modellerin Maliyet Analizi.

Model	FN (Kaçan Fraud)	FP (Yanlış Alarm)	FN Maliyeti	FP Maliyeti	Toplam Maliyet
CNN (Optuna)	17	36	8.500 €	1.800 €	10.300 €
LSTM (Optuna)	15	55	7.500 €	2.750 €	10.250 €
CNN-LSTM (Optuna)	15	26	7.500 €	1.300 €	8.800 €
XGBoost (Optuna)	23	3	11.500 €	150 €	11.650 €
Rastgele Orman (Optuna)	20	16	10.000 €	800 €	10.800 €

Modellerin tahmini toplam maliyetlerinin karşılaştırması Şekil 4.10'da görselleştirilmiştir.



Şekil 4.10 : Maliyet Analizi Bar Grafiği.

Çizelge 4.8'de sunulan maliyet analizi sonuçları, klasik metriklerin ötesinde önemli içgörüler sağlamaktadır. CNN-LSTM hibrit modeli hem en düşük yanlış alarm sayısına (26) hem de en düşük toplam maliyete (8.800 €) sahip olarak finansal açıdan en avantajlı seçenek olmuştur. CNN (10.300 €) ve LSTM (10.250 €) modelleri ise birbirine yakın ve daha yüksek maliyet seviyelerinde kalmıştır. CNN modelinin görece yüksek maliyeti, daha fazla dolandırıcılık vakası kaçırmasından (17 FN); LSTM modelinin maliyeti ise yüksek yanlış alarm sayısından (55 FP) kaynaklanmaktadır. Bu bulgu, hata türlerinin maliyet yapısının model seçiminde belirleyici olduğunu ortaya koymaktadır. Karşılaştırma amacıyla tabloya dâhil edilen ağaç tabanlı topluluk modellerinde ise XGBoost, yalnızca 3 yanlış alarm üretmesine karşın 23 dolandırıcılık vakasını kaçırması nedeniyle 11.650 €, Rastgele Orman ise 10.800 € toplam maliyete ulaşmış; her iki model de bu çerçevede CNN-LSTM hibrit modelinin gerisinde kalmıştır.

Birim maliyet varsayımlarının sonuçlar üzerindeki etkisini sınamak amacıyla, FN ve FP birim maliyetlerinin farklı kombinasyonları üzerinde bir duyarlılık analizi gerçekleştirilmiştir. Çizelge 4.9'da altı farklı maliyet senaryosu altında tüm modellerin toplam maliyetleri sunulmakta; Şekil 4.11 ise FP birim maliyeti 50 € sabit tutulurken FN birim maliyetinin değiştirilmesinin derin öğrenme modellerinin toplam maliyetine etkisini göstermektedir. Sonuçlar, tezde esas alınan 500 €/50 €

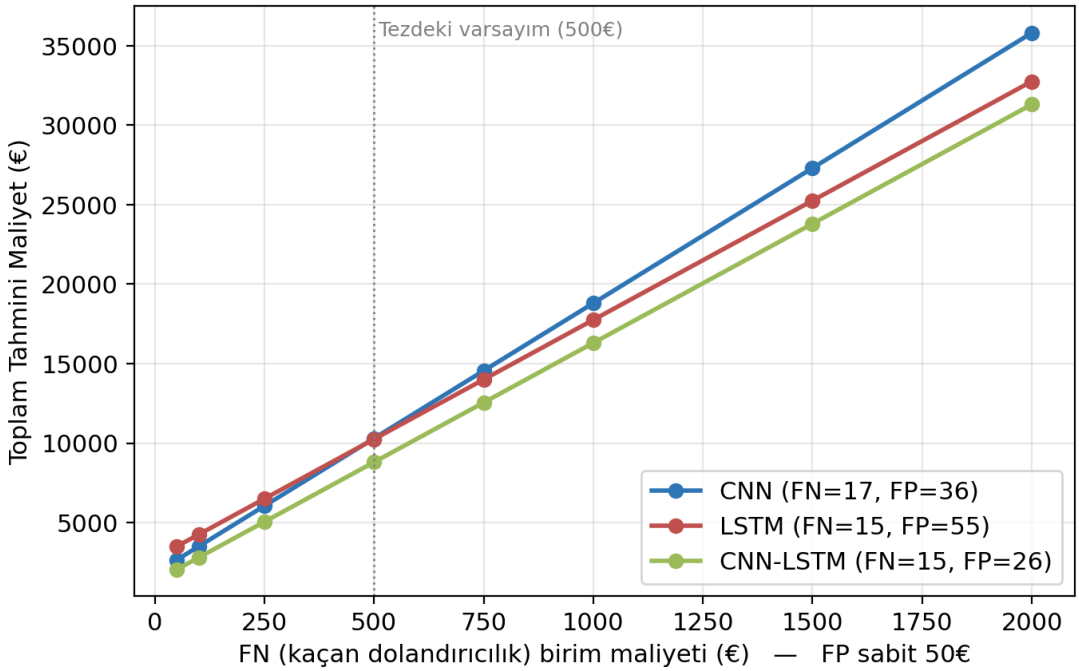
senaryosu dahil olmak üzere yanlış alarm maliyetinin görece ağırlık taşıdığı senaryolarda CNN-LSTM hibrit modelinin en düşük maliyetli seçenek olduğunu doğrulamaktadır. Buna karşılık, kaçan dolandırıcılık maliyetinin baskın hâle geldiği senaryolarda (1.000 €/50 € ve 2.000 €/50 €) düşük FN sayısına sahip KNN modeli toplam maliyet açısından öne geçebilmektedir. Bu bulgu, model seçiminin kurumun kendi maliyet yapısına göre kalibre edilmesi gerektiğini nicel olarak ortaya koymaktadır.

Çizelge 4.9 : Farklı Maliyet Senaryoları Altında Toplam Maliyetler (€).

Senaryo (FN €/FP €)	CNN (Optuna)	LSTM (Optuna)	CNN-LSTM (Optuna)	KNN (k=5)	Naive Bayes	Lojistik regresyon	XGBoost (Optuna)	Rastgele Orman (Optuna)	Kazanan
250/50	6.050 €	6.500 €	5.050 €	8.000 €	76.850 €	77.350 €	5.900 €	5.800 €	CNN-LSTM
500/25	9.400 €	8.875 €	8.150 €	8.125 €	42.925 €	42.425 €	11.575 €	10.400 €	KNN
500/50	10.300 €	10.250 €	8.800 €	10.750 €	79.850 €	79.850 €	11.650 €	10.800 €	CNN-LSTM
500/100	12.100 €	13.000 €	10.100 €	16.000 €	153.700 €	154.700 €	11.800 €	11.600 €	CNN-LSTM
1000/50	18.800 €	17.750 €	16.300 €	16.250 €	85.850 €	84.850 €	23.150 €	20.800 €	KNN
2000/50	35.800 €	32.750 €	31.300 €	27.250 €	97.850 €	94.850 €	46.150 €	40.800 €	KNN

Ağaç tabanlı topluluk modelleri de aynı senaryolar altında değerlendirilmiştir. XGBoost ve Rastgele Orman, çok düşük yanlış alarm sayıları sayesinde yanlış pozitif maliyetinin ağırlık kazandığı senaryolarda avantaj sağlasa da, görece yüksek yanlış negatif sayıları nedeniyle incelenen hiçbir senaryoda en düşük toplam maliyeti üretememiştir.

