
GAZİ ÜNİVERSİTESİ
Gazi
1926

SINIF DENGESİZ BÜYÜK VERİDE DOLANDIRICILIK TESPİTİ VE AÇIKLANABİLİRLİK

Duygu SİNANÇ TERZİ

DOKTORA TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

EKİM 2020

Duygu SİNANÇ TERZİ tarafından hazırlanan "SINIF DENGESİZ BÜYÜK VERİDE DOLANDIRICILIK TESPİTİ VE AÇIKLANABİLİRLİK" adlı tez çalışması aşağıdaki jüri tarafından OY BİRLİĞİ ile Gazi Üniversitesi Bilgisayar Mühendisliği Ana Bilim Dalında DOKTORA TEZİ olarak kabul edilmiştir.

Danışman: Prof. Dr. Şeref SAĞIROĞLU

Bilgisayar Mühendisliği Ana Bilim Dalı, Gazi Üniversitesi

Bu tezin, kapsam ve kalite olarak Doktora Tezi olduğunu onaylıyorum. ……..

İkinci Danışman: Dr. Umut DEMİREZEN

Yapay Zekâ ve Büyük Veri Birimi, Cumhurbaşkanlığı Dijital Dönüşüm Ofisi

Bu tezin, kapsam ve kalite olarak Doktora Tezi olduğunu onaylıyorum. ...……

Başkan: Prof. Dr. Müslim BOZYİĞİT

Bilgisayar Mühendisliği Ana Bilim Dalı, Çankaya Üniversitesi

Bu tezin, kapsam ve kalite olarak Doktora Tezi olduğunu onaylıyorum. ...……

Üye: Prof. Dr. Suat ÖZDEMİR

Bilgisayar Mühendisliği Ana Bilim Dalı, Hacettepe Üniversitesi

Bu tezin, kapsam ve kalite olarak Doktora Tezi olduğunu onaylıyorum. ...……

Üye: Prof. Dr. Volkan ATALAY

Bilgisayar Mühendisliği Ana Bilim Dalı, Orta Doğu Teknik Üniversitesi

Bu tezin, kapsam ve kalite olarak Doktora Tezi olduğunu onaylıyorum. ...……

Üye: Doç. Dr. Hacer KARACAN

Bilgisayar Mühendisliği Ana Bilim Dalı, Gazi Üniversitesi

Bu tezin, kapsam ve kalite olarak Doktora Tezi olduğunu onaylıyorum. .....……

Üye: Dr. Öğr. Üyesi Mehmet DEMİRCİ

Bilgisayar Mühendisliği Ana Bilim Dalı, Gazi Üniversitesi

Bu tezin, kapsam ve kalite olarak Doktora Tezi olduğunu onaylıyorum. ...……

Tez Savunma Tarihi: 02/10/2020

Jüri tarafından kabul edilen bu tezin Doktora Tezi olması için gerekli şartları yerine getirdiğini onaylıyorum.

……

Prof. Dr. Cevriye GENCER

Fen Bilimleri Enstitüsü Müdürü

ETİK BEYAN

Gazi Üniversitesi Fen Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmasında;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmasında yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu,

bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

…

Duygu SİNANÇ TERZİ

…../……./……

SINIF DENGESİZ BÜYÜK VERİDE DOLANDIRICILIK TESPİTİ VE AÇIKLANABİLİRLİK

(Doktora Tezi)

Duygu SİNANÇ TERZİ

GAZİ ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

Ekim 2020

ÖZET

Elektronik ortamlarda gerçekleştirilen dolandırıcılık, doğası gereği dinamik bir problemdir. Çözüm modellerinin geliştirilmesi için eğitim sürecinde kullanılacak olan verilerin temini ve temin edilen verilerin etiketlenmesi zor, dolandırıcı olarak etiketli verinin miktarı ise genele oranla oldukça azdır. Büyük veri çağının sınıflar arası farkı daha da artırması ve geleneksel makine öğrenmesi yaklaşımlarının dengeli sınıf dağılımı varsayımı üzerinde tasarlanmış olması, dolandırıcılık tespit sürecini daha da zorlaştırmaktadır. Bu bağlamda, ilk defa bu tez kapsamında dolandırıcılık tespiti problemi sınıf dengesiz büyük veride veri bilimi bakış açısı ile ele alınmıştır. Telekom ve kredi kartı dolandırıcılığı özelinde, üç yeni tespit yöntemi ve bir açıklanabilir yapay zekâ yaklaşımı geliştirilmiştir. İlk yöntem, hem kullanıcıların belirli bir zaman diliminde yaptığı aktivitelerde oluşan anormal durumlardan hem de bilinen dolandırıcı aktivitelerinden faydalanılan, büyük veri analitiği tabanlı bir tespit sunar. İkinci yöntem, çoğunluk ve azınlık sınıf arasındaki dengesizliği gidermek amacıyla büyük veri analitiği ile kümeleme tabanlı yeniden örnekleme yaptıktan sonra sınıflandırma gerçekleştirerek dolandırıcılığı tespit eder. Üçüncü yöntem ise, dolandırıcılık verisini zamansal ilişkilerini koruyarak görüntüye dönüştürüp, özelliklerin ikili ilişkilerini çıkardıktan sonra derin sinir ağı ile sınıflandırmayı sağlar. Bu yöntemlere ilave olarak, geliştirilen yöntemler açıklanabilirlik açısından ele alınmış ve üçüncü yöntemin ne öğrendiğini daha iyi ortaya koymak amacıyla ısı haritası tabanlı yeni bir açıklanabilirlik yaklaşımı geliştirilmiştir. Bu yaklaşım, dolandırıcı-görüntü dönüşümünü sağlayan üçüncü tekniğin ürettiği ve belirli bir ölçekteki renk haritalarından oluşan görüntüler üzerindeki ilişkilerin daha net ifade edilmesini sağlamaktadır. Bu tez çalışmasında önerilen yöntemlerin; dolandırıcılık problemlerinin karşılaşılmadan çözümlenmesi, mevcudiyeti halinde otomatik olarak tespiti ve kullanıcılar ile hizmet sağlayıcıların karşılaşacakları risklerin azaltılmasına katkılar sağlaması beklenmektedir. Ayrıca, önerilen açıklanabilirlik yaklaşımının ise sadece dolandırıcılık değil diğer yapay zekâ modellerinin sonuçlarının da yorumlanmasına katkılar sağlayacağı öngörülmektedir.

Bilim Kodu	:	92414
Anahtar Kelimeler	:	Büyük veri, sınıf dengesizliği, dolandırıcılık tespiti, açıklanabilirlik
Sayfa Adedi	:	165
Danışman	:	Prof. Dr. Şeref SAĞIROĞLU
İkinci Danışman	:	Dr. Umut DEMİREZEN

FRAUD DETECTION MODELS IN CLASS IMBALANCED BIG DATA AND EXPLAINABILITY

(Ph. D. Thesis)

Duygu SİNANÇ TERZİ

GAZİ UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

October 2020

ABSTRACT

Fraud in electronic environments is a dynamic problem by its nature. It is difficult to obtain and label the data to be used in the training process for the development of solution models, and the amount of data labeled as fraudulent is relatively low compared to general. The fact that the big data age further increases the gap between classes and traditional machine learning approaches are designed on the assumption of balanced class distribution, making the fraud detection process even more difficult. In this context, for the first time in this thesis, the problem of fraud detection has been discussed in class imbalanced big data from a data science perspective. Specific to telecom and credit card fraud, three new detection methods and an explainable artificial intelligence approach have been developed. The first method provides a big data analytics-based detection that takes advantage of both anomalous situations occurring in the activities and known fraudulent activities of users in a certain time period. The second method is a big data analytics-based classifier that resamples before analysis in order to compensate for the imbalance between majority and minority class. The third method, on the other hand, enables processing-based data to be transformed into an image by preserving its temporal relationships, and after extracting the binary relationships of features, classifying it with a deep neural network. In addition to these methods, the developed methods are discussed in terms of Explained Artificial Intelligence (XAI) and a new heat map-based XAI approach has been developed to help better understand what the third method has learned. This approach provides a clearer expression of the relationships on images produced by the technique of fraud-image transformation and made up of color maps of a certain scale. The methods proposed in this thesis are expected to contribute to the resolution of fraud problems encountered in electronic environments, to automatically detect them if they exist, and to reduce the risks that users and service providers will encounter. In addition, the proposed XAI approach is predicted to contribute not only to fraud, but also to the interpretation of the results of other artificial intelligence models.

Science Code	:	92414
Key Words	:	Big data, class imbalanced data, fraud detection, explainability
Page Number	:	165
Supervisor	:	Prof. Dr. Şeref SAĞIROĞLU
Co-Supervisor	:	Dr. Umut DEMİREZEN

TEŞEKKÜR

Bu tez çalışması esnasında ve akademik gelişim sürecimde büyük emekleri olan kıymetli danışmanlarım Prof. Dr. Şeref SAĞIROĞLU ve Dr. Umut DEMİREZEN'e, desteğini hiçbir zaman esirgemeyen sevgili eşim Dr. Öğr. Üyesi Ramazan TERZİ'ye, hayatımı güzelleştiren oğlum Fatih TERZİ'ye ve her zamanda yanımda olan annem, babam ve ablama sonsuz teşekkürlerimi sunarım.

Bu tez; TÜBİTAK TEYDEB 1501 destekli, 3151243 numaralı "Ağ ve Uygulama Seviyesinde Büyük Veri Analitiği ile Davranışsal Anomali ve Kural İhlali Tespiti Yapan, Raporlayan ve Güvenlik Araçlarını Bilgilendiren Yeni Nesil Analiz Aracı Tasarımı" projesi kapsamında tez danışmanım Prof. Dr. Şeref SAĞIROĞLU'nun yapmış olduğu proje danışmanlığı kapsamında gerçekleştirilmiştir.

Gazi Üniversitesi Bilimsel Araştırma Projeleri (BAP) Birimi tarafından desteklenen 06/2015-04 numaralı "Büyük Veri Analiz Merkezi" altyapı projesi (BIDISEC) kapsamında kurulan laboratuvar altyapısı kullanılarak, büyük veri teknolojileri ile model geliştirme ve test süreçleri tamamlanmıştır.

Bu tez kapsamında çalışmalarımıza destek veren; NETAŞ, Gazi Üniversitesi BAP Birimi ve Gazi BIDISEC'e teşekkür ederim.

İÇİNDEKİLER

Sayfa

Sayfa

Sayfa

ÇİZELGELERİN LİSTESİ

Çizelge **Sayfa**

ŞEKİLLERİN LİSTESİ

Şekil **Sayfa**

Şekil **Sayfa**

Şekil **Sayfa**

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Kısaltmalar	Açıklamalar
3V	Hacim, Hız, Çeşitlilik (Volume, Velocity, Variety)
3^2V	Hacim, Hız, Çeşitlilik, Görünürlük, Değer, Karar, Gerçeklik, Geçerlilik, Değişkenlik (Volume, Velocity, Variety, Visibility, Value, Verdict, Veracity, Validity, Variability)
AI	Yapay Zekâ (Artificial Intelligence)
AUC	Eğrinin Altındaki Alan (Area Under The Curve)
CDR	Çağrı Detay Kaydı (Call Detail Recording)
CNN	Evrişimli Sinir Ağı (Convolutional Neural Network)
DT	Karar Ağacı (Decision Tree)
DCRFD	Dağıtık Yeniden Örnekleme İle Dolandırıcılık Tespiti (Distributed Cluster-based Resampling for Fraud Detection)
DFDDS	Dinamik İmzalar İle Dağıtık Dolandırıcılık Tespiti (Distributed Fraud Detection with Dynamic Signatures)
DIKW	Veri, Bilgi, Öz Bilgi, Hikmet (Data, Information, Knowledge, Wisdom)
DL	Derin Öğrenme (Deep Learning)
FDIC	Görüntüye Dönüştürme İle Dolandırıcılık Tespiti (Fraud Detection with Image Conversion)
FMS	Dolandırıcılık Yönetim Sistemi (Fraud Management System)
FN	Yanlış Negatif (False Negative)
FP	Yanlış Pozitif (False Positive)
GAF	Gramyan Açısal Alanları (Gramian Angular Fields)
GADF	Gramyan Açısal Fark Alanları (Gramian Angular Difference Fields)
GASF	Gramyan Açısal Toplama Alanları (Gramian Angular Summation Fields)
GBT	Gradyan Yükseltme Ağaçları (Gradient-Boosted Trees)

Kısaltmalar	Açıklamalar
Grad-CAM	Gradyan Ağırlıklı Sınıf Aktivasyon Haritalama (Gradient-weighted Class Activation Mapping)
GSM	Mobil İletişim İçin Küresel Sistem (Global System for Mobile Communications)
ML	Makine Öğrenmesi (Machine Learning)
MTF	Markov Geçiş Alanı (Markov Transition Field)
NoSQL	İlişkisel Olmayan Yapılandırılmış Sorgu Dili (Non-relational Structured Query Language)
OVR	Bire-Kalan Lojistik Regresyonu (One-vs-Rest Logistic Regression)
RDD	Esnek Dağıtılmış Veri Kümeleri (Resilient Distributed Datasets)
RF	Rastgele Orman (Random Forest)
RP	Yineleme Grafiği (Recurrence Plot)
ROS	Rastgele Aşırı Örnekleme (Random Over Sampling**)**
RUS	Rastgele Örneklem Azaltma (Random Under Sampling)
SMOTE	Sentetik Azınlık Aşırı Örnekleme Tekniği (Synthetic Minority Over-sampling Technique)
SSE	Karesel Hataların Toplamı (Sum of Squared Errors)
TN	Doğru Negatif (True Negative)
TP	Doğru Pozitif (True Positive)
VoIP	İnternet Protokolü Üzerinden Ses (Voice over Internet Protocol)
XAI	Açıklanabilir Yapay Zekâ **(**Explainable Artificial Intelligence)

1. GİRİŞ

İsabetli kararlara kaliteli bilgi sayesinde ulaşılır. Kaliteyi oluşturan erişilebilir, kullanılabilir ve güvenilir bilgiye ulaşmak ise gerçek dünya problemleri için oldukça zordur. Veri kümeleri sıklıkla; sınıf dengesizliği, dağılım çarpıklığı, gürültü ve eksik özellikler gibi çeşitli düzensizlikleri barındırır. En önemli veri düzensizliklerinden biri olan sınıf dengesizliği, hedef değeri niteleyen etiket verisinin yetersiz olmasından kaynaklanmaktadır. Problem uzayının detaylıca anlaşılmasını engelleyen bu durum, daha az kapsayıcı modellerin geliştirilmesine sebep olur.

Geleneksel veri toplama ve veri analiz etme araç ve tekniklerinin sınırlarını zorlayan büyük veri kavramının literatüre girmesiyle, sınıf dengesizliği de ciddi boyutlara ulaşmıştır. Sınıf dengesizliği konulu çalışmalara bakıldığında, çözüme ulaşmak için farklı amaçların benimsendiği görülmektedir. Mevcut çalışmalar; çoğunluk ve azınlık sınıfları arasındaki dengesiz dağılımı telafi etmek için veri kümesini değiştirmeyi amaçlayanlar [1], çoğunluk sınıfa karşı ön yargıyı azaltmak için geleneksel öğrenme algoritmalarında değişiklik yapanlar [2] ve önceki iki yaklaşımı beraber kullananlar olarak üç başlık altında özetlenebilir [3]. Küçük ve orta ölçekli veri kümelerinde iyi çalışan bu teknikler, büyük veri söz konusu olduğunda doğruluklarını sürdürmekte ve hızlı sonuç üretmekte zorlanmaktadırlar [4]. Ayrıca sınıf dengesizliği problemi veri kümesine ve uygulama alanına bağlı olarak değişiklik gösterdiği için, ideal bir çözüm yöntemi bulunmamakta ve çözümler probleme özgü olarak değişiklik göstermektedir.

Veri toplama sürecinin belirli nedenlerle sınırlı olduğu durumlardan ya da veri edinimi sürecindeki teknik hatalardan kaynaklanan sınıf dengesizliği, pek çok uygulama alanında ortaya çıkmaktadır. Bu alanlardan biri de dolandırıcılık tespitidir. Hem doğası gereği tüm işlemler arasında dolandırıcılık işlemlerinin az olmasından hem de dolandırıcılık etiketinin edinilmesinin maliyetli olmasından kaynaklanan sınıf dengesizliği, dolandırıcılık tespiti çözümlerinin hızlı ve gürbüz olarak çalışmasının önündeki en büyük engellerden biridir.

Sisteme izinsiz girişleri ve sahte etkinlikleri keşfetmek, tanımlamak ve bunları sistem yöneticisine bildirmek amacıyla geliştirilen dolandırıcılık tespiti çalışmalarına bakıldığında, sürece engel olan bu durumların ortadan kaldırılmasının hedeflendiği görülmektedir.

Mevcut çalışmalar; sınıf dengesizliğinin giderilmesi [5], kavram sapmasının engellenmesi [6] ve tespitin çevrim içi gerçekleştirilmesi [7] gibi konulara odaklanmaktadır. Fakat geliştirilen modellerin eğitim sürecini hızlandırmak, karmaşık saldırılara karşı duyarlılığını artırmak, başarısını yükseltmek ve dolandırıcı olmayan sınıfa doğru yanlılığı engellemek için güncel teknik ve teknolojilere ihtiyaç duyulmaktadır.

Bu tez, meşru yani dolandırıcılık şüphesi olmayan işlemlere nazaran çok daha az dolandırıcılık işleminin bulunduğu büyük veri kümelerinden gürbüz dolandırıcılık tespit yöntemleri geliştirebilmeyi amaçlayan üç araştırma sorusu kapsamında gerçekleştirilmiştir. Bu araştırma soruları doğrultusunda, üç yeni yöntem önerilmiştir.

"Kullanıcı davranışlarındaki sapmalar ve bilinen dolandırıcılık işlemleri beraber değerlendirilerek dolandırıcılık tespit yöntemi geliştirilebilir mi?" sorusuna cevaben geliştirilen ilk yöntem, *"Dinamik İmzalar İle Dağıtık Dolandırıcılık Tespiti"* olarak adlandırılmıştır ve kullanıcı davranışlarındaki sapmaları temel alarak dolandırıcılık tespiti gerçekleştirmektedir. Büyük veri analitiği kullanılarak, verideki anormal davranışlar yani şüpheliler tespit edilip, bilinen dolandırıcı davranışları ile karşılaştırılmakta ve muhtemel dolandırıcı faaliyetleri yakalanmaktadır.

"Veri kümesindeki dolandırıcı ve meşru örnek dağılımı değiştirilerek daha iyi tespit performansı elde edilebilir mi?" sorusuna cevaben geliştirilen ikinci yöntem, *"Dağıtık Yeniden Örnekleme İle Dolandırıcılık Tespiti"* olarak adlandırılmıştır ve veri kümesindeki dolandırıcı ve meşru örnek dağılımını değiştirerek daha iyi tespit performansı elde etmeyi amaçlamaktadır. Büyük veri analitiği ile yeniden örnekleme yöntemleri dağıtık analize uygun hale getirilip, sınıflandırıcı için daha iyi bir eğitimin sağlanacağı yeni bir uzayın oluşturulması hedeflemektedir.

"Tekil özelliklerin dolandırıcılık tespitine etkisini modellemek yerine, özelliklerin ikili ilişkisinin etkisini ortaya koyan bir yöntem geliştirilebilir mi?" sorusuna cevaben geliştirilen üçüncü yöntem ise, *"Görüntüye Dönüştürme İle Dolandırıcılık Tespiti"* olarak adlandırılmıştır ve özelliklerin iki boyutlu ilişkilerinin dolandırıcılık tespitine etkisini ortaya koymayı amaçlamaktadır. Derin öğrenme tabanlı bu yöntemde, zamansal işlemler önce kutupsal koordinatlarda ifade edilmiş, sonra da görüntü formuna dönüştürülüp ikili ilişkiler görsel bağlamda çözülerek dolandırıcılık tespit edilmiştir.

Dolandırıcılık tespitinde, yapay zekâ kendini açıklama yeteneğinden yoksun olduğunda karar vermenin hızı, doğruluğu ve etkinliği açısından getirebileceği faydalardan ziyade yanlış karar verme riski daha ağır basabilir. Bu sebeple geliştirilen yöntemler yetenekleri bağlamında, açıklanabilir yapay zekâ açısından değerlendirilmiştir. Dinamik İmzalar İle Dağıtık Dolandırıcılık Tespiti için uzman görüşü alma ve kural çıkarma, Dağıtık Yeniden Örnekleme İle Dolandırıcılık Tespiti için kural çıkarmaya ek olarak özellik önemlerinin çıkarılması ve son olarak Görüntüye Dönüştürme İle Dolandırıcılık Tespiti için ısı haritalarının oluşturulması stratejileri kullanılmıştır. Ek olarak, Görüntüye Dönüştürme İle Dolandırıcılık Tespiti yönteminden yola çıkılarak, *"ilk bakışta anlamlandırılması zor olan GAF görüntüleri ile oluşturulmuş kara kutu modelleri için yapay zekâ açıklanabilirliği sağlanabilir mi?*" sorusuna cevaben geliştirilen ve *"GAF için Grad-CAM"* olarak adlandırılan, yeni bir açıklanabilir yapay zekâ yaklaşımı geliştirilmiştir. Önerilen yaklaşım, zamansal ilişkilerden elde edilen ve ilk bakışta anlamlandırılması zor olan görüntüler için açıklanabilirlik sağlamaktadır.