Farklı maliyet senaryolarında modellerin toplam maliyetlerinin nasıl değiştiği Şekil 4.11'de görselleştirilmiştir.



Şekil 4.11 : Maliyet Duyarlılık Analizi.

4.8 Bulguların Değerlendirilmesi

Uygulama bölümünde elde edilen bulgular dört ana sonuca işaret etmektedir. Birincisi, Optuna tabanlı sistematik hiperparametre optimizasyonu, LSTM (+16 puan) ve CNN-LSTM hibrit (+21 puan) modellerinin Kesinlik metriğinde belirgin iyileşmeler sağlamış; varsayılan parametrelerle dahi güçlü performans gösteren CNN modelinde ise mevcut seviyeyi korumuştur. Bu bulgu, derin öğrenme modellerinin başarısının yalnızca mimari seçimine değil, sistematik optimizasyona da büyük ölçüde bağlı olduğunu doğrulamaktadır.

İkincisi, sınıf ağırlıklandırma ({0:1, 1:10}) ve eşik ayarlamanın birlikte uygulanması, sentetik örnekleme (SMOTE) tek başına uygulandığında elde edilemeyecek başarıların önünü açmıştır. Bu üç tekniğin birlikte kullanılması, sınıf dengesizliğine yönelik bütünleşik bir çerçeve oluşturmaktadır.

Üçüncüsü, klasik metriklerin (Kesinlik, Anma, F1-skoru, AUC-ROC) yanı sıra maliyet odaklı değerlendirme, modellerin pratik kullanımına ilişkin daha şeffaf bir görüntü sunmaktadır. CNN-LSTM hibrit modelinin en düşük toplam maliyeti üretmesi, model seçiminde sadece istatistiksel metriklere değil, operasyonel maliyetlere de dikkat edilmesi gerektiğini göstermektedir.

Dördüncüsü, üç modelin karşılaştırması, bütünleşik optimizasyon sonrasında CNN-LSTM hibrit mimarisinin öne çıktığını ortaya koymaktadır. Hibrit model en yüksek Kesinlik (%76), en yüksek F1-skoru (%80), en yüksek PR-AUC (0,8506), en düşük yanlış alarm sayısı (26) ve en düşük toplam maliyet (8.800 €) değerlerine ulaşmıştır. CNN modeli en yüksek AUC-ROC (0,9575) ile dikkat çekerken, LSTM modeli zamansal bağımlılıkları yakalama yeteneği ile diğer iki mimariyi tamamlayıcı bir rol oynamıştır. Bununla birlikte maliyet senaryosu analizi (Çizelge 4.9), kaçan dolandırıcılık maliyetinin baskın hâle geldiği koşullarda model sıralamasının değişebileceğini göstermektedir.

4.9 Klasik Makine Öğrenmesi Modelleri ile Karşılaştırma

Derin öğrenme modellerinin üstünlüğünü deneysel olarak ortaya koymak amacıyla, aynı veri seti ve ön işleme koşulları altında üç klasik makine öğrenmesi algoritması da değerlendirilmiştir: K-En Yakın Komşu (KNN, k=5), Gaussian Naive Bayes ve Lojistik regresyon (max_iter=1000). Klasik makine öğrenmesi modelleri, bu çalışmada derin öğrenme mimarilerinin temel performans alt sınırını (baseline) belirlemek amacıyla varsayılan parametreleriyle doğrusal bir referans noktası olarak konumlandırılmıştır. Çalışmanın odak noktası, derin öğrenme mimarilerinin hiperparametre uzayının sistematik biçimde optimize edilmesi olduğundan, klasik modeller için ayrı bir hiperparametre arama süreci yürütülmemiştir. Bu yöntemsel tercih, hiperparametre optimizasyonunun yalnızca derin öğrenme modellerinde sağladığı katkıyı izole biçimde gözlemleyebilmek ve karşılaştırma çerçevesini bilinçli olarak "optimize edilmemiş klasik ML vs. optimize edilmiş derin öğrenme" ekseninde kurabilmek için yapılmıştır. Klasik modellere de Optuna veya GridSearch uygulanması, ileri çalışma kapsamında ele alınabilecek bir genişletme alanı olarak öngörülmektedir. Bu modeller, SMOTE ile dengelenmiş eğitim seti üzerinde eğitilmiş ve orijinal dengesiz test seti üzerinde değerlendirilmiştir.

Klasik makine öğrenmesi modellerinin test setindeki performansları Çizelge 4.10'da sunulmaktadır.

Çizelge 4.10 : Klasik Makine Öğrenmesi Modellerinin Performans Analizi.

Model	TP	FP	FN	Kesinlik	Anma	F1-skoru	AUC-ROC
KNN (k=5)	87	105	11	%45,3	%88,8	%60,0	0,9485
Naive Bayes	86	1.477	12	%5,5	%87,8	%10,4	0,9645
Lojistik regresyon	88	1.497	10	%5,6	%89,8	%10,5	0,9740

Çizelge 4.10'da sunulan sonuçlar, klasik makine öğrenmesi algoritmalarının bu problem üzerindeki ciddi sınırlılıklarını ortaya koymaktadır. Naive Bayes ve Lojistik regresyon modellerinin Kesinlik değerleri yalnızca %6 düzeyinde kalmıştır; bu, modelin dolandırıcılık olarak işaretlediği her 100 işlemin 94'ünün yanlış alarm olduğu anlamına gelmektedir. Pratik olarak, bir banka bu modeli kullandığında her gerçek dolandırıcılık vakası için yaklaşık 17 yanlış alarm üreteceğinden, müşteri deneyimi ve operasyonel maliyetler açısından sürdürülebilir bir çözüm sunamamaktadır.

KNN algoritması %45 Kesinlik ile klasik modeller arasında en iyi sonucu vermiş; ancak bu değer dahi pratik uygulamalar için yetersizdir. Dikkate değer biçimde, klasik modellerin AUC-ROC değerleri (0,95-0,97) yüksek görünmektedir; ancak %0,17 dolandırıcılık oranı gibi ileri düzey dengesizlikte AUC-ROC yanıltıcı olabilmektedir. Bu nedenle Kesinlik ve F1-skoru gibi sınıf dengesizliğine duyarlı metrikler, bu problem için çok daha bilgilendirici ölçütlerdir.

4.10 Ağaç Tabanlı Topluluk Modelleri ile Karşılaştırma

Klasik makine öğrenmesi modellerinin (KNN, Naive Bayes, Lojistik regresyon) yetersiz kaldığı gösterildikten sonra, derin öğrenme mimarilerinin gerçekten güçlü bir referans karşısında değerlendirilebilmesi amacıyla, dolandırıcılık tespiti literatüründe en rekabetçi geleneksel yaklaşımlar olarak öne çıkan iki ağaç tabanlı topluluk (tree ensemble) modeli de analize dâhil edilmiştir: Rastgele Orman ve XGBoost (Extreme Gradient Boosting). Bölüm 4.9'daki klasik modellerden farklı olarak, bu iki model derin öğrenme mimarileriyle birebir aynı optimizasyon protokolü altında değerlendirilmiştir. Böylece karşılaştırma, "optimize edilmemiş klasik ML" ile "optimize edilmiş derin öğrenme" arasına, güçlü ve optimize edilmiş bir ara katman (ağaç toplulukları) yerleştirilerek daha adil ve kapsamlı bir zemine oturtulmuştur.

Ağaç tabanlı topluluk modelleri, derin öğrenme modelleriyle tümüyle tutarlı bir deney kurgusu içinde eğitilmiştir. Veri seti önce %80/%20 oranında ve sınıf dağılımı korunacak biçimde (stratified, random_state=42) eğitim ve test kümelerine ayrılmış; eğitim kümesinin %20'si karar eşiği seçimi için doğrulama (validation) kümesi olarak ayrılmıştır. StandardScaler yalnızca eğitim kümesi üzerinde uyarlanmış, SMOTE (k=5) yalnızca eğitim kümesine uygulanarak veri sızıntısı önlenmiştir. Hiperparametre araması, Bölüm 4.5'te açıklanan Optuna tabanlı TPE örnekleyicisi ile gerçekleştirilmiş; XGBoost için 50, Rastgele Orman için 30 deneme (trial) yürütülmüştür. Optimizasyon hedefi, Anma ≥ 0,75 kısıtı altında Kesinlik değerini en yükseğe çıkarmak olarak belirlenmiş; nihai karar eşiği ise doğrulama kümesinde F1-skorunu en yükseğe taşıyacak biçimde ayarlanmıştır. Bu tercih, azınlık sınıfının doğru yakalanması ile yanlış alarm sayısının kontrol altında tutulması arasında dengeli bir çözüm sağlamayı amaçlamaktadır.

Optuna ile optimize edilmiş ağaç tabanlı topluluk modellerinin orijinal (dengesiz) test seti üzerindeki performansları Çizelge 4.11'de sunulmaktadır.

Çizelge 4.11 : Ağaç Tabanlı Topluluk Modellerinin Performans Analizi.

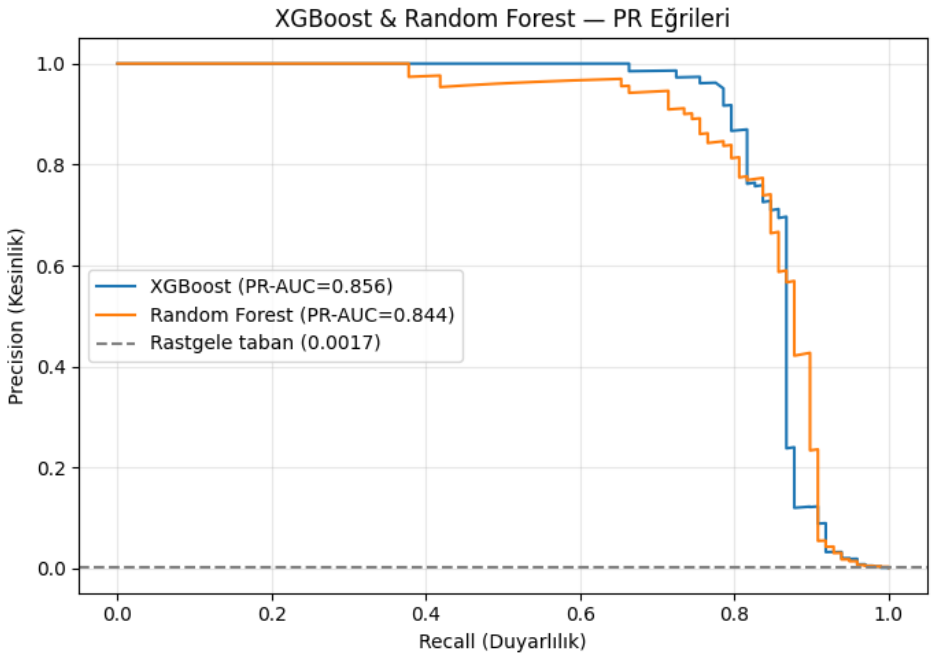
Model	TP	FP	FN	Kesinlik	Anma	F1-skoru	AUC-ROC
XGBoost (Optuna)	75	3	23	%96,2	%76,5	%85,2	0,9785
Rastgele Orman (Optuna)	78	16	20	%83,0	%79,6	%81,3	0,9821

Çizelge 4.11'deki sonuçlar, ağaç tabanlı topluluk modellerinin bu problem üzerinde son derece güçlü birer sınıflandırıcı olduğunu ortaya koymaktadır. XGBoost modeli %96 Kesinlik değeriyle tüm modeller arasındaki en yüksek kesinliğe ulaşmış; test kümesindeki 98 gerçek dolandırıcılık vakasının 75'ini yakalarken yalnızca 3 yanlış alarm üretmiştir (karar eşiği 0,992). Rastgele Orman ise 0,9821 ile en yüksek AUC-ROC değerini elde etmiş, %83 Kesinlik ve %80 Anma ile dengeli bir profil sergilemiştir (karar eşiği 0,722). Kesinlik-Anma eğrisi altındaki alan (PR-AUC) değerleri de bu tabloyu doğrulamaktadır: XGBoost 0,8556 ve Rastgele Orman 0,8440 PR-AUC değerlerine ulaşmıştır. Bu değerler, %0,17 gibi ileri düzey bir dengesizlik altında, rastgele sınıflandırıcının PR-AUC tabanının yalnızca 0,0017 olduğu göz önüne alındığında, her iki modelin de azınlık sınıfını çok güçlü biçimde ayırt edebildiğini göstermektedir.