Bu tez kapsamında, büyük veri ve derin öğrenme yaklaşımları bağlamında güncel teknik ve teknolojiler kullanılarak yeni dolandırıcılık tespiti yöntemleri geliştirilmiştir. Önerilen yöntemler, telekom dolandırıcılığına ve kredi kartı dolandırıcılığına ait veri kümeleri üzerinde uygulanmıştır. Her bir yöntem; yeteneklerine uygun olarak ayrı ayrı ele alınmış, test edilmiş, literatürdeki benzer çalışmalar ile karşılaştırılmış, üstünlükleri gösterilmiştir ve açıklanabilir yapay zekâ açısından değerlendirilmiştir.

Bu tez, Giriş Bölümü'nü izleyen yedi bölümden oluşmaktadır.

İkinci bölümde; veri, veri bilimi, veri bilimi sürecini zorlaştıran veri düzensizlikleri, büyük veri, büyük veri analitiği, yapay zekâ ve açıklanabilir yapay zekâ gibi temel terim ve kavramlara açıklık getirilmiştir.

Üçüncü bölümde; literatürdeki çalışmalar incelenmiştir. Ele alınan problem için literatürde net bir çalışma alanı bulunmadığından, mevcut çözümler; sınıf dengesiz veri analizi, dolandırıcılık tespiti ve sınıf dengesiz dolandırıcılık tespiti için büyük veri analizi olarak üç ayrı bağlamda değerlendirilmiştir.

Dördüncü bölümde; MapReduce mimarisine uygun olarak geliştirilen: Dinamik İmzalar İle Dağıtık Dolandırıcılık Tespiti ve Dağıtık Yeniden Örnekleme İle Dolandırıcılık Tespiti olarak isimlendirilen iki yöntem detaylıca açıklanmış ve elde edilen sonuçlar karşılaştırmalı olarak değerlendirilmiştir.

Beşinci bölümde; derin öğrenme tabanlı bir yöntem olan Görüntüye Dönüştürme İle Dolandırıcılık Tespiti yaklaşımı ayrıntılı bir şekilde ele alınmış ve elde edilen sonuçlar karşılaştırmalı olarak değerlendirilmiştir.

Altıncı bölümde; önerilen üç yöntemin çıktıları, açıklanabilir yapay zekâ açısından ele alınmıştır. Ek olarak, GAF için Grad-CAM adlı yeni bir açıklanabilir yapay zekâ yaklaşımı geliştirilmiş ve Beşinci bölümde önerilen yöntem üzerinde örneklendirilerek açıklanmıştır.

Son olarak yedinci bölümde ise; yapılan tüm çalışmalar özetlenmiş, sonuçlar sunulmuş, yorumlanmış ve gelecek çalışmalar hakkında değerlendirmeler yapılmıştır.

2. BÜYÜK VERİ VE YAPAY ZEKÂ

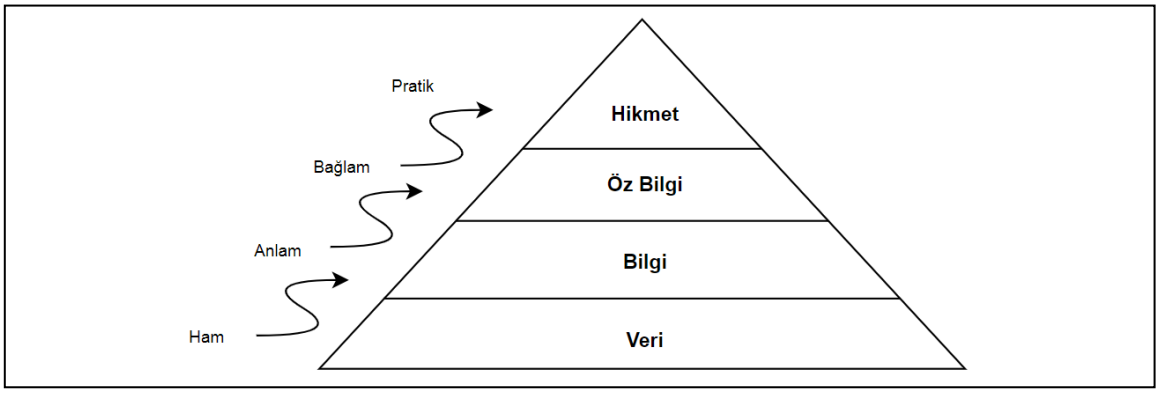
Toplumda devrimci bir değişim yaşanmaktadır. Küçük yerel şirketlerden küresel işletmelere kadar herkes, veri varlıklarını sayısallaştırma ve veri odaklı olma potansiyelini fark etmeye başlamıştır. Analitik, makine öğrenmesi, yapay zekâ tekniklerini kullanarak ve veri bilimini bir disiplin olarak ele alarak verimin nasıl artırılacağı ve büyük verinin yeni değerlere nasıl dönüştürüleceği güncel araştırma konuları haline gelmiştir. Bu yeni teknolojilerin kullanılması şirketlerin operasyonlarını basitleştirmelerine ve maliyetlerini düşürmelerine yardımcı olacak olsa da, geliştirilen yaklaşımların veri bilimi yatırımı için doğru hale getirilmesi konusu oldukça zordur. Bununla beraber esas zorluk, değer yaratmak için veri biliminin nasıl kullanacağı anlamaya çalışıldığında ve bu anlayışın yürütülebilir stratejilere yerleştirilmesi sırasında ortaya çıkmaktadır. Çünkü gerçek dünyada veri çeşitli kaynaklardan ve farklı formatlarda elde edilir, düzensizdir ve dinamik olarak değişmektedir. Ek olarak, verimli bir veri bilimi süreci işletildiği düşünülse dahi, üretilen modeller sonucunda elde edilen değerlerin doğrulanması için, modellerin açıklanabilirliğinin ortaya koyulması gerekmektedir.

Bu bağlamda, bu bölüm üç alt bölümde sunulmuştur. İlk bölümde; veri kavramına ve verinin sistematik analizini ifade eden veri bilimi sürecine açıklık getirilmiştir. Daha sonra, bu süreci olumsuz olarak etkileyen veri düzensizlikleri ele alınmıştır. İkinci bölümde, veri biliminin kaynağı olan büyük veri, beraberinde getirdiği zorluklar, büyük veri analitiği ve büyük veri araçları hakkında detaylı bilgi sunulmuştur. Üçüncü bölümde ise büyük veri, yapay zekâ ve makine öğrenmesi kavramları arasındaki ilişki ortaya koyulmuştur. Son olarak, büyük veri üzerinde makine öğrenmesi yöntemleri kullanılarak geliştirilen modellerin öğrendiği örüntülerin daha net ortaya koyulmasını sağlayan açıklanabilirlik konusu ele alınmıştır.

2.1. Veri ve Veri Düzensizlikleri

Veri yönetimi veya bilgi sistemleri gibi hususlarda belirsiz ya da çelişkili tanımların çeşitliliği sebebiyle ilk olarak DIKW (Data, Information, Knowledge, Wisdom) hiyerarşisinin [8] netleştirilmesi gerekmektedir. Şekil 2.1'de verildiği üzere, hiyerarşi; *"veri, bilgi, öz bilgi, hikmet*" varlıklarını birbirlerine göre bağlamsallaştırmak ve düşük seviyedeki bir varlığın daha yüksek seviyedeki bir varlığa dönüştürülmesine ilişkin süreçleri

tanımlamak için kullanılmaktadır. *"Veri"* nesnelerin, olayların ve ortamların özelliklerini temsil eden semboller, *"bilgi"* açıklamaları ya da kim, ne, ne zaman, kaç tane gibi soruların cevapları, *"öz bilgi"* talimatlarla ya da deneyimlerle çıkarılan teknik bilgi, son olarak *"hikmet"* bilinenleri ve en iyi olanı dikkate alarak en uygun davranışları eyleme geçirme kapasitesi şeklide ifade edilmektedir [9].



Şekil 2.1. DIKW hiyerarşisi [8]

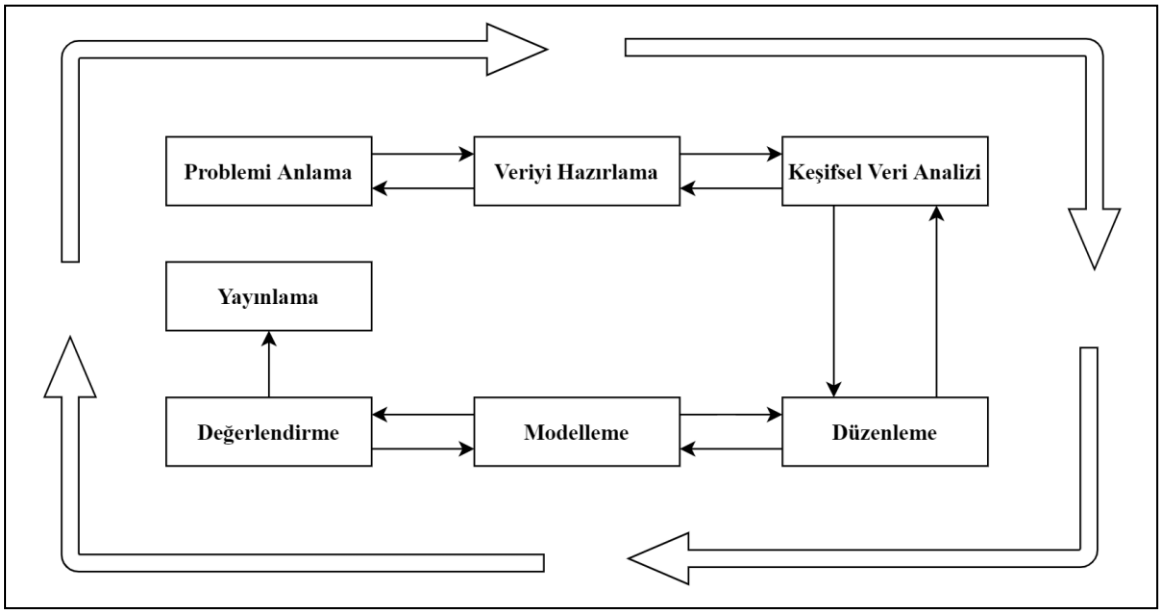
Verinin bilgeliğe dönüştürülmesini amaçlayan süreç; veri tabanlarından bilgi keşfi, veri madenciliği, veri analizi, veri analitiği gibi isimlendirmelerden sonra zaman içerisinde "*veri bilimi"* (data science) kavramına evrilmiştir [10]. Veri bilimindeki *"veri"*, birkaç sayısal gözlemden oluşan basit bir diziden, binlerce değişkenli milyonlarca gözlemden oluşan karmaşık bir matrise kadar değişen aralığı, *"bilim"* ise yöntemlerin kanıta dayalı olduğunu ve deneysel bilgiye dayandığını ifade etmektedir [11]. Kısaca veri bilimi, en uygun modelleri ortaya çıkarmak için verinin bilimsel bir çerçevede sistematik olarak analiz edilmesidir. Disiplin perspektifinden bakıldığında ise veri bilimi, verileri öngörülere ve kararlara dönüştürmek amacıyla kayıtları ve ortamları incelemek için istatistik, bilişim, bilgi işlem, iletişim, yönetim ve sosyolojiyi sentezleyen ve geliştiren yeni bir disiplinler arası alandır [12].

Birçok parçacığın heterojen bir topluluğu olan veri bilimi dünyasında, "*veri bilimcisi"* büyük miktardaki biçimsiz veriye yapı getirip analizini mümkün kılan, zengin veri kaynaklarını tanımlayıp potansiyel olarak eksik veri kaynaklarıyla birleştiren, zorlukların değişmeye devam ettiği ve verinin akmayı asla durdurmadığı rekabetçi bir ortamda yenilik üreten kişilerdir [13]. İyi bir veri bilimcisi; bilgisayar bilimleri, istatistik ve matematik gibi bilimsel disiplinlerde uzmanlığa, problemi test edilebilecek ve açık bir hipoteze dönüştürecek

meraka, mevcut veriyi senaryolaştıracak iletişim becerisine ve son olarak soruna farklı ve yaratıcı bakış açıları sunabilecek zekâya sahip olmalıdır [14].

Veri bilimi taksonomisinde sıkça kullanılan diğer kavramlara da açıklık getirmek gerekirse, *"veri kümesi"* tanımlanmış bir yapıya sahip veri topluluğunu, *"kayıt"* veri kümesindeki tek bir örneği, *"öznitelik"* veri kümesinin tek bir özelliğini, son olarak *"etiket"* tüm girdi özelliklerine göre tahmin edilecek özniteliği ifade eder [15].

Veri kümesindeki ilişkilerin ve örüntülerin yöntemsel keşfi, *"veri bilimi süreci"* olarak bilinen bir dizi yinelemeli etkinlik tarafından gerçekleştirilmektedir [16]. Veri bilimi uygulamalarının evrimi boyunca, çeşitli akademik ve ticari kurumlar tarafından süreç için farklı çerçeveler ortaya koyulmuştur. Veri tabanlarında bilgi keşfinde kullanılan seçim, ön işleme, dönüşüm, veri madenciliği, yorumlama ve değerlendirme adımlarını içeren KDD (Knowledge Discovery in Databases) [17], veri madenciliği için sektörler arası standart süreç olan CRISP-DM (Cross Industry Standard Process for Data Mining) [18], altı sigma uygulamasında kullanılan tanımlama, ölçme, analiz etme, geliştirme ve kontrol adımlarından oluşan DMAIC (Define, Measure, Analyze, Improve, Control) [19], ve son olarak örnekleme, keşfetme, değiştirme, modelleme ve değerlendirme, aşamalarından oluşan SEMMA (Sample, Explore, Modify, Model, Assess) [20] en önemli çerçevelerden birkaçıdır.



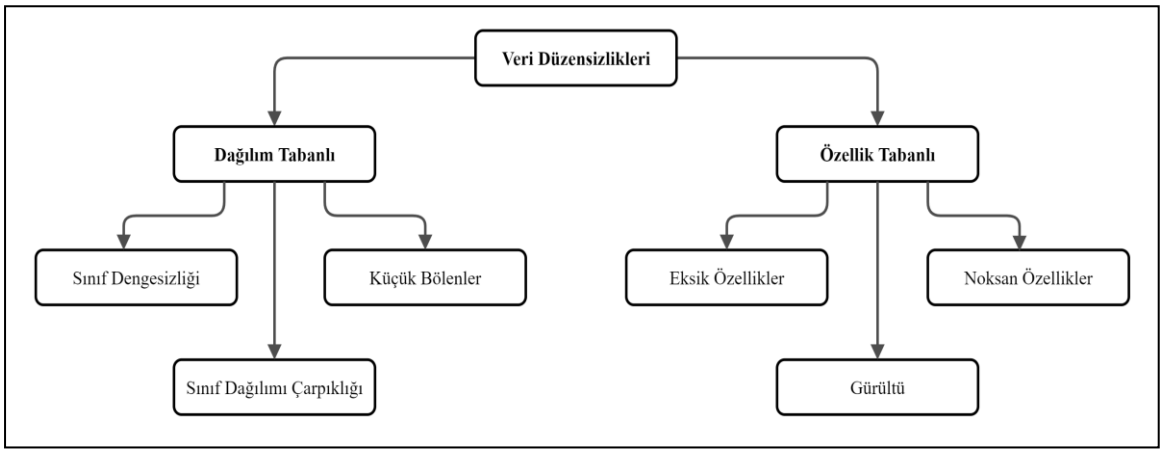
Şekil 2.2. Veri bilimi süreci

Herhangi bir süreç çerçevesinde olduğu gibi, veri bilimi süreci de en uygun çıktıyı üretmek için belirli bir dizi görevin yürütülmesini önerir. Ancak, Şekil 2.2'de görselleştirildiği gibi, veri bilimi sürecindeki adımlar doğrusal değildir ve birçok döngüye girmeli, adımlar arasında ileri geri gitmeli ve bazen de veri bilimi problemini yeniden tanımlamak için ilk adıma geri dönmelidir. Veri bilimi süreci uyarlanabilir ve yinelemeli olarak aşağıda belirtilen 7 aşamadan oluşur [21].

1. Problemi Anlama: İlk olarak, proje hedefleri açıkça belirtilir ve bu hedefler veri bilimi kullanılarak çözülebilecek bir sorunun kesin ifadesine dönüştürülür.
2. Veriyi Hazırlama: Veri bilimi sürecinin en yoğun emek harcanan aşamasıdır. Aykırı değerlerin ve bunlar hakkında ne yapılacağının belirlemesi, verilerin dönüştürülmesi ve standartlaştırılması gibi işlemleri içerir.
3. Keşifsel Veri Analizi: Bu aşamada, öznitelikler arasındaki tek değişkenli ve çok değişkenli ilişkileri keşfetmek, mevcut öznitelikleri gruplandırmak ya da yeni öznitelikler türetmek gibi işlemler gerçekleştirilir.
4. Düzenleme: Veri bölümlerinin gerçekten rastgele olduklarından emin olmak için çapraz doğrulama, sınıf dağılımlarını dengeleme ve karşılaştırmalar için temel performans oluşturma gibi hazırlıkların yapıldığı aşamadır.
5. Modelleme: Veride gizlenmiş ilişkileri ortaya çıkarmak için uygun modelleme algoritmalarının seçildiği, uygulandığı, ince ayar yapıldığı ve temel modellerden daha iyi performans göstermesinin sağlandığı, veri bilimi sürecinin ana aşamasıdır.
6. Değerlendirme: Modelin problemi anlama aşamasında belirlenen hedeflere ulaşıp ulaşmadığı kontrol edilir ve model kurulum aşamasındaki temel performans ve maliyet ölçümleri doğrultusunda en iyi model belirlenir.
7. Yayınlama: Elde edilen sonuçlar hem üstünlükleri hem de kısıtları ile raporlanarak süreç sonlandırılır. Oluşturulan modelin, ilgili bir sorunun araştırılmasına girdi olarak hizmet ettiği durumlarda süreç tekrar edilir.

Veri bilimi sürecinin etkin ve etkili bir şekilde gerçekleştirilerek değer üretilmesinin ön koşulu, doğru ve yüksek kaliteli veri kümesine dayanmaktadır. Çeşitli veri düzensizliklerini barındıran düşük kaliteli veri, düşük kaliteli bilginin çıkarılmasına yol açmaktadır. *"Çöp girerse çöp çıkar"* diye adlandırılan bu durum, toplanan verilerin doğru, geçerli ve tam olması gerektiğini aksi takdirde veri analizi, uygulamalar veya iş sürecinin güvenilir olmayacağını özetlemektedir [22].

Veri kalitesi; erişilebilirlik, kullanılabilirlik, güvenilirlik, uygunluk ve sunum olmak üzere beş boyuttan oluşur [23]. Erişilebilirlik veri ve bilgiye erişimin kolaylık derecesini, kullanılabilirlik verinin yararlı olup olmadığını, güvenilirlik verinin doğru ve tutarlı olduğunu, uygunluk veri ile talepler arasındaki ilişki derecesini, sunum ise veriler için geçerli açıklama yöntemlerini ifade eder. Bu boyutlara aykırı bir şekilde gelişen veri düzensizlikleri, veri kalitesini düşürmekte, veri bilimi sürecini zorlaştırmakta, hatta hatalı modellin geliştirilmesine sebep olmaktadır.



Şekil 2.3. Veri düzensizlikleri türleri

Veri düzensizlikleri Şekil 2.3'de görselleştirildiği gibi dağılım tabanlı ve özellik tabanlı olarak ortaya çıkmaktadır [24], [25]. Dağılım tabanlı veri düzensizlikleri nadir olaylar, anormal örüntüler veya veri toplama sırasındaki kesintilerden kaynaklanırken, özellik tabanlı düzensizlikler gürültü, arıza, gizlilik ya da iletim hatası gibi nedenlerle değerlerin bozulmasından kaynaklanmaktadır. Düzensizlikler arasında her ne kadar ayrıştırma yapılsa da unsurlardan birinin varlığının diğerinin varlığı üzerinde önemli etkileri olabilmektedir.

Sınıf dengesizliği, bir veya daha fazla sınıfın veri kümesinde yetersiz temsil edildiği dağılım tabanlı veri düzensizliğinin çok yaygın bir biçimidir. Sınıflar arasındaki dengesiz dağılımı telafi etmek için; çoğunluk sınıf örneklerini azaltmak veya azınlık sınıf örneklerini artırmak ya da çoğunluk sınıfına yönelik ön yargıyı azaltmak için geleneksel öğrenme algoritmalarında değişiklikler yapmak gibi yöntemler kullanılmaktadır [26].