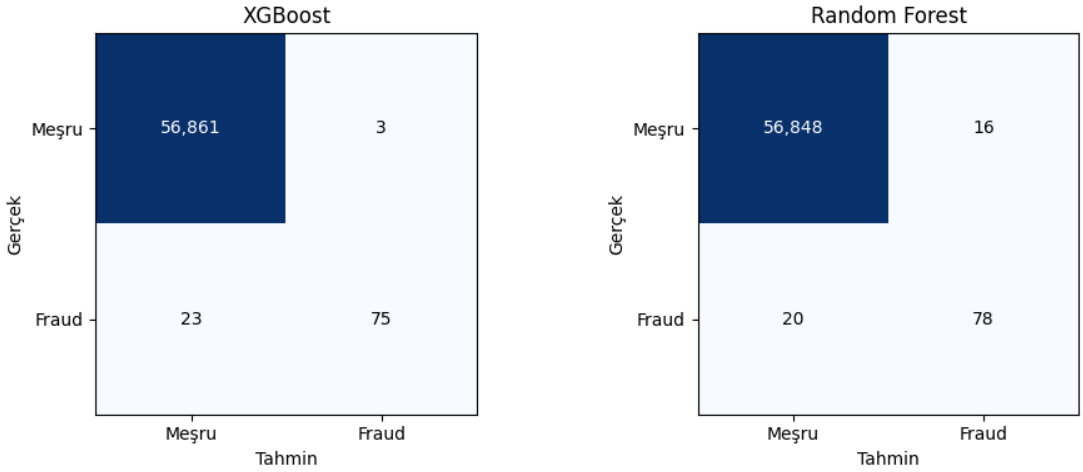
Ağaç tabanlı topluluk modellerinin performansı, bu çalışmanın merkezindeki CNN-LSTM hibrit modeliyle karşılaştırıldığında ilgi çekici bir tablo ortaya çıkmaktadır. Klasik sınıflandırma metrikleri açısından bakıldığında, XGBoost (%96 Kesinlik, %85 F1-skoru) ve Rastgele Orman (0,9821 AUC-ROC), CNN-LSTM hibrit modelinin (%76 Kesinlik, %80 F1-skoru, 0,9527 AUC-ROC) önüne geçmektedir. Bu bulgu, dolandırıcılık tespiti gibi yapısal (tablolar hâlindeki) verilerde ağaç tabanlı toplulukların hâlâ son derece rekabetçi olduğunu vurgulayan literatür ile uyumludur. Ancak bu üstünlük büyük ölçüde daha yüksek Kesinlik'dan kaynaklanmakta; CNN-LSTM hibrit modeli ise %85 ile en yüksek Anma değerine sahip olarak daha fazla gerçek dolandırıcılık vakasını yakalamaktadır (15 kaçan vaka; XGBoost'ta 23, Rastgele Orman'ta 20).

Bu farkın pratikteki karşılığı, Bölüm 4.7'de tanımlanan maliyet çerçevesi (kaçan her dolandırıcılık 500 €, her yanlış alarm 50 €) altında daha da belirginleşmektedir. Aynı formülle hesaplandığında XGBoost modelinin toplam maliyeti 11.650 € (23×500 € + 3×50 €), Rastgele Orman modelinin toplam maliyeti ise 10.800 € (20×500 € + 16×50 €) olmaktadır. Her iki değer de CNN-LSTM hibrit modelinin 8.800 €'luk toplam maliyetinin üzerindedir. Bunun temel nedeni, kaçan dolandırıcılığın (FN) yanlış alarmdan (FP) yaklaşık 10 kat daha maliyetli kabul edilmesi ve ağaç tabanlı modellerin yüksek Kesinlik'ı yüksek kesinlik uğruna görece daha fazla dolandırıcılık vakasını kaçırarak elde etmesidir. Dolayısıyla, salt istatistiksel metrikler ağaç tabanlı toplulukları öne çıkarsa da, maliyet odaklı değerlendirmede CNN-LSTM hibrit modelinin yüksek Anma'undan kaynaklanan finansal avantajı korunmaktadır. Bu sonuç, model seçiminin yalnızca tek bir metriğe değil, kurumun maliyet yapısına ve operasyonel önceliklerine göre yapılması gerektiği yönündeki tezin temel argümanını daha da güçlendirmektedir.

Ağaç tabanlı topluluk modellerinin Kesinlik-Anma eğrileri Şekil 4.12'de, karmaşıklık matrisleri ise Şekil 4.13'te sunulmaktadır. PR eğrileri, her iki modelin de yüksek Anma bölgesine kadar yüksek Kesinlik değerini koruduğunu; karmaşıklık matrisleri ise XGBoost'un yalnızca 3 yanlış alarmla dikkat çeken kesinliğini ve Rastgele Orman'ın daha dengeli hata dağılımını görsel olarak doğrulamaktadır.



Şekil 4.12 : XGBoost ve Rastgele Orman için Kesinlik-Anma Eğrileri.



Şekil 4.13 : XGBoost ve Rastgele Orman için Karmaşıklık Matrisleri.

4.11 Genel Karşılaştırma: Klasik ML, Ağaç Toplulukları ve Derin Öğrenme

Tüm modellerin karşılaştırmalı performansı Çizelge 4.12'de sunulmaktadır. Bu tablo, klasik ML'den derin öğrenme varsayılan parametrelerine ve ardından Optuna ile optimize edilmiş derin öğrenme modellerine doğru sistematik bir iyileşme eğilimini ortaya koymaktadır.

Çizelge 4.12 : Tüm Modellerin Karşılaştırmalı Performans Analizi.

Model	Kesinlik	Anma	F1-skoru	AUC-ROC	Aşama
Naive Bayes	%5,5	%87,8	%10,4	0,9645	Klasik ML

Model	Kesinlik	Anma	F1-skoru	AUC-ROC	Aşama
Lojistik regresyon	%5,6	%89,8	%10,5	0,9740	Klasik ML
KNN (k=5)	%45,3	%88,8	%60,0	0,9485	Klasik ML
Rastgele Orman (Optuna)	%83,0	%79,6	%81,3	0,9821	Ağaç Topluluğu
XGBoost (Optuna)	%96,2	%76,5	%85,2	0,9785	Ağaç Topluluğu
CNN (Varsayılan)	%69,2	%84,7	%76,1	0,9668	DL Varsayılan
LSTM (Varsayılan)	%43,6	%86,7	%58,0	0,9633	DL Varsayılan
CNN-LSTM (Varsayılan)	%55,5	%87,8	%68,0	0,9439	DL Varsayılan
CNN (Optuna)	%69,2	%82,7	%75,3	0,9575	DL Optimize
LSTM (Optuna)	%60,1	%84,7	%70,3	0,9492	DL Optimize
CNN-LSTM (Optuna)	%76,1	%84,7	%80,2	0,9527	DL Optimize

Çizelge 4.12'den açıkça görüleceği üzere, üç aşamalı bir iyileşme eğilimi mevcuttur. Birinci aşamada, klasik ML algoritmaları yüksek Anma değerlerine rağmen Kesinlik açısından ciddi biçimde yetersiz kalmış; özellikle Naive Bayes ve Lojistik regresyon'un her gerçek dolandırıcılık vakası için yaklaşık 17 yanlış alarm ürettiği görülmüştür. İkinci aşamada, derin öğrenme modellerinin varsayılan parametrelerle klasik ML'e kıyasla genel olarak daha iyi sonuçlar ürettiği; CNN modelinin %69 Kesinlik ile KNN'in (%45) açık biçimde üzerine çıktığı, buna karşın LSTM modelinin (%44) KNN ile benzer seviyede kaldığı tespit edilmiştir. Üçüncü ve kritik aşamada ise Optuna tabanlı optimizasyon, eşik ayarlama ve sınıf ağırlıklandırmanın birlikte uygulanmasıyla Kesinlik değerlerinin %60-76 bandına taşındığı ve özellikle hibrit modeldeki asıl performans sıçramasının bu bütünleşik optimizasyon yaklaşımından kaynaklandığı gösterilmiştir.

Bu üç aşamalı çerçeveye ek olarak, Bölüm 4.10'da değerlendirilen ağaç tabanlı topluluk modelleri (Rastgele Orman ve XGBoost), optimize edilmiş klasik olmayan bir ara katman olarak tabloya dâhil edilmiştir. XGBoost'un %96 Kesinlik ve %85 F1-skoru, Rastgele Orman'ın ise 0,9821 AUC-ROC değeriyle salt istatistiksel metriklerde derin öğrenme modellerinin dahi önüne geçebildiği görülmektedir. Bu bulgu, yapısal verilerde ağaç tabanlı toplulukların güçlü birer aday olduğunu teyit etmekte; ancak CNN-LSTM hibrit modelinin en yüksek Anma (%85) ve buna bağlı en düşük toplam maliyet (8.800 €) değerini koruması, maliyet duyarlı dolandırıcılık tespiti bağlamında derin öğrenme temelli yaklaşımın pratik üstünlüğünü sürdürdüğünü göstermektedir. Dolayısıyla model seçimi, tek bir performans metriğine değil, kurumun maliyet yapısına ve öncelikli hata türüne bağlı olarak yapılmalıdır.

5. SONUÇLAR VE ÖNERİLER

Bu tez çalışmasında, kredi kartı dolandırıcılık tespiti gibi ileri düzey sınıf dengesizliği içeren bir problem üzerinde, hem klasik makine öğrenmesi (KNN, Naive Bayes, Lojistik regresyon) hem de derin öğrenme tabanlı üç farklı mimarinin (CNN, LSTM ve CNN-LSTM hibrit) performansı sistematik biçimde karşılaştırılmış; Optuna tabanlı Bayes hiperparametre optimizasyonu, sınıf ağırlıklandırma ve karar eşiği ayarlamasının birlikte uygulanmasıyla elde edilen iyileşmeler sayısal olarak ortaya konmuştur.

Çalışma kapsamında elde edilen temel bulgular aşağıdaki gibi özetlenebilir:

- Klasik makine öğrenmesi algoritmalarının (KNN, Naive Bayes, Lojistik regresyon) ileri düzey sınıf dengesizliği altında Kesinlik değerleri %6-45 aralığında kalmış; pratik olarak yetersiz oldukları deneysel olarak gösterilmiştir. Bu bulgu, derin öğrenme yaklaşımlarının tercih edilmesinin gerekçesini somut verilerle desteklemektedir.
- Optuna tabanlı hiperparametre optimizasyonu, LSTM ve CNN-LSTM modellerinde Kesinlik metriğinde belirgin iyileşmeler sağlamıştır. LSTM modelinde Kesinlik değeri %44'ten %60'a (+16 puan), CNN-LSTM hibrit modelinde ise %55'ten %76'ya (+21 puan) yükselmiş; varsayılan parametrelerle dahi %69 Kesinlik değerine ulaşan CNN modeli bu seviyesini korumuştur.
- F1-skoru değerlerinde de tutarlı iyileşmeler gözlemlenmiştir. LSTM modelinde F1-skoru değeri %58'den %70'e, CNN-LSTM hibrit modelinde ise %68'den %80'e yükselmiş; hibrit model %80 F1-skoru ile en yüksek değere ulaşmıştır.
- AUC-ROC değerlerinde üç model de 0,94 üzerinde değerler elde etmiş; optimize edilmiş modeller arasında CNN 0,9575 ile en yüksek AUC-ROC değerini sağlarken, CNN-LSTM hibrit modeli 0,8506 ile en yüksek PR-AUC değerine ulaşmıştır.

- Maliyet analizi sonuçlarına göre CNN-LSTM hibrit modeli, 8.800 € ile en düşük toplam maliyeti üreten model olmuştur.
- Optimize edilmiş ağaç tabanlı topluluk modelleri (XGBoost ve Rastgele Orman) güçlü birer referans olarak değerlendirilmiş; XGBoost %96 Kesinlik ve %85 F1-skoru, Rastgele Orman ise 0,9821 AUC-ROC ile klasik metriklerde öne çıkmıştır. Buna karşın, kaçan dolandırıcılığın baskın maliyet unsuru olduğu değerlendirme çerçevesinde, en yüksek Anma (%85) ve en düşük toplam maliyete (8.800 €) sahip CNN-LSTM hibrit modelinin pratik üstünlüğü korunmuştur.
- SMOTE, sınıf ağırlıklandırma ve eşik ayarlamanın birlikte uygulanması, dengesizlik problemine yönelik bütünleşik ve etkili bir çerçeve oluşturmaktadır.

5.1 Çalışmanın Akademik ve Pratik Katkıları

Bu tezin literatüre sunduğu akademik katkılar birkaç açıdan değerlendirilebilir. Birincisi, klasik ML algoritmalarının bu problem üzerindeki yetersizliğini deneysel olarak göstererek, derin öğrenme yaklaşımlarının tercih edilmesinin nesnel gerekçesini ortaya koymaktadır. İkincisi, üç farklı derin öğrenme mimarisi eşit koşullarda ve bütünleşik bir optimizasyon çerçevesi içinde karşılaştırılmıştır. Üçüncüsü, Optuna tabanlı hiperparametre optimizasyonunun eşik ayarlama ve sınıf ağırlıklandırma ile birlikte sistematik biçimde uygulanmasının üstünlüğü deneysel olarak gösterilmiştir. Dördüncüsü, model değerlendirmesinin yalnızca klasik metrikler üzerinden değil, finansal anlamda yorumlanabilir bir maliyet çerçevesi üzerinden de yapılmasıdır. Beşincisi, derin öğrenme modelleri yalnızca görece zayıf klasik yöntemlerle değil, yapısal verilerde güçlü referans kabul edilen ve Optuna ile optimize edilmiş ağaç tabanlı topluluk modelleriyle (Rastgele Orman ve XGBoost) de karşılaştırılmış; bu sayede bulguların geçerliliği ve karşılaştırmanın adilliği güçlendirilmiştir.

Pratik açıdan ise çalışma, gerçek bankacılık ve finans kurumları için doğrudan uygulanabilir nitelikte bir çerçeve sunmaktadır. Model seçiminin operasyonel önceliklere göre değiştirilebilir olması, karar vericilere esnek ve yorumlanabilir alternatifler sunmaktadır.