Küçük bölenler sorunu, sınıflar içinde az temsil edilen alt kavramlar olduğunda ortaya çıkar. Bu sorun, veri kümesinin artırılmaya çalışılmasıyla, daha etkin metriklerin kullanılmasıyla,

budamanın devre dışı bırakılmasıyla ya da yinelemeli algoritmaların kullanılmasıyla giderilmeye çalışılır [27].

Sınıf dağılımı çarpıklığı, sınıfların koşullu dağılımlarının birbirinden oldukça farklı olmasından kaynaklanır ve özellikle sınıflar arasında çakışma olması durumunda, dağılımların yerel özelliklerine bağlı olarak ciddi sorunlar ortaya çıkabilir. Sınıf çarpıklığı ya da sınıf çakışması genellikle sınıfların sınır bölgelerinde meydana geldiği için bu düzensizliklerin çözümünde komşuluk tabanlı yaklaşımlar kullanılmaktadır [28].

Gürültü, veri kümesindeki öznitelik değerlerinin bozulmasından, örnekler için nitelendirmenin yetersiz olmasından ya da örneklerin hatalı etiketlenmesinden kaynaklanmaktadır. Modern sınıflandırıcıların çoğu büyük değişiklikler yapılmadan gürültü varlığında etkili bir şekilde öğrenme gerçekleştirebilmelerine rağmen, genellikle gürültülü örnekleri kaldırmayı veya düzeltmeyi amaçlayan ön işlemler gerçekleştirilerek bu problem giderilmeye çalışılır [29].

Eksik özellik problemi, tüm örneklerin tam olarak gözlemlenemediği ya da farklı kaynaklardaki veri kümelerinin birleştirildiği zamanlarda ortaya çıkan bir durumdur. Eksik özelliklerin giderilmesi için ya bu veri noktaları atılır ya da tahminleme yöntemleriyle bu eksiklikler doldurulur [30].

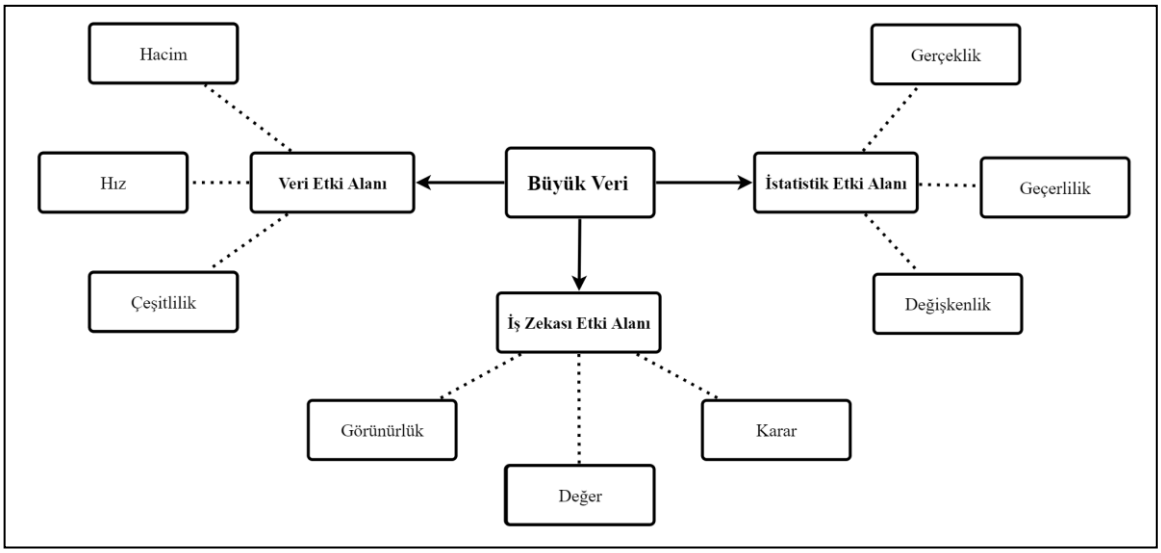
Noksan özellik, verinin doğası gereği bazı özellikler için bazı veri noktalarının tanımsız olduğu durumlarda ortaya çıkar. Bu sorunun giderilmesi için ön işlemlerin gerçekleştirilmesi anlamsız olduğundan, veri kümesinin mevcut haliyle doğrudan analiz edilmesini sağlan değiştirilmiş algoritmalar veya melez yöntemler tercih edilmektedir [31].

Veri bilimi süreci, tüm düzensizliklere rağmen sadece veri ile başa çıkmak anlamına gelmez. Veri bilimi süreci, değerini verilerin kendisinden alır ve sonuç olarak daha fazla veri oluşturur yani veri bilimi veri ürünlerinin oluşturulmasını sağlar [32]. Bu ürünleri oluşturmak da çoğunlukla analitik aşamasının gerçekleştirim kalitesine, yani verinin planlanan amaca hizmet etmesinin nitelikli olarak sağlanmasına bağlıdır. Veri kümesinin boyutu arttıkça, geleneksel veri analizini bile zorlu kılan bu sorunların ortaya çıkma ihtimali artmaktadır. Bu sebeple, daha detaylı planlama süreçleri ve karmaşık mekanizmalar gerekmektedir.

2.2. Büyük Veri ve Büyük Veri Analitiği

Web, sosyal medya, mobil cihazlar ve sensör ağların çoğalması, depolama ve bilgi işlem kaynaklarının maliyetlerinin düşmesi ile birlikte yaygın eylemlerin ve iletişimin her yerde ve her geçen gün artması "*büyük veri"* (big data) olarak adlandırılan dijital kayıtların üretilmesine yol açmıştır. Büyük veri: geleneksel süreçlerin veya araçların etkili bir şekilde yakalama, depolama, yönetme, analiz etme, görselleştirme ve kullanma kapasitesinin ötesine ulaşan heterojen ve otonom kaynaklardan gelen yüksek boyutlarda, çeşitli türde ve dinamik ilişkilerde olan veri kümelerini kapsamaktadır [33].

Büyük veri, hem yol açtığı zorlukları hem de sağladığı avantajları anlamaya yardımcı olan belirli özellikler ile tanımlanmaktadır. Bu özellikler hacim, hız ve çeşitlilik kelimelerinin İngilizce baş harflerini ifade eden *"3V"* (Volume, Velocity, Variety) ile simgelenmektedir [34]. Zaman içerisinde bu tanımlama, Şekil 2.4'de verildiği gibi üç farklı etki alandaki üçer özelliği ifade eden *"*3^2V*"* (Volume, Velocity, Variety, Visibility, Value, Verdict, Veracity, Validity, Variability) haline dönüşmüştür [35].



Şekil 2.4. Büyük verinin etki alanı özellikleri

Veri etki alanı örüntü aramayı, iş zekâsı etki alanı tahminlerde bulunmayı ve istatistik etki alanı varsayımlarda bulunmayı işaret etmektedir. Bu konsept dahilinde ele alınan özelliklerden: hacim (volume) artan veri akışı veya kümülatif veri miktarını, hız (velocity) etkileşimler sonucu süratle oluşan veriyi ve dinamik analiz ihtiyacını, çeşitlilik (variety) veri

formatlarının ve veri yapılarının farklılığını, görünürlük (visibility) öngörü ve kavrayış sağlamayı, değer (value) uzun vadeli fayda veya stratejik kazanç elde etmeyi, karar (verdict) potansiyel veya olası seçimler yapmayı, gerçeklik (veracity) veri belirsizliğinin çözülmesini, geçerlilik (validity) verinin kalitesinin doğrulanmasını, son olarak değişkenlik (variability) veri karmaşıklığı ve varyasyonunun sonuçlarını ifade etmek için kullanılmaktadır [36].

Büyük verinin tarihi, ilgilenilen veri boyutu açısından değerlendirildiğinde; büyüklük sırasıyla genişleyen boyut sınırlamaları ile daha büyük veri kümelerini daha verimli bir şekilde depolama ve yönetme yeteneği, beş gelişim süreci içerisinde incelenmektedir [37].

- Megabayt - Gigabayt: 1970'ler ve 1980'ler büyük veri sorunun ortaya çıktığı en erken dönemdir. Dönemin ihtiyacı, veriyi barındırmak ve iş analizleri ya da raporlama için ilişkisel sorgular yürütmekti. Sorunları çözmek için entegre donanım ve yazılım içeren veri tabanı araştırmaları başlatıldı.
- Gigabayt - Terabayt: 1980'lerin sonunda veri hacminin artması, tek bir büyük bilgisayar sisteminin depolama ve işleme yeteneklerinin ötesinde geçilmesine neden oldu. Bu sorun, belleğin ya da diskin paylaşıldığı çeşitli paralel veri tabanlarının geliştirilmesiyle çözüldü.
- Terabayt - Petabayt: 1990'ların sonunda Web 1.0'ın hızlı gelişimi tüm dünyayı internet çağına sürüklerken, yarı yapılandırılmış veya yapılandırılmamış web sayfalarının ortaya çıkmasına sebep oldu. 2000'lerde web içeriğini indekslemek ve sorgulamak amacıyla, ölçeklendirilebilir ve genişletilebilir Google Dosya Sistemi ve MapReduce programlama modeli oluşturuldu. 2000'lerin ortalarında ise şema içermeyen, hızlı, yüksek düzeyde ölçeklenebilir ve güvenilir olan NoSQL (Non-relational Structured Query Language) veri tabanları ortaya çıkmaya başladı.
- Petabayt - Eksabayt: Mevcut gelişim ve değişim yönelimleri altında, büyük şirketler tarafından saklanan ve analiz edilen veriler eksabayt büyüklüğüne ulaşmıştır ve hem daha büyük veri kümeleriyle başa çıkmak hem de veri yönetimi ve analiz yöntemi için devrimci teknolojileri geliştirmek için çalışmalar hala devam etmektedir.

Büyük veri; organizasyon performansının artırılması, iş süreçlerinin iyileştirilmesi, ürün ve hizmetlerin yenilenmesi veya pazarın geliştirilmesi gibi doğrudan büyük veri analitiğinin benimsenmesinden kaynaklanan fonksiyonel değerler ya da büyük veri analitiğinin sinyal

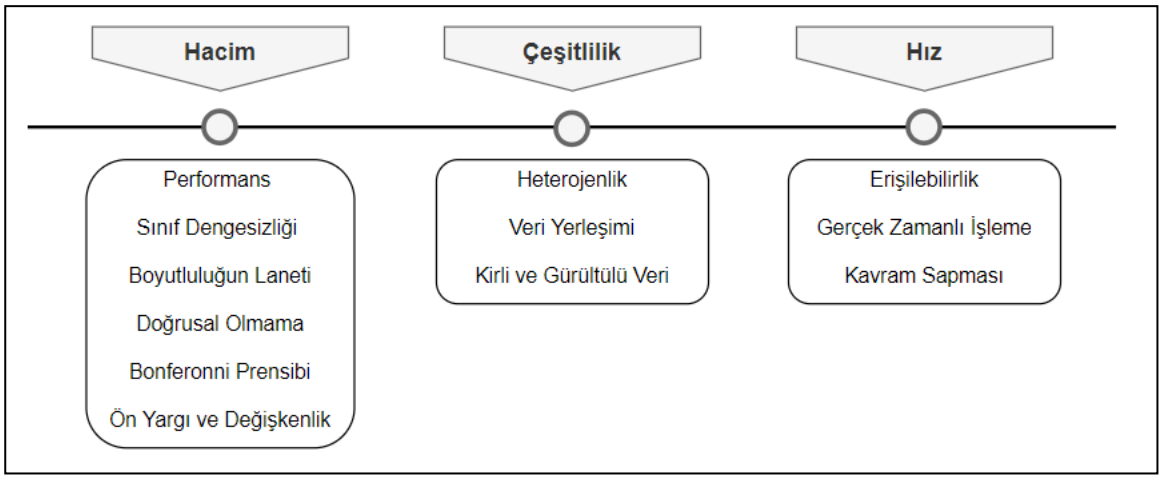
etkisinden kaynaklanan sembolik değerler yaratmaktadır [38]. Çeşitli organizasyonların çeşitli iş ve bilimsel uygulamada temel operasyonlarını desteklemek için gerçekleştirdikleri büyük veri analitiği uygulamaları aşağıda verildiği şekilde örneklendirilebilir [39].

- Üretim: Ürün kalitesi takibi, arz planlaması, hata takibi, verimliliğin artırılması, çıktı tahmini, yeni ürünlerin geliştirilmesi, yeni süreçlerin test edilmesi, vb.
- Satış: Market sepeti analizi, kampanya yönetimi, müşteri sadakat programları, tedarik zinciri yönetimi, tüketici grupları oluşturma, kişiselleştirilmiş reklam, talep analizi, vb.
- Sağlık: Hastalık teşhisi, ilaç keşfi, hasta bakım kalitesinin artırılması, tıbbi cihaz ve ilaç tedarik zinciri yönetimi, tedavinin kişiselleştirilmesi, vb.
- Eğitim: Öğrenci gelişimi takibi, öğretmen performansının değerlendirilmesi, müfredatın güncellenmesi, davranışsal sınıflandırma, can sıkıntısı tespiti, vb.
- Finans: Dolandırıcılık tespiti, kredi puanlaması, risk yönetimi, müşteri davranış analizi, anormal ticaret örüntüsü tespiti, ticari görünürlük, uyumluluk analizi, vb.
- Telekomünikasyon: Çağrı detay kayıtlarının analizi, dolandırıcılık tespiti, müşteri kaybının önlenmesi, ağ güvenliğinin sağlanması, ağ eniyileme, vb.
- Taşımacılık: Trafik kontrolü, rota planlama, akıllı ulaşım sistemleri, trafik sıkışıklığı yönetimi, yakıt tasarrufu, turizm seyahat düzenlemeleri, vb.
- Tarım: Ürün verimliliğini artırma, çevresel zorlukları yönetme, tasarruf etme, zirai ilaç kullanımının en uygun seviyeye getirilmesi, tedarik zinciri yönetimi, tahmine dayalı modelleme, vb.
- Dijital Medya: Tıklama akışı analizi, sosyal çizge analizi, kötüye kullanım tespiti, tıklama dolandırıcılığı önleme, reklam hedefleme, kullanıcı bölütleme, hedef kitle analizi, vb.
- Bilimsel Araştırma: Veriden bilgi elde eden yeni yöntemler geliştirme, veri yönetim modelleri önerme, topluluklara veri sunmak için yeni altyapılar oluşturma, vb.

Büyük veri analitiği uygulamaları mevcut operasyonlar için verimlilik sağlarken yeni ürün ve hizmetlerin de geliştirilmesini sağlamaktadır. Fakat büyük veri analitiği; veri, süreç ve yönetim olmak üzere üç boyutta ele alınan zorlukları da beraberinde getirmektedir [40]. Veri zorlukları; hacim, çeşitlilik, hız ve kalite gibi verinin kendi özelliklerinden kaynaklanır. Süreç zorlukları; verilerin nasıl yakalanacağı, bütünleştirileceği, dönüştürüleceği, analiz için doğru modelin nasıl seçileceği ve sonuçların nasıl doğrulanacağı gibi bir dizi teknik ile

ilgilidir. Yönetim zorlukları ise veriye erişim ve denetim aşamalarında ortaya çıkan gizlilik, güvenlik, paylaşım ve etik gibi konuları kapsar.

Analitik sürecini zorlaştıran ya da engelleyen bu zorluklar temelde; paralelliğin algoritmik olarak sağlanması, çeşitli belleklerin bir arada kullanımı, büyük ölçekli dinamik veri güncellemeleri ve akışları ile başa çıkabilmek, kaynak kısıtlamaları altında makine öğrenmesi gerçekleştirmek, hesaplamaların gürbüzlüğünü sağlamak ve enerji tüketimini azaltmak gibi başlıklar altında ortaya çıkmaktadır. Bu zorluklar büyük veriyi tanımlamaya yarayan en temel bileşenler olan; hacim, çeşitlilik ve hız açısından değerlendirildiğinde Şekil 2.5'de verilen zorluklar ortaya çıkmaktadır.



Şekil 2.5. Büyük veri kaynaklı zorluklar

"*Hacim*" kaynaklı zorluklar, ya kayıtların sayısına göre dikey olarak ya da özniteliklerin sayısına göre yatay olarak gerçekleşen aşırı büyümenin etkisiyle ortaya çıkmaktadır. Performans, sınıf dengesizliği, boyutluluğun laneti, doğrusal olmama, Bonferonni prensibi, önyargı ve değişkenlik olarak sınıflandırılabilir [41]. Performans sorunu, artan veri boyutunun hesaplamaların gerçekleştirilmesindeki zaman ve uzay karmaşıklığını artırarak çok büyük veri kümelerinde, algoritmaların kullanılamaz hale geldiği durumdur. Sınıf dengesizliği; bazı sınıfların çok sayıda örnekle, bazı sınıfların ise çok az sayıda örnekle sunulduğu problemlerin, makine öğrenmesi algoritmalarında sınıfların eşit olarak temsil edildiği varsayımını yıkmasıyla ortaya çıkar. Boyutluluğun laneti, boyutsallık arttıkça algoritmanın öngörme yeteneğinin ve etkinliğinin azalmasından kaynaklanır. Doğrusal olmama problemi; birçok metriğin ve tekniğin geçerliliği doğrusallık varsayımına

dayanmasına rağmen, genellikle değişkenler arasında doğrusal bir ilişki olmadığı büyük veride ortaya çıkan bir durumdur. Bonferonni prensibi, belirli miktardaki veri içinde belirli bir olay türü arandığında bu olayın bulunma olasılığının yüksek olduğunu savunan fikirdir. Fakat sık olmamakla birlikte, bu oluşumlar sahte ilişkilerden kaynaklanabilir. Oldukça büyük bir hacimde de, çoğu ilişki sahte olma eğilimindedir. Ön yargı ve değişkenlik, analiz ve tahmin sağlamak için genelleştirme yapan makine öğrenmesi algoritmalarının genelleme hatasını ölçmek için kullanılır. Yüksek yanlılığa sahip bir modelin eksik öğrenme olasılığı artarken, yüksek varyans da aşırı öğrenmeye yol açabilir. Veri hacmi arttıkça, bu durumlar yeni veriler için genelleme yapılmasını zorlaştırmaktadır.

"*Çeşitlilik*" kaynaklı zorluklar, yalnızca veri kümesinin içerdiği veri türlerinin yapısal varyasyonunu değil, aynı zamanda anlamsal yorumunu ve kaynaklarını da kapsamaktadır [42]. Heterojenlik, veri yerleşimi, kirli ve gürültülü veri olarak sınıflandırılabilir. Heterojenlik problemi, çeşitli kaynaklardan gelen çeşitli verilerin birleştirilmesini içeren büyük veride tür, biçim, model ve anlam bilim açısından bulunan farklılıkların, verinin algoritmaya adaptasyon sürecini zorlaştırması durumudur. Birçok öğrenme algoritması, işlenen verinin tamamen bellekte veya bir diskteki tek bir dosyada tutulduğu varsayımına dayanmasına rağmen, büyük veri farklı fiziksel konumlarda bulunan çok sayıda dosyaya dağıtılır. Bu bağlamda ortaya çıkan veri yerleşiminden kaynaklanan sorunlar, işlem gecikmesine ve ağ trafiğine neden olmaktadır. Kirli ve gürültülü veri, farklı ve bilinmeyen kaynaklardan gelen büyük veride ölçüm hataları, aykırı değerler ve eksik değerler olduğunu işaret etmektedir.