5.2 Yönetsel Çıkarımlar ve Politika Önerileri

Bu çalışmada elde edilen bulgular, yalnızca teknik bir model karşılaştırması olarak değil, finansal kurumların operasyonel ve stratejik karar süreçlerine somut katkı sağlayan bir çerçeve olarak değerlendirilmelidir. Aşağıda, deneysel bulguların bankacılık ve finans sektörü için yönetsel çıkarımları (managerial implications) üç ana eksen üzerinden ele alınmaktadır: (i) operasyonel verimlilik, (ii) müşteri deneyimi ve sadakat, (iii) stratejik bütçe planlaması ve risk yönetimi.

5.2.1 Operasyonel verimlilik ve iş yükü yönetimi

Optuna optimizasyonu sonrasında Kesinlik değerinin %55'ten %76'ya yükselmesi, banka Dolandırıcılık İzleme Operasyon Merkezi'ndeki (Fraud Operations Center) analist iş yükünü doğrudan etkileyen bir gelişmedir. Varsayılan parametrelerle çalışan CNN-LSTM modeli her gerçek dolandırıcılık vakası için yaklaşık 0,8 yanlış alarm üretirken, optimize edilmiş model bu oranı yaklaşık 0,3'e düşürmektedir. Bu iyileşme, operasyonel ölçekte günde binlerce işlemi inceleyen analist ekiplerinin gerçek riskli vakalara odaklanabilmesini sağlamakta; çağrı merkezi yüklerinin, manuel doğrulama süreçlerinin ve müşteri hizmetleri operasyonlarının verimliliğini önemli ölçüde artırmaktadır. Bu durum aynı zamanda mevcut insan kaynağının daha stratejik görevlere (gelişmiş dolandırıcılık kalıplarının analizi, model performansının izlenmesi gibi) yönlendirilmesine olanak tanımaktadır.

5.2.2 Müşteri deneyimi ve sadakat

Bankacılık literatüründe yanlış alarmların müşteri memnuniyeti ve sadakat üzerindeki olumsuz etkisi sıkça vurgulanmaktadır. Meşru bir işlem sırasında kartının haksız yere bloke edilmesi veya işlemin reddedilmesi, müşterinin bankasına olan güvenini sarsmakta ve Net Promoter Score (NPS) gibi sadakat göstergelerini olumsuz etkilemektedir. Bu çalışmada önerilen CNN-LSTM hibrit modelinin yalnızca 26 yanlış alarm üretmesi (varsayılan parametrelere kıyasla yaklaşık 2,7 kat daha az), bankanın müşterilerine sunduğu deneyimi doğrudan iyileştirmektedir. Özellikle yurtdışında veya yüksek tutarlı işlemlerde haksız kart bloklamalarının azalması, müşteri kaybının (churn) önüne geçmekte ve uzun vadeli müşteri yaşam boyu değeri (Customer Lifetime Value) açısından stratejik bir kazanım oluşturmaktadır.

5.2.3 Stratejik bütçe planlaması ve risk yönetimi

Bu çalışmada uygulanan maliyet odaklı değerlendirme yaklaşımı, model seçimini yalnızca istatistiksel bir tercih olmaktan çıkarıp finansal bir karar olarak konumlandırmaktadır. CNN-LSTM hibrit modelinin 8.800 € ile en düşük tahmini toplam maliyeti üretmesi, bankanın risk yönetimi ve siber güvenlik bütçesi planlamasında somut bir girdi sağlamaktadır. Dahası, modelin sağladığı maliyet tasarrufunun bankanın Ar-Ge, MLOps altyapısı veya açıklanabilir yapay zekâ (XAI) çalışmalarına yeniden yatırılması mümkündür. Yönetim kademesi, tezde esas alınan maliyet yapısı altında her iki hata türünde de en iyi dengeyi sunan CNN-LSTM hibrit modelini; kaçan dolandırıcılık maliyetinin baskın hâle geldiği (yüksek hacimli perakende bankacılık operasyonları gibi) kurumlar için ise Çizelge 4.9'daki senaryo analizinde öne çıkan yüksek Anma'lu alternatifleri tercih edebilir. Bu çoklu seçenek yapısı, modelin gerçek bankacılık operasyonlarına uyarlanabilirliğini artırmaktadır.

5.3 Çalışmanın Sınırlılıkları

Her bilimsel çalışmada olduğu gibi bu tezin de bazı sınırlılıkları bulunmaktadır. Birinci sınırlılık, kullanılan veri setinin 2013 yılına ait olmasıdır. Bu durum, dolandırıcılık tespit literatüründe sıkça tartışılan kavram kayması (concept drift) olgusunu gündeme getirmektedir. Dolandırıcılık yöntemleri zaman içinde evrilmekte; saldırganlar yeni teknolojilerden ve davranış kalıplarından yararlanarak modellerin daha önce karşılaşmadığı örüntüler üretmektedir. Bu nedenle 2013 verisi üzerinde eğitilmiş bir modelin 2026 koşullarındaki performansının doğrulanabilmesi için, modelin canlı bankacılık sistemlerine entegrasyonu sırasında belirli periyotlarla (örneğin aylık veya çeyrek dönemlik) yeni verilerle yeniden eğitilmesi (continuous training) ve eğitim ile üretim verisi arasındaki dağılımsal farkların izlenmesi (data drift monitoring) gerekmektedir. MLOps altyapısı içinde bu süreçlerin sistematik biçimde yürütülmesi, çalışmanın canlı sistemlere transferi açısından kritik öneme sahiptir. İkinci sınırlılık, veri setindeki özelliklerin PCA ile anonimleştirilmiş olmasıdır; bu durum özellik mühendisliği (feature engineering) çalışmalarının kapsamını daraltmıştır. Üçüncü sınırlılık, maliyet analizinde kullanılan birim maliyet değerlerinin (500 €/50 €) Bhattacharyya ve diğ. (2011), Bahnsen ve diğ. (2014) ile Dal Pozzolo ve diğ. (2018) çalışmalarından alınan ortalama tahminler olmasıdır; bu değerlerin gerçek bir bankanın iç maliyet yapısına göre kalibre edilmesi pratikte

mümkündür. Son olarak, çalışmada doğrulama yöntemi olarak Hold-out yaklaşımı tercih edilmiş olup, hesaplama maliyeti üstlenildiği takdirde K-Katlı Çapraz Doğrulama (K-Fold Cross Validation) ile sonuçların ek bir güvence katmanı kazandırılması da mümkündür.

Gelecek çalışmalarda aşağıdaki konular ele alınabilir:

- Focal Loss gibi dengesiz veri problemine yönelik özel kayıp fonksiyonlarının uygulanması.
- SMOTE-ENN, ADASYN ve Borderline-SMOTE gibi gelişmiş örnekleme tekniklerinin karşılaştırılması.
- Stacking, Voting ve Bagging gibi topluluk öğrenme yöntemleri ile modellerin birleştirilmesi.
- Dikkat (Attention) mekanizmaları ve Transformer tabanlı mimarilerin değerlendirilmesi.
- SHAP ve LIME gibi açıklanabilir yapay zekâ tekniklerinin entegrasyonu.
- Federated Learning ile veri gizliliğini koruyan modellerin geliştirilmesi.
- Optuna trial sayısının 200-500'e çıkarılması.
- Graph Neural Networks (GNN) ile işlemler arası ilişkisel bağlantıların modellenmesi.

Sonuç olarak, bu tez çalışması; klasik makine öğrenmesi algoritmalarının ileri düzey sınıf dengesizliği altındaki yetersizliğini ortaya koyarak derin öğrenme yaklaşımlarının tercihini somut verilerle desteklemiş; ardından bu derin öğrenme modellerinin başarısının yalnızca mimari seçimiyle değil, sistematik hiperparametre optimizasyonu, sınıf ağırlıklandırma ve karar eşiği ayarlaması gibi tamamlayıcı stratejilerin birlikte uygulanmasıyla şekillendiğini deneysel olarak ortaya koymuştur. Önerilen Optuna tabanlı çerçeve, hem akademik literatüre özgün bir katkı sunmakta hem de finansal kurumlara doğrudan uygulanabilir pratik araçlar sağlamaktadır. Ayrıca, güçlü referans yöntemler olan XGBoost ve Rastgele Orman ile yürütülen karşılaştırma, önerilen çerçevenin göreli konumunu netleştirmiş ve elde edilen bulguların sağlamlığını (robustness) artırmıştır.

KAYNAKLAR

Akiba, T., Sano, S., Yanase, T., Ohta, T. ve Koyama, M. (2019). Optuna: A next-generation hyperparameter optimization framework. In *Proceedings of the 25th ACM SIGKDD* (pp. 2623-2631). ACM.

Al Balawi, S. ve Aljohani, N. (2023). Credit-card fraud detection system using neural networks. *International Arab Journal of Information Technology*, *20*(2), 234-241.

Alamri, M. ve Ykhlef, M. (2026). A novel hybrid machine learning model for credit card fraud detection using anomaly detection and optimised classification. *PeerJ Computer Science*, *12*, e3748.

Alarfaj, F. K. ve Shahzadi, S. (2024). Enhancing fraud detection in banking with deep learning: Graph neural networks and autoencoders for real-time credit card fraud prevention. *IEEE Access*, *12*, 134342-134353.

Ali, A., Abd Razak, S., Othman, S. H., Eisa, T. A. E., Al-Dhaqm, A., Nasser, M., Elhassan, T., Elshafie, H. ve Saif, A. (2022). Financial fraud detection based on machine learning: A systematic literature review. *Applied Sciences*, *12*(19), 9637.

Ali, M. N. Y., Kabir, T., Raka, N. L., Toma, S. S., Kulsum, M. L. ve Ferdaus, J. (2022). SMOTE based credit card fraud detection using convolutional neural network. *International Journal of Computer Applications*, *184*(15), 22-29.

Almalki, S. ve Masud, M. (2022). Financial fraud detection using explainable AI and stacking ensemble methods. *IEEE Access*, *10*, 89456-89472.

Alzahrani, R. A. ve Aljabri, M. (2023). AI-based techniques for ad click fraud detection and prevention: Review and research directions. *Journal of Sensor and Actuator Networks*, *12*(1), 4.

Aurna, N. F., Hossain, M. D., Taenaka, Y. ve Kadobayashi, Y. (2023). Federated learning-based credit card fraud detection. In *Proc. IEEE International Conference on Cyber Security and Resilience* (pp. 1-8). IEEE.

Badjatiya, P., Gupta, M. ve Gupta, V. (2021). Deep learning for credit card fraud detection: A review of algorithms, challenges, and solutions. *Expert Systems with Applications*, *175*, 114820.

Bahnsen, A. C., Aouada, D. ve Ottersten, B. (2014). Example-dependent cost-sensitive logistic regression for credit scoring. *2014 13th International Conference on Machine Learning and Applications*, 263-269. IEEE.

Banu, S. R., Gongada, T. N., Santosh, K., Chowdhary, H., Sabareesh, R. ve Muthuperumal, S. (2024). Financial fraud detection using hybrid

convolutional and recurrent neural networks. *2024 10th International Conference on Communication and Signal Processing (ICCSP)*, 1027-1031. IEEE.

Bao, Y., Hilary, G. ve Ke, B. (2022). Artificial intelligence and fraud detection. In *Innovative Technology at the Interface of Finance and Operations* (pp. 223-247). Springer.

Barman, S., Pal, U., Sarfaraj, M. A., Biswas, B., Mahata, A. ve Mandal, P. (2023). A complete literature review on financial fraud detection applying data mining techniques. *International Journal of Trust Management in Computing and Communications, 6*(4), 305-331.

Bengio, Y., Simard, P. ve Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks, 5*(2), 157-166.

Berhane, T. (2023). A hybrid convolutional neural network and support vector machine-based credit card fraud detection model. *Mathematical Problems in Engineering, 2023*, 8134627.

Bhattacharyya, S., Jha, S., Tharakunnel, K. ve Westland, J. C. (2011). Data mining for credit card fraud: A comparative study. *Decision Support Systems, 50*(3), 602-613.

Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of ML algorithms. *Pattern Recognition, 30*(7), 1145-1159.

Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5-32.

Chawla, N. V., Bowyer, K. W., Hall, L. O. ve Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *JAIR, 16*, 321-357.

Chen, J. I.-Z. ve Lai, K.-L. (2021). Deep convolution neural network model for credit-card fraud detection and alert. *Journal of AI and Capsule Networks, 3*(2), 101-112.

Chen, T. ve Guestrin, C. (2016). XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 785-794). ACM.

Chen, Y., Zhao, C., Xu, Y., Nie, C. ve Zhang, Y. (2025). Year-over-year developments in financial fraud detection via deep learning: A systematic literature review. arXiv preprint arXiv:2502.00201.

Cherif, A., Badhib, A., Ammar, H., Alshehri, S., Kalkatawi, M. ve Imine, A. (2023). Credit card fraud detection in the era of disruptive technologies: A systematic review. *Journal of King Saud University - Computer and Information Sciences, 35*(1), 145-174.

Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C. ve Bontempi, G. (2018). Credit card fraud detection: A realistic modeling and a novel learning strategy. *IEEE Transactions on Neural Networks and Learning Systems, 29*(8), 3784-3797.

Esenogho, E., Mienye, I. D., Swart, T. G., Aruleba, K. ve Obaido, G. (2022). A neural network ensemble with feature engineering for improved credit card fraud detection. *IEEE Access, 10*, 16400-16407.

Fu, K., Cheng, D., Tu, Y. ve Zhang, L. (2016). Credit card fraud detection using convolutional neural networks. In *Neural Information Processing* (pp. 483-490). Springer.