"*Hız*" kaynaklı zorluklar, hem verinin üretilme hızının hem de analiz edilmesi gereken hızın sebep olduğu çıkmazlardır [43]. Erişilebilirlik, gerçek zamanlı işleme ve kavram sapması olarak sınıflandırılabilir. Erişilebilirlik, birçok makine öğrenmesi yaklaşımı öğrenme başlamadan önce tüm veri kümesinin mevcut olduğu varsayımına dayanmasına rağmen, sürekli veri akışı olan büyük veri için yeniden öğrenme ihtiyacı olmayan ve uyarlama becerisine sahip olan algoritmaların gerekmesi durumudur. Gerçek zamanlı işleme, makine öğrenmesi modellerinin yeni veri geldikçe güncellenmesi ihtiyacını ifade eden erişilebilirlikten farklı olarak hızlı gelen veriyi gerçek zamanlı veya gerçek zamana yakın işleme gereksinimini ifade eder. Büyük veri sabit değildir, yani mevcut verinin gelecekteki veriyle aynı dağılımı takip edip etmediği belirlenemez. Bu bağlamda ortaya çıkan kavram

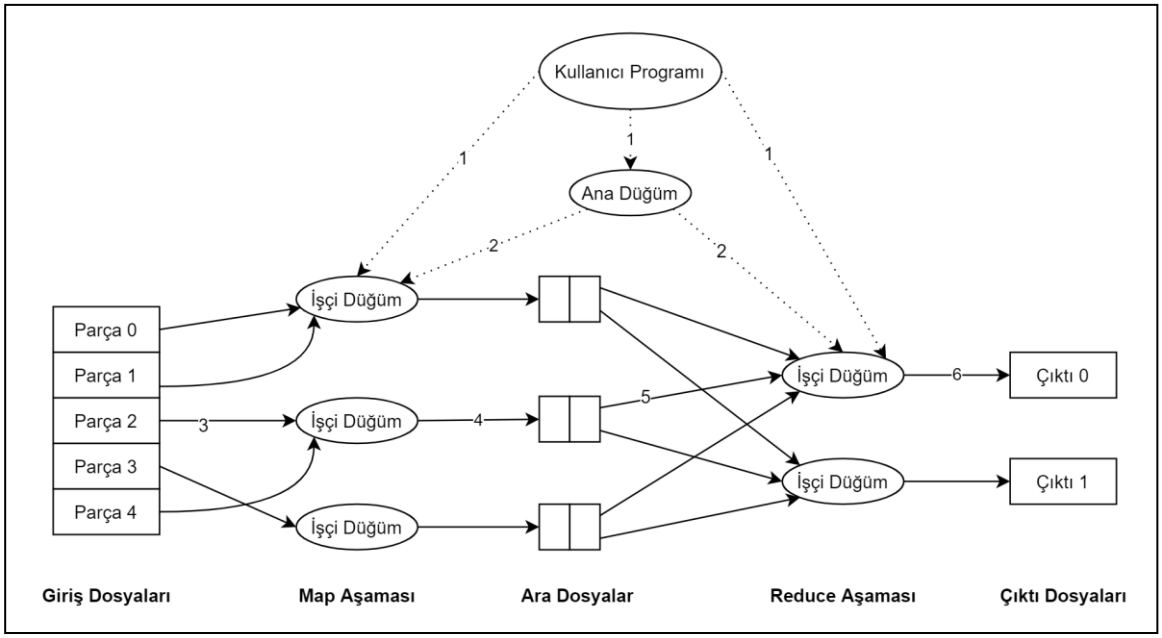
sapması sorunu, modellerin verinin koşullu dağılımındaki değişiklikleri yansıtmayan nüshaları kullanılarak oluşturulduğunda ortaya çıkar.

Kaotik veri kümelerinden faydalı bilgiler elde etmeyi sağlayan büyük veri analitiği projelerinin %60'ı ne yazık ki tüm süreci tamamlayamadan terk edilmekte ve hayal kırıklığı yaratan sonuçlar vermektedir [44]. Bu sebeple; mevcut kaynaklar, beklenen geri dönüşler, işlevsel ekosistem üzerindeki potansiyel etkiler, fırsat maliyeti ve riskler dâhil olmak üzere stratejik planlamayı mümkün kılmak için büyük veri analitiğinin temel unsurlarının hem negatif hem de pozitif etkileri doğrultusunda net olarak değerlendirilmesi gerekmektedir. Büyük verinin potansiyel değerini açığa çıkarmak için ise yüksek hacimli yapılandırılmış ve yapılandırılmamış veri kümelerini analiz eden verimli süreçlere ve yöntemlere ihtiyaç duyulmaktadır.

Geleneksel hesaplamaların çoğu kavramsal olarak doğrusaldır ve büyük veri üzerindeki hesaplamaların makul bir sürede bitirilmesi için pek çok makineye dağıtılması gerekmektedir. Küme olarak adlandırılan büyük veri donanım birimi, her biri düğüm olarak tanımlanan birçok makinenin kaynağını bir araya toplayarak tek bir bilgisayarmış gibi kullanma imkânı sağlamaktadır [45].

Çok büyük boyutlu veride küme üzerinde, hesaplamanın nasıl paralelleştirileceği, verilerin nasıl dağıtılacağı ve hataların nasıl ele alınacağı gibi sorunlarla başa çıkabilmek için çeşitli programlama modelleri önerilmiştir. Google'ın MapReduce paradigması [46], çok sayıda düşük seviye bilgi işlem düğümleri ile paralel olarak büyük miktarda veri işlemeye olanak tanıması sebebiyle öne çıkan paralel programlama modeli haline gelmiştir.

MapReduce modelinin ana fikri, paralel yürütme ayrıntılarını gizlemek ve kullanıcıların yalnızca veri işleme stratejilerine odaklanmalarına izin vermektir. MapReduce modelinin tasarımı, *"böl ve fethet"* ilkesi ve *"veri taşımak yerine işlemi taşımak"* stratejisine dayanmaktadır [47]. Bu durum MapReduce uygulamalarına; basit, kullanımı kolay, esnek, depolamadan bağımsız ve hata toleranslı olma avantajlarını sağlamaktadır [48]. Kullanıcı programı, MapReduce işlevini çağırdığı zaman Şekil 2.6'daki numaralı etiketler aşağıdaki listedeki numaralara karşılık gelecek şekilde sırasıyla altı eylem gerçekleşmektedir [46].



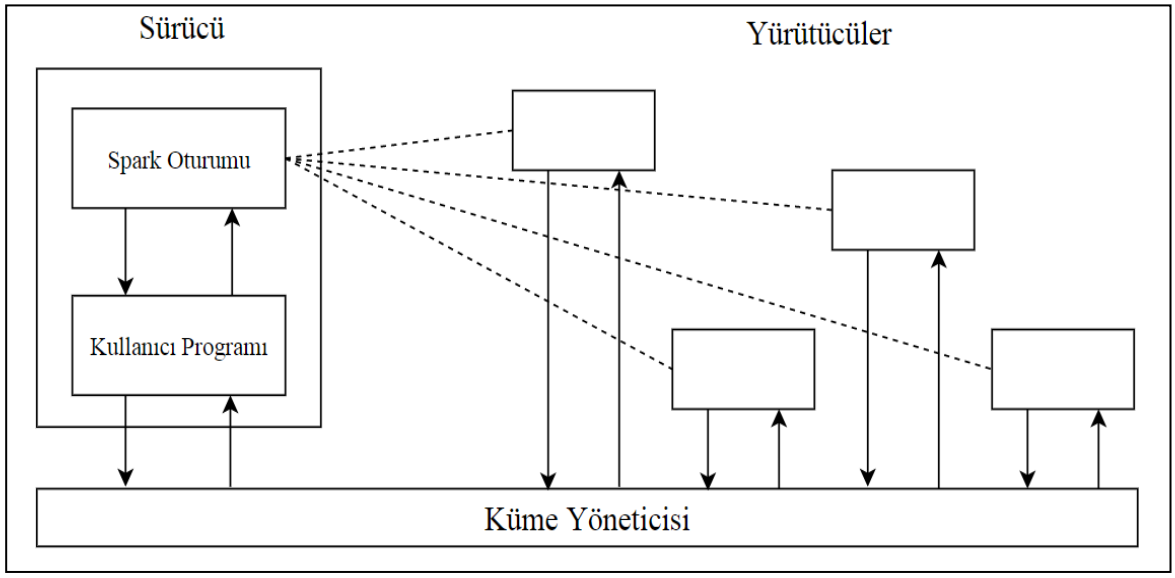
Şekil 2.6. MapReduce paradigması [46]

1. Kullanıcı programındaki MapReduce kütüphanesi ilk olarak girdi dosyalarını parça başına genellikle 16-64 MB olan parçalara ayırır. Daha sonra programın birçok kopyasını bir makine kümesinde başlatır.
2. Programın kopyalarından biri olan ana düğüm özeldir. Geri kalanlar ise ana düğüm tarafından görev atanan işçi düğümlerdir. Ana düğüm boştaki işçi düğümleri seçer ve her birine bir Map görevi veya bir Reduce görevi atar.
3. Map görevi atanan bir işçi düğüm, karşılık gelen girdi bölümünün içeriğini okur ve MapReduce işleminin yürütmeyi kabul ettiği kayıt varlığı olan anahtar/değer çiftlerini giriş verilerinden ayrıştırır. Böylece her anahtar/değer çiftini kullanıcı tanımlı Map işlevine geçirir.
4. Map tarafından üretilen ara anahtar/değer çiftleri periyodik olarak ara belleğe alınır. Bu çiftlerin yerel diskteki konumları, bu konumları Reduce işçilerine iletmekten sorumlu olan ana düğüme geri gönderilir.
5. Bir Reduce işçisine ana düğüm tarafından konumlar hakkında bilgi verildiğinde, ara belleğe alınan verileri Map işçilerinin yerel disklerinden okumak için uzaktan yordam çağrılarını kullanır. Bir Reduce işçisi bölümü için tüm ara verileri okuduğunda, aynı anahtarın tüm oluşumlarının birlikte gruplandırılması için ara anahtarlarla sırama yapar. Sıralama genellikle gereklidir çünkü birçok farklı anahtar aynı Reduce göreviyle eşleşir. Ara veri miktarı belleğe sığmayacak kadar büyükse, harici bir sıralama kullanılır.

6. Reduce işçisi, sıralanan ara veriler üzerinden yinelenir ve karşılaşılan her bir benzersiz ara anahtar için, anahtarı ve karşılık gelen ara değerler kümesini kullanıcının Reduce işlevine geçirir. Reduce işlevinin çıktısı, bu bölüm için bir son çıktı dosyasına eklenir.

MapReduce işlevinin en büyük dezavantajı yinelemeli algoritmaları çalıştırmadaki verimsizliğidir [49]. Disk sınırlandırmalarının üstesinden gelmek ve sistemin performansını artırmak amacıyla bellek içi hesaplamalara olanak tanıyan Spark, MapReduce benzeri RDD (Resilient Distributed Datasets) adı verilen esnek bir programlama soyutlamasıyla paralel işlemlerin yürütülmesine daha verimli bir şekilde olanak sağlar [50].

Şekil 2.7'de, Spark uygulamalarını yürüten küme yöneticisinin çalışma mekanizması görselleştirilmiştir [51]. Spark uygulamaları, bir sürücü işlemi ve bir dizi yürütücü işleminden oluşur. Kümedeki bir düğümde bulunan sürücü, uygulamaya ait tüm bilgilerin saklanması, kullanıcı programına yanıt verilmesi ve işin yürütücüler arasında dağıtılmasından sorumludur. Yürütücüler ise sürücünün kendilerine atadığı işi yapmak ve hesaplamanın durumunu sürücüye iletmek ile görevlidirler.



Şekil 2.7. Spark uygulaması mimarisi [51]

Spark sistemi, Spark çekirdeği ve üst düzey kütüphaneler dahil olmak üzere çeşitli ana bileşenlerden oluşur: makine öğrenimi için MLlib, çizge analizi için GraphX, akan veri işleme için Spark Streaming ve yapılandırılmış veri işleme için Spark SQL [52]. Spark,

MLlib ile veri paralelliğinin veya model paralelliğinin önemli olduğu büyük ölçekli makine öğrenme algoritmalarının geliştirilmesini kolaylaştırmaktadır [53].

Büyük veri analitiği araçları, güçlü sistemler ve dağıtılmış uygulamalarla çoklu bağlantılı sistemleri desteklemektedir. Çizelge 2.1'de çeşitli örnekleri sunulan bu araçlar; veri edinme, depolama, işleme, sorgulama ve yönetme gibi amaçları kapsayan oldukça geniş bir ekosistem içerisinde yer almaktadırlar [54]. Büyük veri ekosistemi, büyük veriyi işlemek ve analiz etmek için kullanılan karmaşık ve birbiriyle ilişkili teknolojiler kümesidir. Ekosistemdeki her bir teknoloji, 3V gösterimindeki bir ya da birkaç büyük veri bileşenine hitap etmektedir.

Çizelge 2.1. Büyük veri analitiğinde kullanılan bazı araçlar

Amaçlar	Araçlar
Veri Edinme	Flume, Sqoop, Chukwa
Veri Depolama	HDFS, HBase, S3
Veri İşleme	MapReduce, YARN, Spark
Veri Görselleştirme	GraphX, Zeppelin, Matplotlib
Veri Yönetimi	Atlas, Falcon, Sentry
Sorgu ve Analiz	Pig Latin, Hive, Mahout, JAQL
Koordinasyon ve İş Akışı	ZooKeeper, Avro, Oozie
Sistem Dağıtımı	Ambari, Whirr, BigTop, Hue

Büyük veri ekosisteminin omurgasını oluşturan ve daha yoğun kullanılan, veri işleme ve depolama amacıyla geliştirilen araçlar; toplu işlem araçları ve akış işleme araçları olarak gruplandırılmaktadır [55]. Toplu işlem araçları için; Hadoop, MapReduce, Pig, Flume, akış işleme araçları için; Storm, S4, Kafka, Flink, hem toplu hem akış işleme aracı için Spark örnek olarak verilebilir.

Tipik bir büyük veri analitiği boru hattı MapReduce benzeri veri yükleme koduna, SQL benzeri sorgulara ve makine öğrenmesi algoritmalarına ihtiyaç duymaktadır [56]. Fakat, makine öğrenmesi algoritmalarının mevcut uygulamaları büyük veri analitiğinde ölçeklenebilirlik, maliyet ve başarım ile ilgili çeşitli zorluklara yol açmaktadır. Bu sebeple, ya büyük veriye özgü yeni yaklaşımların geliştirilmesi ya da mevcut yaklaşımların büyük veriye uyarlanması gerekmektedir [57].

2.3. Yapay Zekâ ve Açıklanabilir Yapay Zekâ

Yapay zekâ (Artificial Intelligence, AI), makinelerin insan düşünme sürecini ve davranışını analiz ve taklit etmesine yardımcı olan bir teori veya yöntemdir [58]. Yapay zekâ, herhangi bir sorunu anında çözebilecek olağanüstü zeki bir makine inşa etmek anlamına gelmez, aksine insan benzeri davranışlar sergileyebilen bir makine tasarlamayı ifade eder. Bu davranış; program tarafından kullanılan veriler, mevcut davranış ile ideal davranış arasındaki hatayı veya mesafeyi ölçen bir metrik ve sonraki olaylarda daha iyi davranış üretmesi için programı yönlendirmek için ölçülen hatayı kullanan bir geri bildirim mekanizması olmak üzere üç faktör doğrultusunda makineye öğretilir [59]. Bu faktörler, yapay zekâ kavramını soyutlayarak akıllı sistemler oluşturmadaki matematiksel derinliğini vurgulamaktadır.

Zaman içerisinde yapay zekâ; 1956-1980 yılları arasında ilk aşama, 1980-2000 yılları arasında sanayileşme aşaması ve 2000 yılı ile sonrasını kapsayan patlama aşaması olmak üzere üç gelişme safhasından geçmiştir [60]. İlk aşamada, yapay zekâ sadece geometrik teoremleri ispatlamak ve cebirsel uygulama problemlerini çözmek gibi amaçlar için kullanılıyordu. Sanayileşme aşamasında, yapay zekâ araştırmalarının odağı bilgi işleme olmuştur ve insan-makine diyaloğunu destekleyen makineler üretilmiştir. Son olarak patlama aşaması ise, araştırmaların artmasıyla pek çok yapay zekâ teknolojisinin sıradan insanların gerçek yaşamında bile yer almasına sebep olmuştur. Bu süreçte, özellikle büyük veri yapay zekâyı güçlendirerek mevcut patlamada önemli rol oynamıştır.

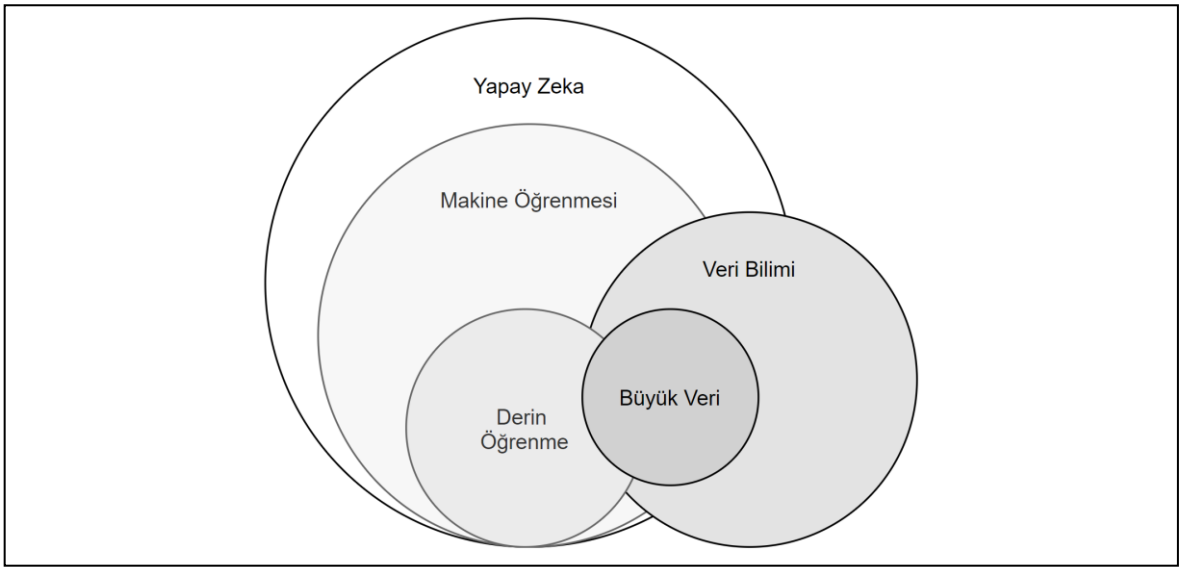
Büyük veriye erişim sağlayarak önceden tanımlanmış bir davranışı üretmeyi öğrenebilen modelleri yaratmak için aşağıda belirtildiği gibi; tanımlayıcı, kestirimci ve kuralcı olmak üzere üç farklı analitik metot kullanılmaktadır [61].

1. Tanımlayıcı Analitik: Örüntüleri tanımlamak ve geçmiş davranışı modellemekle ilgili raporları oluşturmak için büyük miktardaki veri ve bilginin özetlenmesi veya incelenmesidir. Genellikle istatistiksel veri madenciliği araçları kullanılarak yararlı kalıplar veya tanımlanmamış ilişkiler keşfedilir.
2. Kestirimci Analitik: Geçmiş verilere dayanarak, gelecekteki olasılıkları belirlemek için makine öğrenmesi tekniklerinin kullanıldığı tahmin modellerinin üretilmesidir. Amaç,

ne olduğunu bilmenin ötesine geçerek, ne olacağına dair en iyi değerlendirmeyi sağlamaktır.

3. Kuralcı Analitik: Tanımlayıcı ve kestirimci analitikler sonucunda eylemlerin belirlenerek, bu eylemlerin iş hedefleri, gereksinimleri ve kısıtlamaları üzerindeki etkileri doğrultusunda iş analistlerine eniyileme ve karar verme sürecinde yardımcı olmayı amaçlamaktadır.

Yapay zekânın bir alt alanı olarak ortaya çıkan makine öğrenmesi (Machine Learning, ML), kestirimci analitiklerin modern bir uzantısı olarak kabul edilmektedir ve istatistik ile bilgisayar biliminin kesişimi sonucunda karmaşık kalıpları otomatik bir şekilde tanımayı ve mevcut veri kümelerine dayalı akıllı kararlar almayı amaçlamaktadır [62]. Bu süreçte, istatistik veri kümelerinin modellenmesini ve özetlenmesini sağlarken, bilgisayar bilimi veri kümelerini verimli bir şekilde depolamak, işlemek ve görselleştirmek için algoritmalar tasarlayıp kullanabilme becerilerini kazandırmaktadır.

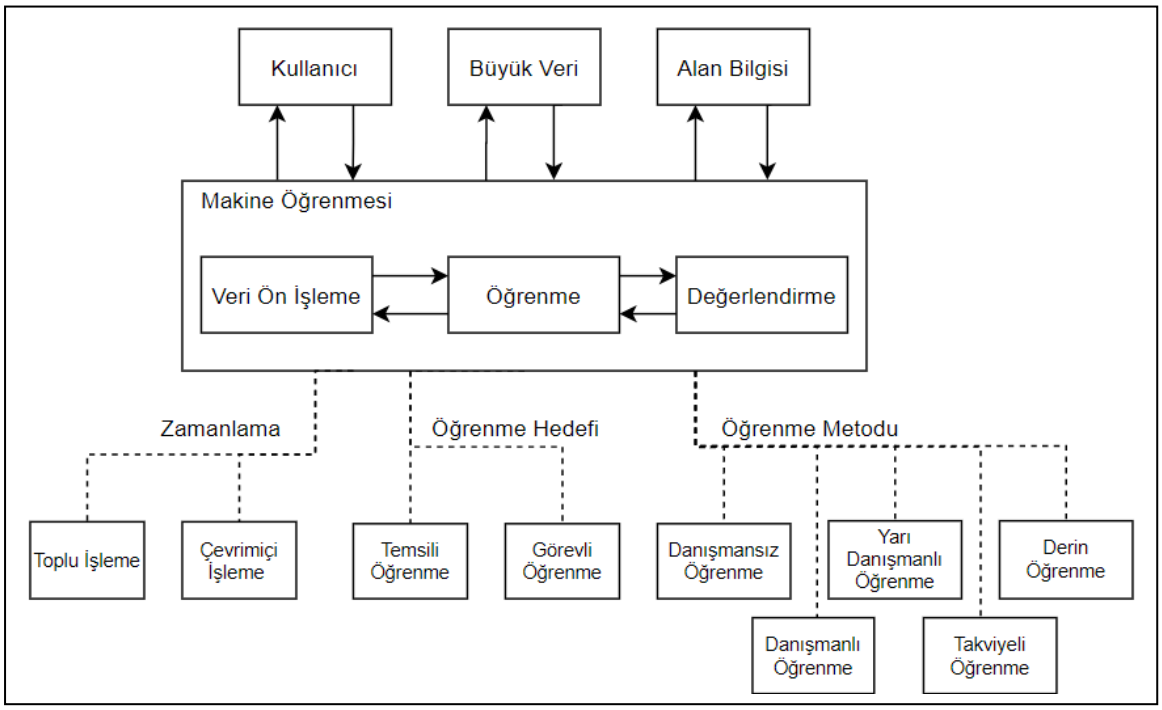


Şekil 2.8. Yapay zekâ ile ilişkili kavramların kesişimi [63]

Yapay zekâ ile ilişkili terminoloji genellikle kavram karmaşıklığı barındırır çünkü sık kullanılan birkaç terim birbiriyle örtüşür [63]. Şekil 2.8, bu anahtar terimlerin nasıl kesiştiğini görselleştirmenin bir yolunu sunmaktadır. Büyük oranda makine öğrenmesi üzerine kurgulanan büyük veri analitiği, Şekil 2.9'da verildiği üzere bütünleyici bir çerçeve ile ele alınabilir. Bu çerçeve; makine öğrenmesinin kaynakları, makine öğrenmesi süreci,

öğrenme zamanlaması, öğrenme hedefi ve öğrenme metodu olmak üzere beş kanal bağlamında oluşturulmuştur.

Büyük veri analitiği için makine öğrenmesi çerçevesi; büyük verinin kendisi, kullanıcılar ve alan bilgisi ile beslenmektedir [64]. Büyük veri kümelerinin tek başına gürbüz modellerin bulunmasını sağlayacak kadar temsil edici olmadığı durumlarda, alan bilgisi uygulamaya özgü gereksinimlerin karşılanmasıyla örüntülerin genelliğini ve sağlamlığını geliştirmeye yardımcı olur. Kullanıcının katılım amacı ise, veri etiketleri sağlamak, alanla ilgili soruları yanıtlamak veya genellikle uygulayıcıların aracılık ettiği uzun ve asenkron yinelemelere yol açan öğrenilmiş sonuçlar hakkında geri bildirim sağlamaktır.



Şekil 2.9. Büyük veri üzerinde makine öğrenmesi çerçevesi

Makine öğrenmesi, veri bilimi sürecinde olduğu gibi tipik olarak sırasıyla veri önişleme, öğrenme ve değerlendirme aşamalarından oluşmaktadır [65]. Büyük veri çerçevesi dâhilinde ise makine öğrenmesi algoritmalarının, dağıtılmış paralel programlama paradigmalarına uyması gerekir. Veri ön işleme; muhtemelen yapılandırılmamış, gürültülü, eksik ve tutarsız ham veriyi temizleme, çıkarma, dönüştürme ve öğrenmeye girdi olarak kullanılabilecek bir forma dönüştürme sürecidir. Öğrenme aşamasında, öğrenme metodu seçilir ve önceden işlenmiş girdi verisi kullanılarak istenen çıktıları üretmek için model parametreleri ayarlanır.

Değerlendirme ise öğrenilen modellerin performans ölçümü ve hata tahmini için yapılır. Değerlendirme sonuçları, seçilen öğrenme algoritmalarının parametrelerinin yeniden ayarlanmasına veya farklı algoritmaların seçilmesine yol açabilir.

Öğrenme verilerini kullanılabilir hale getirme zamanlamasına dayanarak makine öğrenmesi; toplu öğrenme ve çevrimiçi öğrenme olarak sınıflandırılır [66]. Toplu öğrenme, tüm eğitim verisini bir anda alıp öğrenerek modeller üretirken, çevrimiçi öğrenme modelleri her yeni girdiye göre güncellenir.

Öğrenme hedefine göre makine öğrenmesi; temsili öğrenme ve görevli öğrenme olarak kategorize edilebilmektedir [67]. Temsili öğrenme, yararlı bilgilerin çıkarılmasını kolaylaştıran yeni veri temsillerini öğrenmeyi amaçlarken, görevli öğrenme tipik olarak istenen çıktılara sahiptir ve buna göre sınıflandırma, regresyon ve kümeleme olarak kategorize edilir.

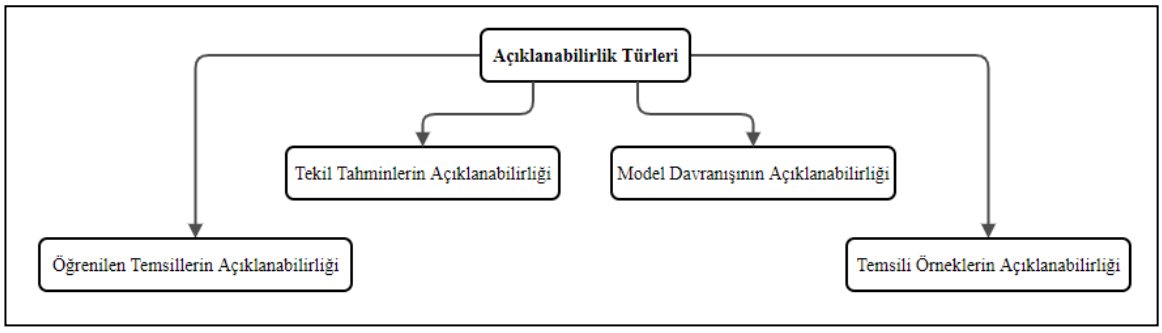
Öğrenme metotlarına göre makine öğrenmesi algoritmaları; danışmansız öğrenme, danışmanlı öğrenme, yarı danışmanlı öğrenme, takviyeli öğrenme ve derin öğrenme olmak üzere, aşağıda verildiği gibi beş sınıfa ayrılmaktadır [68].

1. Danışmansız Öğrenme: Etiketlenmemiş verideki gizli yapıların bulunması sürecidir. Kümeleme, danışmansız öğrenmenin önemli formlarından biridir. Kümeleme, veri noktaları arasındaki küme içi benzerlik maksimum ve kümeler arası benzerlik minimum olacak şekilde veri kümesini gruplara ayırır.
2. Danışmanlı Öğrenme: Doğrulama için etiketli verilerin kullanıldığı analiz yöntemidir. D öğrenmedeki iki geniş alt kategori sınıflandırma ve regresyondur. Kayıtların sınıfını belirlemek veya gelecekteki eğilimleri tahmin etmek için, sınıflandırma kategorik etiket kullanırken, regresyon sürekli etiket kullanır.
3. Yarı Danışmanlı Öğrenme: Veri kümelerinin az miktarda etiketli ve çok miktarda etiketlenmemiş veri içerdiği durumlarda kullanılır. Kayıtlarının çoğunun etiketsiz olduğu büyük veri problemleri için çok uygundur. Eş eğitim ve aktif öğrenme, çıktı elde etmek için kaynağı etkileşimli olarak sorgulayan iki önemli tekniktir.
4. Takviyeli Öğrenme: Daha iyi kararlar almak için deneme yanılma yoluyla en büyük ödülleri veren eylemleri keşfeder. Genellikle oyun ve robotik gibi alanlarda kullanılır. Markov Karar Süreci ve Q-Learning takviyeli öğrenmenin iyi bilinen iki örneğidir.

5. Derin Öğrenme: Derin mimarideki hiyerarşik temsilleri otomatik olarak öğrenmek için danışmanlı, danışmansız ya da takviyeli öğrenme stratejileri kullanan makine öğrenme tekniklerini ifade eder ve genellikle el yapımı özelliklerle oluşturulan modellerden daha iyi performans gösterir. Derin sinir ağları, derin inanç ağları, tekrarlayan sinir ağları ve evrişimli sinir ağları gibi pek çok derin öğrenme mimarisi bulunmaktadır.

Makine öğrenmesi metodu seçimi, araştırma sorusuna ve veri toplama stratejisine bağlı olarak değişkenlik göstermektedir. Veri kaynağının deneysel veya gözlemsel olması, veri kümesinin temsil ettiği topluluk ve varılmak istenen sonucun türü gibi sebepler bu seçime etki etmektedir. Bununla birlikte uygulanan makine öğrenmesi metodu sonunda elde edilen model ve modelin ürettiği sonuçlar, çoğunlukla kolayca yorumlanabilir değildirler. Bu durum da, açıklanabilir yöntemlere ihtiyaç duyulmasına sebep olmuştur.

Yapay zekâ yöntemleri giderek karmaşıklaşan görevleri çözmeyi öğrenerek etkileyici performans seviyelerine ulaşsalar bile, genellikle girdi verilerinin karara ulaşma üzerindeki etkisinin net olmadığı şeffaf olmayan bir yapıya sahiptirler. İlk yapa zekâ sistemleri kolayca yorumlanabilirken, son yıllarda yükselişe geçen derin sinir ağları genellikle kara kutular olarak kabul edilir. Kullanıcılar sistemi, mevcut kararlarla alakalı olduğuna inanılan değişkenlerle beslese bile ortaya çıkan model yalnızca verideki örüntüleri yansıtır ve bu örüntülerin altında yatan nedensel ilişkilerin dayanağı net olarak bilinmemektedir. Son yıllarda, kara kutu modellerini açıklamaya yönelik teknikler, modelin ne öğrendiğini daha iyi anlamaya yardımcı olmaya başlamıştır.



Şekil 2.10. Yapay zekâ modelleri için açıklama türleri

Alıcısına ve amacına bağlı olarak farklı açıklama türleri bulunmaktadır. Genel model davranışının değerlendirilmesinden, öğrenilen temsiller hakkındaki bilgilerden farklı tahmin

stratejilerinin tanımlanmasına kadar modelin farklı yönleri hakkında bilgi sunan açıklama türleri Şekil 2.10'da verildiği üzere dört başlık altında incelenmektedir [69]. Öğrenilen temsilleri açıklama, bir derin sinir ağının nöronlarının anlaşılması gibi temsillerin rolünü araştırmayı ya da öğrenilen kavramı temsil eden prototipler oluşturarak modelin öğrendiklerini yorumlamayı amaçlamaktadır. Tekil tahminleri açıklama, modelin karara varması için hangi piksellerin en alakalı olduğunu görselleştiren veya bir girdinin en hassas kısımlarını vurgulayan ısı haritaları gibi yöntemlerle tahminlerin doğrulanmasına yardımcı olan tekniklerdir. Model davranışını açıklama, farklı tahmin stratejilerinin belirlenmesi gibi model davranışı hakkında daha genel bir anlayış sunarak tekil tahminlerin ötesine geçer. Tekil ısı haritalarının kümelenmesiyle hesaplanabilirler. Temsili örnekler ile açıklama ise, eğitim kümesini ve modeli nasıl etkilediğini daha iyi anlamak için temsili eğitim örneklerini tanımlayarak sınıflandırıcıları yorumlar. Bu temsili örnekler potansiyel olarak verideki önyargıların tanımlanmasına ve modelin eğitim kümesinin varyasyonları için daha gürbüz olmasına yardımcı olabilir.

Açıklama, yapay zekâ sistemlerinin evrimi ile yakından bağlantılıdır. Erken yapay zekâ açıklama çabaları sistem geliştiricilerinin yanlış muhakeme yollarını teşhis etmesine yardımcı olmayı amaçlayan yorumlanabilirlik üzerine kurgulanmışken, modern şeffaf açıklama yöntemleri ise muhtemel ilişkileri ortaya çıkarmayı ve görselleştirmeyi hedeflemektedir [70]. Benzer bir deyişle; yorum, çıktı sınıfı gibi soyut bir kavramın bir etki alanı örneğine eşlenmesi olarak tanımlanırken, açıklama modelin çıktı kararına katkıda bulunan bir görüntünün pikselleri gibi bir etki alanı özellikleri kümesi olarak tanımlanmaktadır [71].

Yorumlanabilirlik; lojistik regresyon, saklı markov modelleri ve karar ağaçları gibi daha yapılandırılmış istatistiksel ve çizgesel modellerin çıktılarının bu tekniklerde uzman kişiler tarafından analiz edilmesini kapsamaktadır [72]. Bir makine öğrenmesi modeli geliştirirken, ek bir tasarım sürücüsü olarak yorumlanabilirliğin dikkate alınması, aşağıda verilen üç nedenden dolayı uygulanabilirliğini artırmaktadır [73].

- Karar vermede tarafsızlığın sağlanmasına yani eğitim kümesindeki önyargının tespit edilmesine ve düzeltilmesine yardımcı olur.
- Öngörüyü değiştirebilecek potansiyel düzensizlikleri vurgulayarak gürbüzlük sağlanmasını kolaylaştırır.

- Yalnızca anlamlı değişkenlerin çıktıya etki ettiğini yani model akıl yürütmesinde altta yatan gerçek nedenselliği ortaya çıkarır.

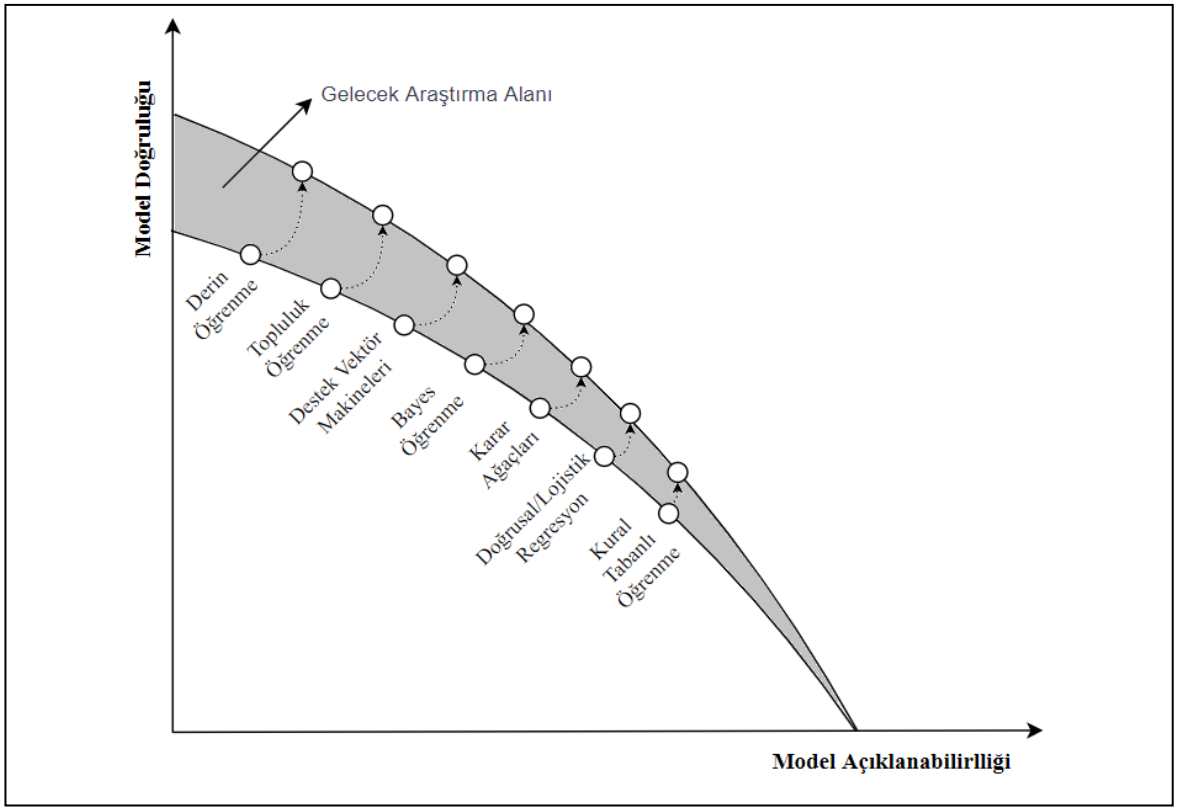
Yapay sinir ağları ve takviyeli öğrenme gibi performansı oldukça yüksek olmasına rağmen yorumlanabilirliği düşük tekniklerin gelişmesiyle, yorumlanabilirlik yerini zamanla *"açıklanabilir yapay zekâ"* (Explainable Artificial Intelligence, XAI) kavramına bırakmaya başlamıştır. Açıklanabilir yapay zekâ; yapay zekâ kararlarının yüksek kalitede yorumlanabilir, sezgi ile kavranabilir ve insan tarafından anlaşılabilir açıklamalarını oluşturabilen bir dizi araç, teknik ve algoritmayı destekleyen yapay zekâ alanıdır [74]. Yapay zekâ sistemlerini açıklama ihtiyacı aşağıda verilen dört nedenden kaynaklanmaktadır [75].

- Sistemin önyargılı veya ayrımcı sonuçlar verme ihtimaline dayanan endişeleri gidererek, verilen kararları gerekçelendirir.
- Modelin kullanıcı ve makine arasındaki tekrarlanan etkileşimleri sayesinde iyileştirilmesini sağlar.
- Yeni gerçekleri öğrenmek, bilgi toplamak ve öz bilgi kazanmak için yararlı olan açıklanabilirlik bu süreçte yeni keşiflerin yapılmasına zemin hazırlar.
- Bilinmeyen açıklar ve kusurlar üzerinde daha fazla görünürlük sağlayarak, hataları hızlı bir şekilde belirlemeye ve düzeltmeye yardımcı olur.

Makine öğrenmesi modellerinin performansı ile açıklanabilirlik arasında doğal bir denge olduğu düşünülmektedir. Şekil 2.11'de verildiği üzere, performans karmaşıklıkla bir araya geldiğinde, açıklanabilirlik mevcut durum için kaçınılmaz görünen aşağı doğru bir eğimle karşılaşır [76]. Genellikle en yüksek performans gösteren yapay sinir ağları gibi modeller en az açıklanabilirken, kural tabanlı yöntemler gibi en açıklanabilir modeller ise en az doğrudur. Açıklanabilir yapay zekâ tekniklerinin ve araçlarının potansiyelinin bulunduğu gelecek araştırma alanı ise beklenen iyileştirmeleri kapsamaktadır. Bununla birlikte, açıklanabilirlik için daha gelişmiş yöntemlerin ortaya çıkması bu eğimi değiştirebilir.

Model açıklanabilirliği ve performansıyla yakın bağlantısı nedeniyle bu noktada değinilmesi gereken bir diğer husus da *"yaklaşım ikilemi"*dir. Yaklaşım ikilemine göre, bir makine öğrenmesi modeli için yapılan açıklamalar tesirli ve gereksinimleri karşılayacak kadar yaklaşık yapılmalıdır, böylelikle açıklamaların incelenen modeli temsil etmesi sağlanırken

temel özellikleri fazla basitleştirmemesi sağlanır [77]. Şekil 2.12'de verildiği üzere XAI çerçevesi; bir göreve ait giriş verisini alarak onu hem öngörü, tavsiye ya da eyleme dönüştürür hem de kullanıcıya bunları gerekçelendiren açıklamalarını sunar ve kullanıcı bu açıklamalara göre karar verir [78].

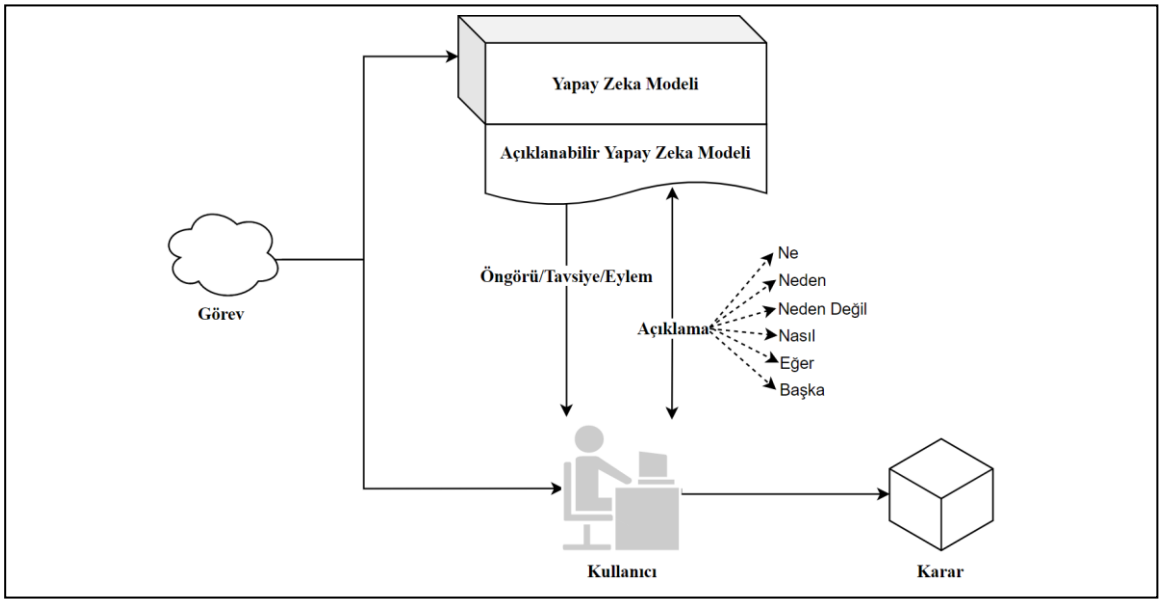


Şekil 2.11. Yaklaşımların modelin doğrulukları ve açıklanabilirlik dengesi [76]

Kullanıcılar karmaşık bir akıllı sistemle karşılaştıklarında, farklı türde açıklayıcı bilgiler talep edebilirler ve her açıklama türü kendi tasarımını gerektirebilir. Açıklanabilir yapay zekâ sistem tasarımlarında kullanılan altı yaygın açıklama türü aşağıda özetlenmiştir [79].