Gambo, M. L., Zainal, A. ve Kassim, M. N. (2022). A convolutional neural network model for credit card fraud detection. In *2022 ICoDSA* (pp. 198-202). IEEE.

Gers, F. A., Schmidhuber, J. ve Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural Computation, 12*(10), 2451-2471.

Goodfellow, I., Bengio, Y. ve Courville, A. (2016). *Deep learning*. MIT Press.

Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R. ve Schmidhuber, J. (2017). LSTM: A search space odyssey. *IEEE Transactions on Neural Networks and Learning Systems, 28*(10), 2222-2232.

Güven, Ö. ve Aras, S. (2022). Fraud detection by machine learning algorithms: A case from a mobile payment system. *Uluslararası Yönetim İktisat ve İşletme Dergisi, 18*(3), 895-911.

He, H. ve Garcia, E. A. (2009). Learning from imbalanced data. *IEEE TKDE, 21*(9), 1263-1284.

Hernandez Aros, L., Bustamante Molano, L. X., Gutierrez-Portela, F., Moreno Hernandez, J. J. ve Rodríguez Barrero, M. S. (2024). Financial fraud detection through the application of machine learning techniques: A literature review. *Procedia Computer Science, 231*, 56-72.

Hochreiter, S. ve Schmidhuber, J. (1997). Long short-term memory. *Neural Computation, 9*(8), 1735-1780.

Ileberi, E., Sun, Y. ve Wang, Z. (2022). A machine learning based credit card fraud detection using the GA algorithm for feature selection. *Journal of Big Data, 9*(1), 24.

Jain, M. ve Shah, A. (2024). Anomaly detection using convolutional neural networks (CNN). *ESP International Journal of Advancements in Computational Technology, 2*(3), 12-22.

Karim, M. R., Kamal, M., Rahman, A. ve Hossain, A. (2025). An explainable ensemble model for credit card fraud detection using advanced data balancing and feature selection technique. arXiv preprint arXiv:2503.21160.

Kingma, D. P. ve Ba, J. (2015). Adam: A method for stochastic optimization. In *Proc. ICLR*.

Kiranyaz, S., Avci, O., Abdeljaber, O., Ince, T., Gabbouj, M. ve Inman, D. J. (2021). 1D convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing, 151*, 107398.

LeCun, Y., Bengio, Y. ve Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436-444.

LeCun, Y., Bottou, L., Bengio, Y. ve Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE, 86*(11), 2278-2324.

Makki, S., Assaghir, Z., Taher, Y., Haque, R., Hacid, M.-S. ve Zeineddine, H. (2019). An experimental study with imbalanced classification approaches for credit card fraud detection. *IEEE Access, 7*, 93010-93022.

Marazqah Btoush, E. A. L., Zhou, X., Gururajan, R., Chan, K. C., Genrich, R. ve Sankaran, P. (2023). A systematic review of literature on credit card cyber fraud detection using machine and deep learning. *PeerJ Computer Science, 9*, e1278.

Mienye, I. D. ve Sun, Y. (2023). A deep learning ensemble with data resampling for credit card fraud detection. *IEEE Access, 11*, 30628-30638.

Mizher, A. ve Nassif, A. B. (2023). Deep CNN approach for unbalanced credit card fraud detection data. *2023 Advances in Science and Engineering Technology International Conferences (ASET)*, 1-6. IEEE.

Papasavva, A., Johnson, S., Lowther, E., Lundrigan, S., Mariconti, E., Markovska, A. ve Tuptuk, N. (2024). Application of AI-based models for online fraud detection and analysis. arXiv preprint arXiv:2409.19022.

Provost, F. ve Fawcett, T. (1997). Analysis and visualization of classifier performance: Comparison under imprecise class and cost distributions. In *Proceedings of the Third International Conference on Knowledge Discovery and Data Mining (KDD-97)*, 43-48. AAAI Press.

Raghavan, P. ve El Gayar, N. (2019). Fraud detection using machine learning and deep learning. In *2019 ICCIKE* (pp. 334-339). IEEE.

Roy, A., Sun, J., Mahoney, R., Alonzi, L., Adams, S. ve Beling, P. (2018). Deep learning detecting fraud in credit card transactions. In *2018 Systems and Information Engineering Design Symposium (SIEDS)*, 129-134. IEEE.

Saito, T. ve Rehmsmeier, M. (2015). The Precision-Recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLoS ONE, 10*(3), e0118432.

Sánchez-Aguayo, M., Urquiza-Aguiar, L. ve Estrada-Jiménez, J. (2021). Fraud detection using the fraud triangle theory and data mining techniques: A literature review. *Computers, 10*(10), 121.

Shen, J. (2021). Credit card fraud detection using autoencoder-based deep neural networks. In *2021 IEEE ICBAIE* (pp. 673-677). IEEE.

Sinap, V. (2024). Comparative analysis of machine learning techniques for credit card fraud detection: Dealing with imbalanced datasets. *Turkish Journal of Engineering, 8*(2), 196-208.

Soltani, M., Kythreotis, A. ve Roshanpoor, A. (2023). Two decades of financial statement fraud detection literature review. *Journal of Financial Crime, 30*(5), 1367-1388.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. ve Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research, 15*(1), 1929-1958.

Yang, Y., Yu, Y. ve Li, T. (2022). Deep learning techniques for financial fraud detection. In *2022 4th ICITIIT*. IEEE.

Yu, C., Xu, Y., Cao, J., Zhang, Y., Jin, Y. ve Zhu, M. (2024). Credit card fraud detection using advanced transformer model. arXiv preprint arXiv:2406.03733.

Zhou, H., Zheng, X. ve Du, J. (2025). A deep learning hybrid RNN-LSTM model for credit card fraud detection. Research Square (preprint). DOI: 10.21203/rs.3.rs-8327749/v1.

ÖZGEÇMİŞ

Ad-Soyad: Elvin MAMMADLİ

ÖĞRENİM DURUMU:

Yüksek Lisans: İstanbul Teknik Üniversitesi, Lisansüstü Eğitim Enstitüsü, İşletme Anabilim Dalı, İşletme Programı, 2024 – 2026

Lisans: Yıldız Teknik Üniversitesi, Bilgisayar Mühendisliği Bölümü, 2016 – 2021

MESLEKİ DENEYİM VE ÖDÜLLER:

HITIT Computer Solutions, Bilgisayar Mühendisi, Eylül 2021 – Hâlen.

YABANCI DİL:

Azerice (anadil), Türkçe, İngilizce.

TEZDEN TÜRETİLEN YAYINLAR, SUNUMLAR VE PATENTLER:

MAMMADLİ, E. ve **ÖZCAN, T.** (2026). Comparison of Deep Learning Models with Optuna-Based Hyperparameter Optimization for Credit Card Fraud Detection. International Symposium for Production Research (ISPR 2026), sözlü bildiri, Bildiri No: 43 (kabul edildi).