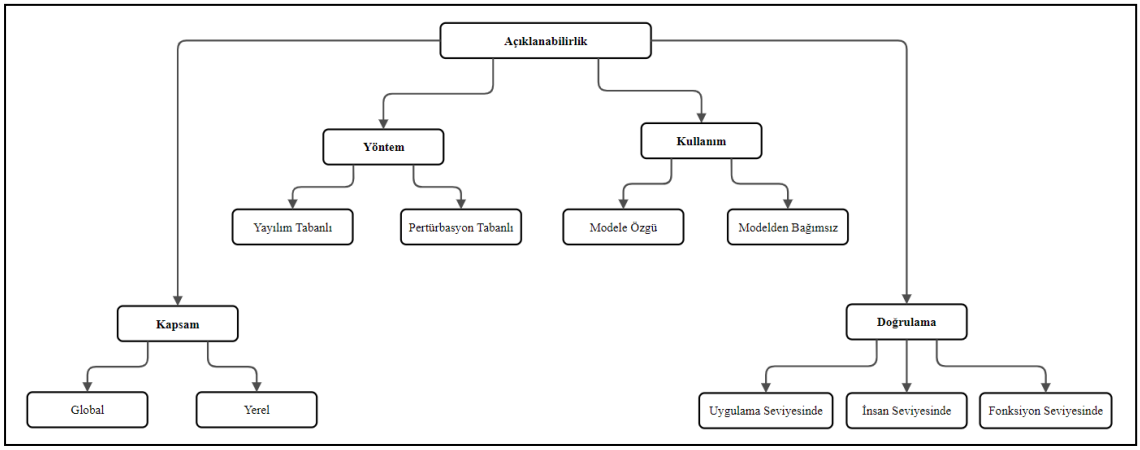
- Ne: Modelin ne öğrendiğini açıklamak için makine öğrenmesi algoritmasının bütünsel bir temsilini gösterir. Model çizgeleri ve karar sınırları, bu açıklama türüne yönelik yaygın tasarım örnekleridir.
- Neden: Belirli bir girdi yapılan tahminin nedenini açıklar. Bu tür açıklamalar, girdi verilerindeki hangi özelliklerin veya modeldeki hangi mantığın algoritma tarafından belirli bir tahmine yol açtığını iletmeyi amaçlamaktadır. Özellik önemi, bu açıklama türü için bir yorumlanabilirlik tekniği olarak kullanılır.

- Neden Değil: Kullanıcıların sistemin çıktısında neden belirli bir çıktının bulunmadığını anlamalarına yardımcı olur. Model tahmini ile kullanıcının beklenen sonucu arasındaki farklılıkların nedenlerini tanımlar.
- Nasıl: Girdiye veya modele yönelik varsayımsal ayarlamaları belirtir. Model yapısı ve iç değerleri açıklamanın bir parçası olmadığı için bu açıklama türleri geçici ve modelden bağımsızdırlar.
- Eğer: Farklı algoritma veya veri değişiklilerinin modelin parametrelerini veya çıktısını nasıl etkilediğini ortaya koyar. Sistem tarafından otomatik olarak önerilen farklı durum senaryoları ya da etkileşimli kullanıcı kontrolü aracılığıyla gerçekleştirilebilirler.
- Başka: Bu açıklamalar, kullanıcılara modelden aynı veya benzer çıktılar üreten benzer girdi örnekleri sunar. Eğitim kümesi düzenli dağılımından yoksun olduğunda, örnek tabanlı bu açıklamalar yanıltıcı olabilmektedir.



Şekil 2.12. Açıklanabilir yapay zekâ çerçevesi [78]

Açıklanabilir yapay zekâ yaklaşımlarının temel amacı, tasarımcılarına önemli olan açıkça tanımlanmış bir dizi unsuru ortaya çıkarmak ve bu unsurları dikkate almalarında yardımcı olmaktır. Bu bağlamla açıklanabilir yapay zekâ yaklaşımları, Şekil 2.13'de verildiği üzere; kapsamına, kullanımına, yöntemine ve doğrulama stratejisine göre çeşitlilik göstermektedir [80], [81].



Şekil 2.13. Açıklanabilir yapay zekâ taksonomisi

Açıklamaların kapsamı yerel veya global olabilir. Yerel olarak açıklanabilir yöntemler, veri kümesindeki tek bir girdi verisi örneğinin bireysel niteliklerini ifade etmek için tasarlanmıştır. Global olarak açıklanabilir modeller, bir bütün olarak modelin kararına ilişkin içgörü sağlar ve bir dizi girdi verisi için nitelikler hakkında bir anlayış ortaya koyar.

Açıklanabilir modelin arkasındaki temel algoritma, uygulama yöntemine dayalı olarak geri yayılım tabanlı veya pertürbasyon tabanlı olarak kategorize edilebilir. Yayılım tabanlı yöntemlerde açıklanabilir algoritma, sinir ağından bir veya daha fazla ileri geçiş yapar ve aktivasyonların kısmi türevlerini kullanarak yayılım aşamasında haritalar üretir. Pertürbasyona dayalı açıklanabilir algoritmalar, belirli bir girdi örneğinin özellik kümesini, özellikleri kısmen ikame etme, maskeleme ya da koşullu örnekleme gibi teknikler kullanarak bozmaya odaklanır.

Belirli bir kapsam ve yönteme sahip iyi geliştirilmiş açıklanabilir bir yöntem, ya sinir ağı modelinin kendisine gömülebilir ya da açıklama için harici bir algoritma olarak uygulanabilir. Model mimarisine bağlı olan açıklanabilir herhangi bir algoritma, modele özgü kategorisine girmektedir. Bu bağlamda değerlendirilen regresyon, kural tabanlı öğrenici veya karar ağaçları gibi makine öğrenmesi yaklaşımları; doğrudan modellerden bilgi çıkarmak, sadeleştirme teknikleri kullanmak, görselleştirmek ya da özellik önem tahminleri ile açıklanabilirlik açısından ele alınmaktadır. Modelden bağımsız yöntemler ise, dâhili model ağırlıklarına veya yapısal parametrelere doğrudan erişimi olmayan geçici açıklanabilir yöntemlerdir. Bu başlıkta ele alınan destek vektör makineleri ya da yapay sinir ağları gibi makine öğrenmesi yaklaşımları ise model özgü açıklanabilirlik tekniklerine ek

olarak mimari modifikasyon, örnekle açıklama ya da şeffaf modeller kullanılarak gerçekleştirilmektedir.

Son olarak, açıklanabilir sistemler, amaçlanan geliştirme senaryosuna benzer şartlar altında örnek çalışmalar veya sentetik deneyler ile doğrulanmalıdır. Uygulama seviyesinde yapılan doğrulamalar, alan uzmanlarının görüşleri ile açıklanabilir yapay zekâ yaklaşımının sonuçlarını karşılaştırmak gibi ile gerçek bir görevde gerçekleştirilir. İnsan seviyesinde doğrulama, zamanı kısıtlı ve pahalı olan alan uzmanlarını kullanmak yerine görevi oldukça basitleştirerek sıradan bir kitlenin doğrulama yapmasını amaçlar. Fonksiyon seviyesinde doğrulama ise, zaten kanıtlanmış bir açıklanabilir yapay zekâ yöntemini vekil yaklaşımlarla doğrulamayı ifade eder.

Özellikle üretim, sağlık, ulaşım, tarım askeri ve ekonomi gibi kritik altyapıları barındıran alanlarda geniş yer tutmaya başlayan yapay zekâ modellerinin hem doğru hem de anlaşılabilir olması gerekmektedir. Yapay zekâ modelleri tahmin yaptığında ne olacağını anlamak; bu sistemlerin yaygın olarak benimsenmesini hızlandırır, kullanıcıların sisteme olan güvenini artırır, açıklanabilir yapay zekânın bir zorunluluk ya da yasal kısıtlama olduğu sektörlerde gelişmiş modellerin entegrasyonunu sağlar ve işletme, araştırma veya karar verme süreçlerinde daha fazla iyileştirme fırsatına zemin hazırlar [82]. Son yıllarda açıklanabilir yapay zekâ alanında önemli ilerlemeler kaydedilmiş olsa da; yöntemler, teori ve açıklamaların pratikte nasıl kullanıldığına ilişkin zorluklar hala mevcuttur ve gelişim süreci devam etmektedir.

3. SINIF DENGESİZ VERİDE DOLANDIRICILIK TESPİTİ YAKLAŞIMLARI

Dolandırıcılık, bir dizi alanda meydana gelen oldukça maliyetli bir suç faaliyetidir. Dolandırıcılık tespiti danışmanlı yaklaşımlarla çözülmek istendiğinde, karşılaşılan en büyük sorunlardan biri ise, sınıf dengesiz verinin varlığıdır.

Bu bölümde; sınıf dengesiz verinin analizi, dolandırıcılık tespiti ve sınıf dengesiz dolandırıcılık tespiti için büyük veri analitiği konularında literatürde önerilen yaklaşımlar kategorize edilerek özetlenmiştir. Sınıf dengesiz veri çözümleri; veri düzeyinde yöntemler, algoritma düzeyinde yöntemler, hibrit yöntemler ve derin öğrenme yöntemleri olarak dört kategoride, dolandırıcılık tespiti çalışmaları ise; kavram sapması engelleme yöntemleri, sınıf dengesizliği giderme yöntemleri, veri azaltma yöntemleri ve gerçek zamanlı tespit yöntemleri olarak dört kategoride incelenmiştir. Son olarak, bu tezin kapsamı olan sınıf dengesiz dolandırıcılık tespiti için büyük veri analitiği gerçekleştiren çalışmalar verilmiştir.

3.1. Sınıf Dengesiz Veri Analizi

Sınıf dengesiz veri problemi, genellikle ilgilenilen kavrama (pozitif veya azınlık sınıfı) atıfta bulunan sınıf, veri kümesinde yetersiz temsil edildiğinde ortaya çıkar; diğer bir deyişle, negatif (çoğunluk) örneklerin sayısının, pozitif örneklerin sayısından daha fazla olduğu durumdur [83]. Spesifik olarak bu problem *"sınıflar arası dengesizlik"* olarak adlandırılırken, sınıf içerisindeki alt kavramların temsilinin yetersiz olması ise *"sınıf içi dengesizlik"* (küçük bölenler sorunu) olarak tanımlanmaktadır [2]. Sınıf dengesiz verinin şiddeti yani dengesizlik oranı, IR (Imbalance Ratio) olarak tanımlanan çoğunluk sınıfın azınlık sınıfa oranı ile ifade edilir ve genellikle 100:1 ve 10000:1 aralığında değişiklik göstermektedir [84].

Sınıf dengesizliği problemi, veri bilimi için büyük önem taşıyan çok sayıda alanda yaygın olarak görülmektedir. Bu sorun, veri toplama süreci belirli nedenlerle sınırlı olduğu durumlarda oluşabileceği gibi, veri edinimi sürecindeki teknik hatalardan da kaynaklanabilir. Dolandırıcılık tespiti, hastalık teşhisi, spam filtreleme, müşteri kaybı tahmini, saldırı tespiti, doğal afet tahmini ve aykırılık tespiti gibi pek çok konu sınıf dengesizliği bakış açısıyla değerlendirilmektedir [85].

Çizelge 3.1. Karışıklık matrisi

		Öngörülen Sınıf	
		Pozitif	Negatif
Gerçek Sınıf	Pozitif	TP	FN
	Negatif	FP	TN

Sınıf dengesizliği probleminde, kümeleme veya birliktelik kural madenciliği gibi danışmansız öğrenme teknikleri; tek başına grupların keşfi süreci olarak, sorunun karmaşıklığını azaltma yöntemi olarak veya azınlık sınıfı yapısının analizine bir çözüm olarak kullanılmaktadır [1]. Bu sebeple sınıf dengesizliği, literatürde sıklıkla sınıflandırma problemi olarak ele alınmaktadır. Veri kümesindeki örneklerin dağılımının değiştirilmesiyle ya da sınıflandırma algoritmalarının modifiye edilmesiyle gerçekleştirilen yaklaşımların başarısı, karışıklık matrisi tabanlı metrikler doğrultusunda değerlendirilmektedir [86]. Çizelge 3.1'de pozitif ve negatif olarak ifade edilen iki sınıf için verilen karışıklık matrisi, veri kümesindeki bu sınıflara ait örneklerin sınıflandırma algoritması tarafından aşağıda açıklanan öngörü türlerine atanma sayısını ifade etmektedir.

- Doğru pozitif (True Positive, TP): Pozitif olduğu öngörülen ve gerçekte pozitif olan örnek sayısı
- Yanlış Pozitif (False Positive, FP): Pozitif olduğu öngörülen ve gerçekte negatif olan örnek sayısı
- Yanlış Negatif (False Negative, FN): Negatif olduğu öngörülen ve gerçekte pozitif olan örnek sayısı
- Doğru Negatif (True Negative, TN): Negatif olduğu öngörülen ve gerçekte negatif olan örnek sayısı

Pozitif ve negatif sınıfların sınıflandırma performansını bağımsız olarak ölçebilmek için, karmaşıklık matrisinden; Eş. 3.1'de verilen doğru pozitif oranı (TPR, sensitivity, recall), Eş. 3.2'de verilen doğru negatif oranı (TNR, specificity), Eş. 3.3'de verilen yanlış negatif oranı (FNR, miss rate), Eş. 3.4'de yanlış pozitif oranı (FPR, fall out), Eş. 3.5'de verilen pozitif tahmin değeri (PPV, precision), Eş. 3.6'da verilen negatif tahmin değeri (NPV), Eş. 3.7'de verilen doğruluk (ACC, accuracy), Eş. 3.8'de verilen F_1-skor (F_1-score), Eş. 3.9'da verilen

G-ortalama (G-mean), ve Eş. 3.10'da verilen eğri altındaki alan (AUC, Area Under The Curve) metrikleri elde edilmektedir [87].

$$Doğru\ Pozitif\ Oranı = \frac{TP}{TP+FN} \tag{3.1}$$

$$Doğru\ Negatif\ Oranı = \frac{TN}{TN+FP} \tag{3.2}$$

$$Yanlış\ Negatif\ Oranı = \frac{FN}{FN+TP} \tag{3.3}$$

$$Yanlış\ Pozitif\ Oranı = \frac{FP}{FP+TN} \tag{3.4}$$

$$Pozitif\ Tahmin\ Değeri = \frac{TP}{TP+FP} \tag{3.5}$$

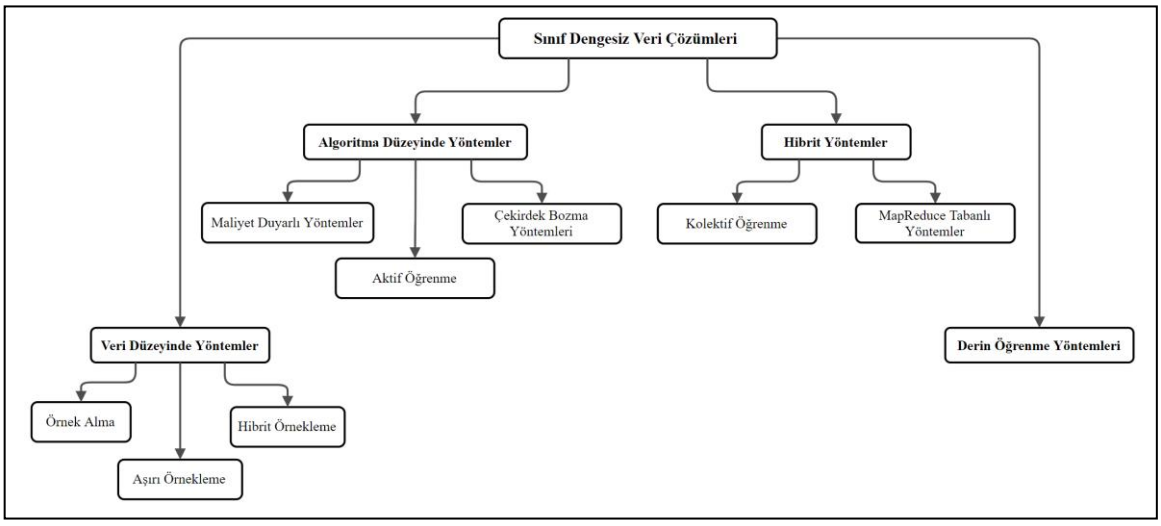
$$Negatif\ Tahmin\ Değeri = \frac{TN}{TN+FN} \tag{3.6}$$

$$Doğruluk = \frac{\mathrm{TP+TN}}{\mathrm{TP+TN+FP+FN}} \tag{3.7}$$

$$F1\text{-}Skor = 2 . \frac{PPV.TPR}{PPV+TPR} \tag{3.8}$$

$$G\text{-}Ortalama = \sqrt{TPR\ .TNR} \tag{3.9}$$

$$Eğri\ Altındaki\ Alan = \frac{TPR+TNR}{2} \tag{3.10}$$



Şekil 3.1. Sınıf dengesiz veri çözümleri

Sınıf dengesiz veri, sınıflandırma algoritmasının eğitim sürecini baskılayarak, daha önemli ve daha maliyetli olan azınlık sınıfını gürültü gibi varsayıp çoğunluk sınıfına doğru bir ön yargı yaratır [3]. Örneğin, %1 negatif etikete sahip bir veri kümesinde %99 doğruluk değeri elde etmek gerçek bir başarı sayılamaz. Sınıflandırma her iki sınıf için de kaliteli sonuçlar elde etmeyi amaçladığından, pek çok sınıflandırma algoritması için tercih edilen doğruluk

metriği, sınıf dengesiz veri için aldatıcı sonuçlara sebep olmaktadır. Bu sebeple, sınıf dengesiz veri çözümlerinin değerlendirilmesi amacıyla sıklıkla doğruluk yerine F_1-skor, G-ortalama ve AUC metrikleri tercih edilmektedir.

Geleneksel sınıflandırma yaklaşımlarında ve başarı değerlendirmesi metriklerinde güncellemeleri gerektiren sınıf dengesiz veri çözümleri, Şekil 3.1'de verildiği üzere; veri düzeyinde yöntemler, algoritma düzeyinde yöntemler, hibrit yöntemler ve derin öğrenme yöntemleri olarak dört kategoride incelenebilir.

3.1.1. Veri düzeyinde yöntemler

Veri düzeyinde yöntemler, çoğunluk ve azınlık sınıfları arasındaki dengesiz dağılımı telafi etmek için eğitim kümesini değiştirmeyi amaçlar. Örnek alma, aşırı örnekleme ve hibrit örnekleme olarak üç grupta incelenebilir. Uygulaması oldukça kolay olmasına rağmen, bu yöntemler çoğunluk sınıfından ayrımcı noktaların kaldırılmasına ya da azınlık sınıfına anlamsız yeni noktaların eklenmesine yol açabilme dezavantajına sahiptir.

Örnek alma yöntemlerinin en basiti ve en popüleri olan RUS (Random Under Sampling), dengesizliğin etkisi önemli ölçüde hafifletilene kadar çoğunluk sınıflarından örnekleri rastgele atar [88]. Yoğunlaştırılmış en yakın komşu olan CoNN (Condensed Nearest Neighbor), karar sınırlarından uzak olan çoğunluk sınıfı örneklerini, öğrenme için daha az alakalı olduğu gerekçesiyle eğitim kümesinden çıkarır [89]. TL (Tomek Links), CoNN'nin tersine çoğunluk sınıfındaki gürültülü ve sınırdaki örnekleri güvensiz olarak değerlendirerek, atar [90]. Yen ve Lee, ön işleme sonrası çoğunluk ve azınlık sınıflarındaki veri dağılımlarını korumak için kümelenmeye dayalı bir örnek alma yöntemi önermiştir [91]. Garcia ve Hererra, sınıflandırma doğruluğu ve indirgeme oranları ile ilgili uygunluk fonksiyonuna dayanan çoğunluk sınıfı için örneklerin verimli bir şekilde seçilmesi için evrimsel algoritmalara dayalı bir dizi örnekleme yöntemi önermiştir [92]. Wong ve diğerleri, çoğunluk sınıfı örneklerini seçmek üzere bulanık kural tabanlı bir sistem ortaya koymuşlardır [93]. Ng ve diğerleri, yeniden örneklemenin çeşitliliğini artırmak için çoğunluk sınıfının kümelendiği, çeşitlendirilmiş duyarlılığa dayalı bir örnek alma yöntemi önermiştir [94]. Fu ve diğerleri, çoğunluk sınıfının karakteristiğini oluşturan örnekleri korurken, çoğunluk sınıfından gereksiz örnekleri atmak için kapsamlı bir değerlendirme modeli kullanan, temel bileşen analizi tabanlı bir yöntem önermişlerdir [95]. Ha ve Lee,

çoğunluk sınıfının orijinal ve örnek alınmış dağılımları arasındaki kayıp fonksiyonunu en aza indirecek şekilde sınıflandırıcının performansını en üst düzeye çıkarmayı amaçlayan genetik algoritma tabanlı bir örnek alma tekniği sunmuşlardır [96]. Devi ve Biswajit, doğru sınıf etiketlerinin tahmininde en az katkısı olan aykırı değerlerin, fazlalıkların ve gürültülü örneklerin algılanmasının dâhil edildiği genişletilmiş TL yöntemi geliştirmişlerdir [97]. Bunkhumpornpat ve Sinapiromsaran, çoğunluk sınıfından her örnek ile azınlık kümesinin sözde merkezi arasındaki en kısa yol mesafesine dayanan bir yoğunluk fonksiyonunu belirlemek için küme grafı kullanmışlardır [98].

Aşırı örnekleme yöntemlerinin en temel formu olan ROS (Random Over Sampling), rastgele bir şekilde azınlık sınıfı örneklerinin tam kopyalarını yaparak çoğaltılmalarını sağlar [99]. Bir diğer popüler teknik olan sentetik SMOTE (Synthetic Minority Over-sampling Technique), mevcut azınlık noktaları ve komşuları arasında ilaveler yaparak rastgele yeni azınlık sınıf noktaları üretmeyi amaçlar [100]. SMOTE, sentetik örnekler oluştururken diğer sınıfların komşu örneklerinin konumlarını ihmal eder. Bu sebeple oluşabilecek sınıf çakışmalarının ihtimalini azaltmak için; yalnızca sınır çizgisi yakınındaki azınlık örnekleri aşırı örnekleyen Borderline-SMOTE [101], sınıflandırma karar sınırının zor örneklere uyarlanabilir şekilde kaydırılmasını sağlayan ADASYN (Adaptive Synthetic) [102], dikkate alınan örneklerin yerel komşulukları hakkında daha kesin bir şekilde bilgi kullanan LN-SMOTE [103] ve aynı hattaki farklı ağırlık derecesine sahip azınlık örneklerini güvenli seviyede örnekleyen Safe-Level SMOTE [104] gibi çeşitli yaklaşımlar önerilmiştir. Bunkhumpornpat ve diğerleri, keyfi olarak şekillendirilmiş bir kümeyi aşırı örneklemek için yoğunluk temelli kümeler konseptine dayanan bir yöntem önermişlerdir [105]. Abdi ve Hashemi, diğer azınlık sınıfı örnekleri ile oluşturulan sınıf ortalaması ile aynı mesafeye sahip sentetik örnekler üreten bir teknik geliştirmişlerdir [106].

Hem örnek alma hem de aşırı örnekleme kullanan hibrit yöntemler; SMOTE ile aşırı örnekleme ve kaba set temelli düzenlemeye dayalı temizlik [107], SMOTE ile birlikte eğitilmiş destek vektör makinesi tarafından oluşturulan bir hiper düzleme olan uzaklığa dayalı örnek alma [108], SMOTE ile birlikte çoğunluk sınıfından kümeleme ile örnek alma [109], azınlık sınıfındaki destek vektörleri örneklemek için SMOTE kullanırken çoğunluk sınıfındaki desteksiz vektörleri örneklemek için RUS kullanma [110] olarak sıralanabilir. Hibrit yöntemlerdeki bir diğer yaklaşım, gürültülü örnekleri kaldırdıktan sonra yeniden örnekleme işlemlerini uygulamaktır. Bu amaçla önerilen yöntemlere; sınırda azınlık

noktalarını aşırı örnekleyen ve her iki sınıf için gürültü filtresi içeren SMOTE iteratif bölümleme filtresi [111], k-NN gürültü filtresi sonrası çoğunluk sınıftan örnek alma [112] ve komşu temizleme kuralına dayalı örnek alma ve SMOTE kombinasyonu [113] örnek olarak verilebilir. Farklı bir teknik olan sınıf değiştirme ise, en yakın düşman mesafesi kullanarak çoğunluk sınıfından noktaları kaldırıp azınlık sınıfına yerleştiren kolektif öğrenme tabanlı bir yaklaşımdır [114].

3.1.2. Algoritma düzeyinde yöntemler

Algoritma düzeyinde yöntemler, çoğunluk sınıfına karşı ön yargıyı azaltmak için geleneksel öğrenme algoritmalarında değişiklik yaparlar. Maliyete duyarlı yöntemler, aktif öğrenme ve çekirdek bozma yöntemleri olmak üzere üç grupta incelenebilir.

Maliyete duyarlı yöntemlerde, azınlık sınıfına çoğunluk sınıfına göre daha yüksek bir sınıflandırma maliyeti verilir. Kullanılan sınıflandırıcı türüne bağlı olarak, maliyet seti ya tüm sınıflar için tam ağırlıklar belirtilerek mutlak şekilde ya da yalnızca sınıf içi maliyetlerin oranı belirtilerek göreli şekilde tanımlanmaktadır [115]. Cheng ve diğerleri, azınlık sınıfının marj ortalamasının maliyete duyarlı parametrelerini artırıp varyansının maliyete duyarlı parametrelerini azaltarak, azınlık sınıfının yanlış sınıflandırma cezasının arttırıldığı bir sınır dağıtım makinesi önermişlerdir [116]. Xiao ve diğerleri tarafından, sınıfa özgü maliyet düzenleme şeması ve çekirdek uzantıları, aşırı öğrenme makinesi sınıflandırıcıları ile dengesiz sınıflandırma açısından araştırılmıştır [117]. Geleneksel maliyete duyarlı öğrenme yöntemlerinin aksine, Ohsaki ve diğerleri, herhangi bir kullanıcı müdahalesi olmadan objektif bir işleve dahil ederek değerlendirme kriterlerinin değerlerini doğrudan artırabilen bir karışıklık matrisi tabanlı çekirdek lojistik regresyon sınıflandırıcı önermişlerdir [118].

Aktif öğrenme, etiketlemenin büyük etiketsiz veri kümeleri için pahalı olabileceği varsayımı altında, eğitim sürecinde yeni veri noktalarında istenen çıktıları elde etmek için kullanıcıyla veya eşdeğer bir bilgi kaynağıyla etkileşime girilmesine izin verilen yarı danışmanlı makine öğrenmesi paradigmasıdır [119]. Sınıflandırıcı önce tüm veri kümesi kullanılarak eğitilir, daha sonra sınıflandırıcıyı yinelemeli olarak güncellemek için yanlış sınıflandırılmış azınlık sınıfı örnekleri kullanılır. Bu süreçte, maliyet ayarlama veya yeniden örneklemenin kapsamı gibi çeşitli sorunlar ortaya çıkmaktadır [120]. Ferdowsi ve diğerleri, farklı örnek seçim stratejileri arasında değişen çevrimiçi bir aktif öğrenme çerçevesi önermiştir [121].

Örneklerin her seferinde bir tane seçildiği aktif öğrenme seri modu yerine, You ve diğerleri, çeşitli sorgular oluşturmayı hedefleyen bir toplu iş modu çerçevesi önermiştir [122]. Zhang ve diğerleri, hem etiketli hem de etiketsiz örneklerin daha fazla etiket elde etmek için seçilebildiği, dengesiz kitle kaynak verilerinin sınıflandırılması için aktif öğrenme tabanlı çoklu gürültü etiketleme çerçevesi önermişlerdir [123]. Guo ve Wang, en bilgilendirici etiketlenmemiş veri noktalarını seçerek başlayan çok sınıflı dengesiz bir sınıflandırıcı geliştirmek için destek vektör makineleri ile aktif öğrenmeyi birleştirmiştir [124]. Zhang ve diğerleri, dengesiz akış verilerinin belirli bir sorgu bütçesi altında sınıflandırılması için farklı aktif sorgulama stratejileri üzerinde çalışmışlardır [125]. Dengesiz akış verilerinin sınıflandırılmasını sağlayan bir diğer çalışmada ise, evrimsel hesaplama ilkeleri kullanılarak geliştirilen genetik programlama tabanlı bir çerçeve önerilmiştir [126].

Çekirdek bozma yöntemleri, Gausyan radyal temel fonksiyon çekirdeğini kullanan sınıflandırıcılara özgüdür. Amaç dengesizliğin etkisini azaltmak için çekirdeğin düzenini bozarak azınlık sınıfı örneklerinin yakınlaştırılmasıdır ve bu durum maliyet seti ayarlaması gerektirir [127]. Bu nedenle, yalnızca son yinelemede eğitilmiş destek vektör makinesi kullanmak yerine, kolektif öğrenme gibi farklı yinelemelerde eğitilmiş destek vektör makinelerini birleştirmek daha yararlıdır [128]. Mevcut yöntemler, sınıf dengesizliğiyle başa çıkabilmek için ya çekirdeğin tamamını bozmayı [129], [130] ya da sınıfa özgü bozma uygulamayı [131], [132] tercih etmektedirler.

3.1.3. Hibrit yöntemler

Hibrit yöntemler, veri düzeyinde yöntemleri algoritma düzeyinde yöntemlerle birleştirerek bu yaklaşımların dezavantajlarını azaltırken avantajlarını güçlendirmeyi hedeflemektedir. Torbalama ve artırma ile gerçekleştirilen kolektif öğrenme teknikleri [133] ve MapReduce mimarisi kullanılarak geliştirilen teknikler [4] olarak gruplandırılabilir.

Kolektif öğrenme sınıflandırıcıları ile veri dengeleme tekniklerinin birleştirildiği klasik örnekler şöyle sıralanabilir: SMOTE algoritması ve artırma prosedürünün kombinasyonu olan SMOTEBoost [134], AdaBoost.M1 algoritmasını veri oluşturma stratejisiyle birleştiren DataBoost-IM [135], rastgele bozma, örnek alma ve aşırı örnekleme ile artırmanın bileşimi olan JOUS-Boost [136], artırma ile örnek alma tekniklerinden oluşan RUSBoost [137], torbalama ile örnek alma bileşimi olan UnderBagging, torbalama ile aşırı örneklemenin

bileşimi olan OverBagging, torbalama ile SMOTE bileşimi olan SMOTEBagging [138], çoğunluk grubunun alt kümelerini azınlık grubuyla birleştirerek ve her bir sınıflandırıcı için sahte dengeli eğitim kümeleri oluşturarak birden fazla sınıflandırıcıdan öğrenen EasyEnsemble ve BalanceCascade [139] son olarak artırmada aşırı öğrenme olarak tanımlanan RAMOBoost [140]. Kolektif öğrenmede bir diğer yaklaşım ise örnek alma veya aşırı örnekleme uygulandıktan sonra maliyete duyarlı ayarlara sahip bir sınıflandırıcının kullanılmasıdır. Bu bağlamda yapılan çalışmalar ise; destek vektör makinesi ile SMOTE kombinasyonu [141], destek vektör makinesi, karar ağacı, bulanık kural oluşturma ve k-NN gibi sınıflandırıcılar ile birlikte aşırı örnekleme ve örnek alma kullanımı [142] ve verilerin örnek alma ile önceden işlenmesinin ardından rastgele orman algoritması ile sınıflandırılması [143] olarak sıralanabilir.

MapReduce tabanlı yöntemler, sınıf dengesiz büyük veri kümelerinin ön işlemesi veya analiz edilmesi süreçlerinin büyük veri analitiğine uygun olarak gerçekleştirildiği çözümlerdir. Çeşitli çalışmalara; MapReduce paradigmasının uyarlandığı [144] veya zenginleştirildiği [145] geleneksel yeniden örnekleme teknikleri, metrik öğrenme algoritmaları ve dengeleme tekniklerinin kombinasyonu [146], çeşitli büyük veri tabanlı danışmanlı öğrenme tekniklerinin iyileştirilmesi [147], evrimsel özellik ağırlıklandırma ile rastgele aşırı örneklemenin uygulanması [148], MapReduce şemasına yerleştirilmiş evrimsel örnek alma yöntemi [149], MapReduce tabanlı bir veri azaltma şeması [150], çok sınıflı dengesiz sınıflandırma için geliştirilmiş SMOTE [151], büyük veri için G-mean metriği maksimizasyonu [152], maliyete duyarlı destek vektör makinesi [153], dilsel bulanık kural tabanlı sınıflandırmanın MapReduce uygulaması [154], dağıtık ve ağırlıklı aşırı öğrenme makinesi [155] ve son olarak birçok yaklaşım ve algoritma kullanan sınıflandırma çerçevesi [156] örnek olarak verilebilir.

3.1.4. Derin öğrenme yöntemleri

Derin öğrenme (Deep Learning, DL), hiyerarşik mimarilerde birçok bilgi işlem aşamasının danışmansız özellik öğrenme, örüntü analizi ya da sınıflandırma için kullanıldığı bir makine öğrenmesi teknikleri sınıfıdır. Derin öğrenmenin özü, üst düzey özelliklerin veya faktörlerin alt düzey olanlardan tanımlandığı gözlemsel verilerin hiyerarşik özelliklerini veya temsillerini hesaplamaktır [157]. Derin öğrenme yöntemleri, geleneksel makine öğrenmesi yöntemlerinden daha iyi özellik öğrenme ve genelleştirme sağlama avantajlarını kullanarak

sınıflandırma performansını artırmayı amaçlar [158]. Bu kategorideki yöntemler, görüntülerden oluşan büyük veri kümeleri üzerinde uygulanmıştır.

Hensman ve Masko, dengesiz görüntü verilerinin ROS ile dengelenmesinin, Evrişimli Sinir Ağı olan CNN (Convolutional Neural Network) kullanılarak sınıflandırılmasında iyileştirme oluşturacağını göstermiştir [159]. Lee ve diğerleri, önceden eğitilmiş CNN ile transfer öğrenme uygulayıp orijinal verilere ince ayar yapmıştır [160]. Wang ve diğerleri tarafından, dengesiz veri kümeleri üzerinde derin ağların eğitimi için yeni kayıp fonksiyonları önerilmiştir [161]. Huang ve diğerleri, daha ayrımcı temsiller oluşturmak için yeni bir kayıp fonksiyonu ve örnekleme yöntemi kullanmışlardır [162]. Khan ve diğerleri, hem çoğunluk hem de azınlık sınıfları için güçlü özellik temsillerini otomatik olarak öğrenebilen, maliyete duyarlı bir derin sinir ağı önermiştir [163]. Ando ve Huang, SMOTE'yi CNN tarafından elde edilen derin özellik alanına genişletmek amacıyla derin aşırı örnekleme çerçevesini sunmuştur [164]. Ding ve diğerleri, sinir ağının derinliğinin artmasının yakınsama oranlarını iyileştirip iyileştirmediğini belirlemek için çok derin CNN modelleri ile deney yapmıştır [165]. Pouyanfar ve diğerleri, örnekleme oranlarını sınıf bazında performansa göre ayarlayan yeni bir dinamik örnekleme yöntemi getirmiştir [166]. Buda ve diğerleri, dengesiz çok sınıflı görüntü veri kümelerinde RUS, ROS ve iki fazlı öğrenme yaklaşımlarını karşılaştırmıştır [167]. Zhang ve diğerleri, dengesiz görüntü verilerinin sınıflandırılmasını geliştirmek için transfer öğrenimi, CNN özellik çıkarma ve küme tabanlı en yakın komşu kuralını birleştirmiştir [168].

3.2. Dolandırıcılık Tespiti

Dolandırıcılık, para veya mülk elde etmek, para veya mülk kaybını engellemek, ödeme yapmaktan kaçınmak ya da ticari veya kişisel avantaj elde etmek amacıyla fiziksel olmayan yollarla, gizlenerek ve açgözlülükle yürütülen bir veya bir dizi yasa dışı eylemdir [169].

Bir araştırma alanı olarak dolandırıcılık; hukuk, kriminoloji, psikoloji, etik, muhasebe ve yönetim gibi çeşitli disiplinler arası kaynaklarla kesişmektedir [6]. Bu kaynaklar tarafından gerçekleştirilen hasar tahmini girişimleri sonucunda maalesef net maliyet ve zarar tespiti yapılamamaktadır. Dolandırıcılıkla ilgili veri toplama sürecinde ortaya çıkan eksik veri probleminin sebep olduğu bu durumun bir nedeni dolandırıcılığın her zaman rapor edilmemesi, diğer nedeni ise dolandırıcılığın her zaman tespit edilememesidir, çünkü çoğu

zaman dolandırıcılık kurbanı olanlar bunun farkında bile değildir [170]. 2020 Küresel Ekonomik Suç ve Dolandırıcılık Araştırması; katılımcı şirketlerin %47'sinin son 24 ayda dolandırıcılık faaliyetlerine maruz kaldığını, şirket başına ortalama 6 dolandırıcılık vakasının tespit edildiğini ve toplamda 42 milyar dolar zarar yaşandığını raporlamıştır [171].

İş perspektifinden bakıldığı zaman dolandırıcılık faaliyetleri; iç dolandırıcılık ve dış dolandırıcılık olarak iki şekilde gerçekleştirilmektedir [172]. Mesleki dolandırıcılık olarak da adlandırılan iç dolandırıcılık, bir kurum ya da işletme tarafından istihdam edilen kişi veya grubun pozisyonlarını yasadışı faaliyet yürütmek için kullanmasıyla gerçekleşir. Dış dolandırıcılık ise, kurumsal varlıkları haksızca elde etmek amacıyla yaratılan dış tehditlerle ilgilidir. Dolandırıcılık girişiminin yönü ne olursa olsun, tüm işletmelerin karşılaştığı tehditler hakkında anlayış, farkındalık ve politika geliştirmeleri büyük önem taşımaktadır.

Türk Ceza Kanunu'nun [173] 157. Maddesi gereğince dolandırıcılık, hileli davranışlarla başkasını aldatarak fayda elde etmektir ve gerçekleştirilmesi durumunda bir yıldan beş yıla kadar hapis ve beş bin güne kadar adli para cezası bulunmaktadır. 158. Madde gereğince ise nitelikli dolandırıcılık; dini inanç ve duyguların istismarı, kişinin zayıflığından yararlanılması, kurum ve kuruluşların aracı olarak kullanılması, bilişim sistemlerinin araç olarak kullanılması ve yayın araçlarının sağladıkları kolaylıkların kötüye kullanılması gibi yollarla gerçekleştirilen suçlardır. Nitelikli dolandırıcılığın bedeli de üç yıldan on yıla kadar hapis ve beş bin güne kadar adli para cezasıdır. Aynı zamanda; bilişim sistemine girme, sistemi engelleme, bozma, verileri yok etme veya değiştirme, banka veya kredi kartlarının kötüye kullanılması, yasak cihaz veya programların kullanılması gibi çeşitli sahtecilik suçları da bazı durumlarda dolandırıcılık suçlarının tamamlayıcı parçası işlevini görmektedir.

Para ve hizmetleri içeren hemen hemen her teknolojik sistem, dolandırıcılık riski ile karşı karşıyadır ve dolandırıcılık eylemi meydana geldikten sonra mümkün olan en kısa sürede tanımlanmalıdır. En yaygın altı alanda meydana gelen dolandırıcılık faaliyetleri ve çözüm yolları aşağıda özetlenmiştir [5], [7].

- Telekomünikasyon: Pazardaki rekabet ve düzenlemeler her yıl daha da kalabalıklaştığı için, müşterilerinin ayrılmasını ve maliyetlerin artmasını engellemek amacıyla telekom firmalarının çok ciddiye aldığı bir problemdir. Uygulamalar, çevrimiçi ve gerçek

zamanlı güvenlik sistemlerine ya da çağrı ayrıntısı ve abonelik bilgilerindeki olağandışı davranışları sınıflandıran ön faturalandırma sistemlerine odaklanmaktadır.

- Kredi Kartı: Kredi kartı dolandırıcılığı çevrim dışı ve çevrim içi olmak üzere iki şekilde gerçekleştirilmektedir. Çevrim dışı dolandırıcılık, çalıntı bir fiziksel kart kullanılarak yapılır ve fark edilmesi durumunda izinsiz kullanılmadan önce kartı veren kurum tarafından kart kapatılır. Çevrim içi dolandırıcılık ise; web veya telefon aracılığıyla ve çoğunlukla güvenli elektronik ticaret altyapılarıyla gerçekleştirilir. Buradaki esas problem, çarpık ve dengesiz verinin analizine uygun çözümler önermektir.
- Sigortacılık: Sahte iddiaları tespit etmek, gizli ilişkileri belirlemek ve hatta gizli suç halkalarını ortaya çıkarmak; sigorta endüstrisinde karlılığı korumanın ana kaynağıdır. Tespit sürecinde amaç; en sık rastlanan sağlık sigortası, hayat sigortası ve otomobil sigortası gibi kayıtları gerçek zamanlı veya toplu analiz teknikleri kullanarak otomatik olarak puanlanmaktadır.
- Denetim: Hileli finansal tablolar yayınlayan firmaların tespit edilmesi ve kurumsal iflasın öngörülmesi amacıyla genellikle finansal tabloların dikeyde ve yatayda analizi, oranların incelenmesi ya da yeni metriklerin tanımlanması ile tutarsızlıklar tespit edilmektedir.
- Sağlık: Verilmeyen hizmetler ve ürünler için faturalandırma ya da gereksiz hizmetler gibi sebeplerden kaynaklanan sağlık harcamalarındaki büyümenin üstesinden gelmek için uygunsuz veya olağandışı davranışları tespit etmek amacıyla, bu alandaki çözümler genellikle karar alıcıların uygulama alışkanlıklarının segmentasyonuna odaklanmaktadır.
- İnternet ve Sosyal Medya: İnternet, çok fazla zaman veya para harcamadan geniş kitlelere ulaşmayı sağlar. Web siteleri, çevrimiçi mesajlaşma uygulamaları veya sosyal medya platformları bu süreci kolaylaştıran yollardır. Kurgu siteler ile insanların meraklarını veya ihtiyaçlarını sömüren bu tür dolandırıcılıklardan çoğunlukla kişisel farkındalık ile korunmak mümkündür.

Tüm dolandırıcılık eylemlerinde, dolandırıcılık üçgeni olarak betimlenen, üç ortak unsur bulunmaktadır: teşvik, fırsat ve rasyonelleştirme [174]. Teşvik, dolandırıcılık failinin motivasyonu olan finansal ihtiyaç, işi ile ilgili hayal kırıklıkları, sonuçları gerçek performanstan daha iyi raporlama isteği ve sistemi yenme arzusu gibi durumları ifade etmektedir. Fırsat, zayıf kontrol sistemleri veya belirsiz politikalar gibi zafiyetlerin,

dolandırıcılığın sürdürülmesi için koşulları elverişli kılması durumudur. Son olarak rasyonelleştirme, dolandırıcılık faillerinin eylemlerini gerçek dışı kurallara uygun ve kabul edilebilir bir tutum çerçevesine yerleştirerek gerçekleştirmeleridir. Ek olarak önerilen bir diğer unsur ise dolandırıcılığın gerçekten gerçekleşip gerçekleşmeyeceği konusunda büyük rol oynayan dolandırıcının kabiliyetidir [175]. Dolandırıcılık sürecinde; fırsat kapıyı açar, teşvik ve rasyonelleştirme potansiyel dolandırıcıyı açık kapıya doğru çeker, ancak birey bu açıklıktan geçebilecek yetkiye, zayıflıkları anlama ve kullanma kapasitesine ve stresle başa çıkma yeteneğine sahip olmalıdır.

Dolandırıcılık failleri, suçlarını işlerken en düşük maliyetle en büyük faydaları sağlamayı amaçlamaktadırlar. Bu sebeple, eylemlerini yer değiştirme teorisini kullanarak gerçekleştirmektedirler. Yer değiştirme teorisinin arkasındaki fikir, motive edilen suçluların caydırılması durumunda, başka yerlerde suç işlenmesidir ve aşağıda verilen altı olasılık doğrultusunda gerçekleştirilmektedir [176].

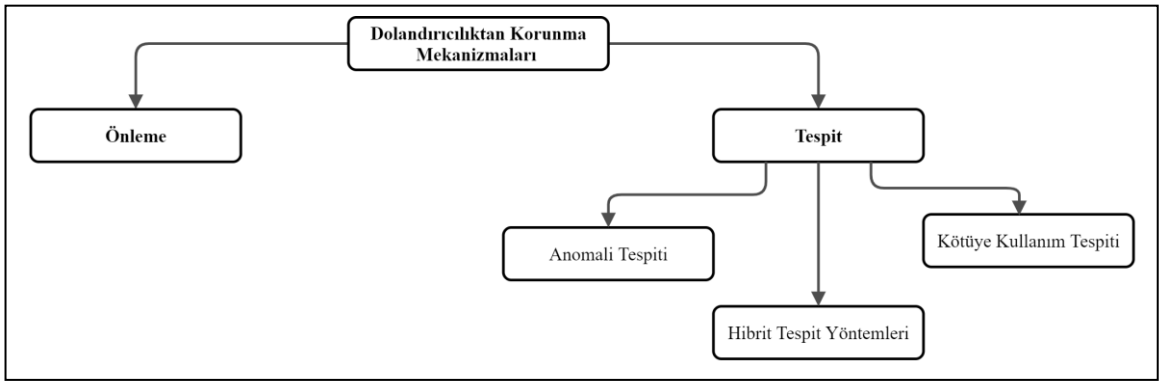
1. Coğrafi Yer Değiştirme: Suç bir yerden başka bir yere taşınabilir.
2. Zamansal Yer Değiştirme: Suç bir zamandan diğerine taşınabilir.
3. Hedef Yer Değiştirme: Suç bir hedeften diğerine yönlendirilebilir.
4. Taktiksel Yer Değiştirme: Suç işlemek için bir yöntem diğeriyle ikame edilebilir.
5. Tür Yer Değiştirme: Suç başta türden bir suç ile ikame edilebilir.
6. Fail Yer Değiştirme: Yakalanan suçluların yerine yenileri gelebilir.

Olası yer değiştirmeleri öngörmek için birbiriyle ilişkili mekanizmaların tasarlanması ve uygulanması gerekmektedir. Dolandırıcılık ile mücadele sürecinde, dolandırıcılıkları engellemek, azaltmak ve yeni tehditleri yakalamak için, Şekil 3.2'de verildiği üzere; önleme ve tespit mekanizmaları kullanılmaktadır [177].

Önleme, sistemleri dolandırıcılığa karşı koruyan ilk koruma katmanıdır ve bilgisayar sistemlerinde, ağlarda veya verilerde istismarların oluşumunu kısıtlar, bastırır, yok eder, kontrol eder ya da kaldırır. Bu mekanizmaya; verilerin şifrelenmesi, güvenlik duvarı kullanımı, ağ akışının aktif incelenmesi gibi aksiyonlar örnek olarak verilebilir.

Tespit ise önlemeden sonraki koruma katmanıdır. Dolandırıcılık tespiti, sisteme izinsiz girişleri ve sahte etkinlikleri keşfetmeyi, tanımlamayı ve bunları sistem yöneticisine

bildirmeyi amaçlar. Dolandırıcılık, başlangıçta manuel denetim teknikleri kullanılarak keşfedilmeye çalışılmasına rağmen, bu karmaşık ve zaman alıcı teknikler yerlerini otomatik dolandırıcılık tespit sistemlerine bırakmışlardır. İyi bir dolandırıcılık tespit sisteminin ise, çarpık dağılıma sahip ve gürültü barındıran veriler üzerinde yasal işlemlere çok benzeyen ve hileli işlemleri tespit edebilir ve değişikliklere uyum sağlayabilir olması gerekmektedir [178].

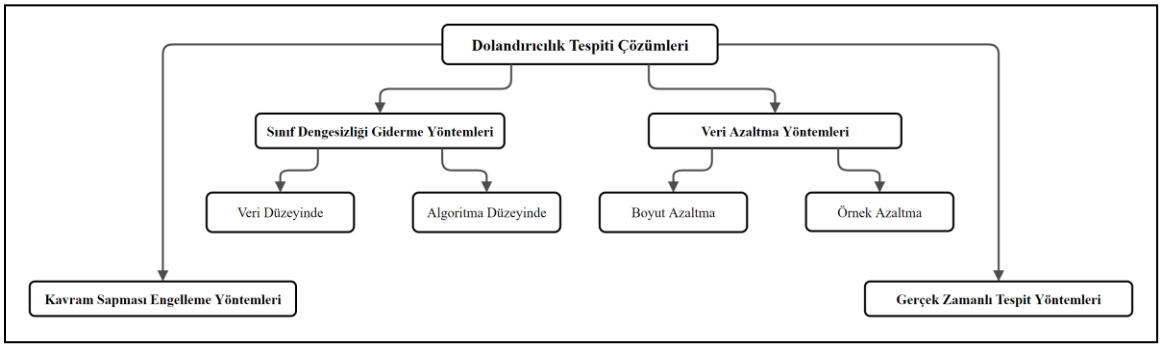


Şekil 3.2. Dolandırıcılıktan korunma mekanizmaları

Son yıllarda dolandırıcılık tespiti için; aykırılık temelli yaklaşımlar, kötüye kullanıma dayalı yaklaşımlar ve hibrit yaklaşımlar önerilmektedir [179]. Anomali veya aykırı durum algılama tabanlı yaklaşımlar, her bireyin davranışının modellenerek normalden sapma durumlarının izlendiği davranışsal profil oluşturma yöntemlerine dayanır. Kötüye kullanım algılama yaklaşımında ise, bilinen dolandırıcılık davranışları dolandırıcı imzaları olarak tanımlanıp diğer davranışlar normal olarak tanımlanır ve şüpheli işlemleri tespit etmek için kural tabanlı, istatistiksel veya sezgisel yöntemler kullanır. Anomali tespiti, genelleme kabiliyeti eksikliğinden yüksek yanlış alarm oranlarının varlığından mustaripken, kötüye kullanım tespiti ise yeni dolandırıcılık türlerini tespit edemez. Bu sebeple, her iki yaklaşımı kapsayan hibrit yöntemler de önerilmektedir. Dolandırıcılık faaliyetleri tespit edildiğinde ve ilgilendirdiği sahtekârlık türü doğrulandığında ise genellikle, dolandırıcılık vakasında kullanılan hileli mekanizmalar ve açıklıklar ayrıntılı bir şekilde araştırılarak düzeltici ve önleyici tedbirler alınmaktadır [180].

Dolandırıcılar ve savunucular arasında sürekli gelişen çatışmada; veri toplama, algılama, hafifletme ve adli bilişim olmak üzere dört bileşen bulunmaktadır [181]. Dolandırıcılık ve kötüye kullanımı önlemek ve tespit etmek için herhangi bir strateji veri toplama aşaması ile

başlar. Maliyet ve depolama kapasitesi gibi kısıtlamalar yüzünden ilgili tüm verinin toplanması genellikle imkânsıza yakındır ve ne kadar verinin yeterli olacağına dair kestirimde bulunmak da zordur. Tespit aşamasındaki belirleyici unsur zamandır. Bir dolandırıcılık dakikalar içinde gelişiyorsa, ancak işleme katmanını veriyi analiz etmesi saatler sürüyorsa, etkili bir şekilde yanıt verilmesi olanaksızdır. Saldırı anlaşıldıktan sonraki aşama ise hafifletmedir. Hafifletme, mevcut veriyi üzerinde işlem yapılabilecek yeni değişkenlere dönüştürme yeteneği sonucunda, istismarın sonlandırılmasını ifade etmektedir. Son olarak, adli bilişim bir saldırının doğasını ve gelecekte bu tür saldırıların nasıl önleneceğinin veya sınırlandırılacağının keşfedilmesini sağlar. Bu aşama sonucunda, yeni analiz süreçleri ve gözetim mekanizmaları devreye sokulabilir.



Şekil 3.3. Dolandırıcılık tespiti çözümleri

Dolandırıcılık tespiti sistemleri, zorlu teknik ve metodolojik problemler içerdiği için; başarısız olmaya eğilimlidir, düşük doğruluk oranlarına sahiptir ve çok sayıda yanlış alarm üretmektedir [182]. Dolandırıcılık tespitindeki sorunlar ve zorluklar açısından literatürdeki çalışmalar değerlendirildiğinde, Şekil 3.3'de verildiği üzere; kavram sapması engelleme yöntemleri, sınıf dengesizliği giderme yöntemleri, veri azaltma yöntemleri ve gerçek zamanlı tespit yöntemleri olarak dört kategoride incelenebilir.

3.2.1. Kavram sapması engelleme yöntemleri

Kavram sapması, meşru kullanıcıların veya dolandırıcıların davranışlarının sürekli olarak değiştiği dinamik ortamı ifade eden kavramdır [183]. Örneğin, bir kredi kartı kullanıcısının işlem tutarı ve harcama alışkanlıkları ile ilgili davranışları, yaşam tarzındaki veya gelirindeki değişikliklerden etkilenmektedir. Aynı zamanda, dolandırıcılık türleri ve teknikleri de zamanla değişmektedir. Girdi verileri ve hedef değişken arasındaki ilişki zaman içinde

değiştiğinden, kavram sapması genellikle çevrimiçi danışmanlı öğrenme senaryosu ile ifade edilmektedir.

Malekian ve Hashemi, durağan olmayan davranışlarla başa çıkmak ve dolandırıcılık tespit modelini dinamik olarak güncellemek amacıyla bir kavram sapması yönetimi çerçevesi önermiştir [184]. Pozzolo ve diğerleri tarafından, biri kolektif öğrenme diğeri kayan pencere tabanlı olan iki yaklaşım önerilmiş ve değerlendirme için geri bildirimlerin ve gecikmiş etiketlerin kullanıldığı bir kazanma stratejisi geliştirilmiştir [185]. Kulkarni ve Ade, durağan olmayan kredi kartı risk değerlendirmesi ortamında, mevcut teknikler üzerinde verimlilik ve iyileştirmeyi korurken, girdi verilerinin dengesiz doğasını artımlı bir şekilde analiz eden algoritmik bir çerçeve geliştirmişlerdir [186]. Veeramachaneni ve diğerleri; büyük veri tabanlı davranışsal analiz platformu, aykırı değer tespiti, analistlerinden geri bildirim alma mekanizması ve danışmanlı öğrenme modülünden oluşan uçtan uca bir sistem sunmuştur [187]. Somasundaram ve Reddy ise, hem kavram sapmasını hem de sınıf dengesizliğini etkili bir şekilde ele almak için artımlı ve maliyete duyarlı öğrenme ile ağırlıklı oylama tabanlı birleştirici kullanan bir model önermiştir [188].

3.2.2. Sınıf dengesizliği giderme yöntemleri

Dolandırıcılık tespitinde sınıf dengesizliği sorunu, normal örnekten çok daha az sayıda hileli örneğin olduğu durumu işaret etmektedir. Geleneksel sınıf dengesizliği çözümlerindeki gibi veri seviyesinde ve algoritma seviyesinde yaklaşımlar önerilmiştir.

Sisodia ve diğerleri, çeşitli aşırı örnekleme teknikleri kullandıktan sonra maliyet duyarlı ve kolektif sınıflandırıcılar ile kredi kartı dolandırıcılığını tespit etmeyi amaçlamışlardır [189]. Perez ve diğerleri, birden fazla alt örnek ile uyarılan ve açıklama kapasitesi kaybı olmayan sınıflandırma ağacı kullanarak araba sigortacılığı dolandırıcılığı sorununa çözüm üretmişlerdir [190]. Hordri ve Farhan, hem farklı sınıflandırma algoritmalarının dolandırıcılık tespitindeki başarısını analiz etmek hem de RUS, ROS ve SMOTE yöntemlerinin sınıflandırma performansına etkisini araştırmak amacıyla çeşitli karşılaştırmalar yapmışlardır [191]. Bauder ve diğerleri ise, farklı IR değerlerine sahip veri kümeleri üzerinde hem farklı örnekleme yöntemlerinin hem de farklı sınıflandırma algoritmalarının etkisini göstermek amacıyla deneyler gerçekleştirmişlerdir [192].

3.2.3. Veri azaltma yöntemleri

Dolandırıcılık veri kümesinin büyük ölçekli ve yüksek boyutlu olması tespit sürecini hem yavaşlatır hem de son derece zor ve karmaşık hale getirir. Bu sebeple çeşitli dolandırıcılık tespiti çalışmaları boyut azaltma veya örnek azaltma yöntemlerini kullanmaktadır. Boyut azaltma, hem sınıf dengesizliğini gidermek amacıyla hem de örneklem almak amacıyla gerçekleştirilmektedir.

Moepya ve diğerleri, finansal tablo dolandırıcılığını tespit etmek amacıyla temel bileşen analizi ve faktör analizi kullanarak seçtikleri öznitelikleri ağırlıklı destek vektör makineleri ile sınıflandırmışlardır [193]. Bahnsen ve diğerleri, işlem zamanının periyodik davranışını analiz etmeye dayalı yeni bir özellik kümesi oluşturma stratejisi önermişlerdir [194]. Herland ve diğerleri, çeşitli sağlık dolandırıcılığı veri kümelerini birleştirip etiketleyerek, bu veri kümelerinin sınıflandırma başarılarını karşılaştırmışlardır [195]. Li ve diğerleri, otomobil sigortası dolandırıcılığı tespit için, temel bileşen analizi ve rastgele orman algoritmasını birleştirerek, uyarlanabilir en yakın komşu perspektifinden analizler gerçekleştirmişlerdir [196]. Shiguihara-Juarez ve Murrugarra-Llerena ise, kredi kartı dolandırıcılığı veri kümesindeki özelliklerin alabileceği olası değerlerin sayısını ifade eden kardinalitesini azaltmak için bir yöntem önermiş ve sınıflandırmada %30 başarı artışı elde etmişlerdir [197].

3.2.4. Gerçek zamanlı tespit yöntemleri

Tespit edilmeye çalışılan dolandırıcılığın türüne göre tespit yöntemi gerçek zamanlı ya da toplu olarak gerçekleştirilmektedir. Örneğin, kredi kartı alanındaki çevrimiçi ödeme uygulamasındaki bir sahtekârlığın anında tespit edilmesi ve uygun bir şekilde yanıtlanması gerekirken, geriye yönelik analizler anında tepki vermek yerine algılama sürecini optimize etmek amacıyla gerçekleştirilir.

Khan ve diğerleri tarafından, gerçek zamanlı senaryoda kredi kartı sahtekârlığı tespitine yönelik yapay sinir ağlarını eğitmek için evrimsel bir simüle tavlama algoritması kullanılmıştır [198]. Niu ve diğerleri, düşük hesaplama maliyetine sahip, hem bilinen hem de yeni telekomünikasyon dolandırıcılık modellerini etkili bir şekilde tespit eden, kayıpları en aza indirgemek için gerçek zamanlı puanlar üreten ve güncelleyen, birleşik akıllı

puanlama algoritmasını önermişlerdir [199]. Manunza ve diğerleri, çevrimiçi dolandırıcılık tespiti gerçekleştirebilen ve aynı zamanda çevrimiçi ödeme sistemleri tarafından oluşturulan olayları düzgün bir şekilde toplayarak standart çağrı detay kayıtları üretebilen dinamik olarak yapılandırılabilir bir sistem geliştirmişlerdir [200]. Carcillo ve diğerleri ise; dengesizlik, durağan olmama ve geri bildirim gecikmesi gibi sorunlarla başa çıkabilen bir makine öğrenme yaklaşımını büyük veri analitiği araçları ile birleştiren ölçeklenebilir gerçek zamanlı dolandırıcılık bulucu çerçevesini önermişlerdir [201].

3.3. Sınıf Dengesiz Dolandırıcılık Tespiti için Büyük Veri Analizi

Dolandırıcılık tespiti, güncel uzay ve zaman gereksinimlerine uyarlanmadığında gürbüz modeller üretmek ve düşük yanlış negatif elde etmek için büyük veri analizi zorlaşmaktadır. Bu sebeple hem geleneksel dolandırıcılık tespiti yöntemlerinin kısıtlarını kaldırmayı hem de büyük verinin yol açtığı sorunları gidererek fayda oluşturmayı amaçlayan sınırlı sayıdaki çalışma, sınıf dengesiz dolandırıcılık tespiti için büyük veri analizi yaklaşımlarını tercih etmektedir.

Kamaruddin ve Ravi, kredi kartı dolandırıcılığını tespit etmek için Spark hesaplama çerçevesini kullanarak tek sınıflı sınıflandırma için parçacık sürüsü optimizasyonu ve otomatik ilişkisel sinir ağı hibrit mimarisini uygulamıştır [202]. Bauder ve Khoshgoftaar, sağlık dolandırıcılığı barındıran bir dengesiz ve büyük veri kümesini RUS kullanarak farklı sınıf dağılımları olan sentetik veri kümeleri üretmek için kullandıktan sonra bu yeni veri kümelerinin sınıflandırma performanslarını değerlendirmişlerdir [203]. Yine Bauder ve Khoshgoftaar tarafından gerçekleştirilen bir çalışmada, farklı IR değerlerine sahip veri kümeleri üzerinde dolandırıcılık tespiti başarıları karşılaştırılmıştır [204]. Belirlenen senaryo dâhilinde elde edilen sonuçlara göre, tespit için en iyi dağılımın 90:10 ve en kötü dağılımın 99.9:0.1 olduğu görülmüştür. Ayrıca, 50:50 ile 99:1 arasında da bariz bir fark olmadığı tespit edilmiştir.

Ayrı ayrı sınıf dengesiz büyük veri analizi yöntemleri ve dolandırıcılık tespiti yöntemlerine ek olarak bütüncül şekilde sınıf dengesiz dolandırıcılık tespiti için büyük veri analizi yöntemleri beraber değerlendirildiğinde, literatürde bu problemlerin farklı bakış açılarıyla ele alınarak pek çok çözüm önerildiği görülmektedir. Buna rağmen, aşağıda verilen, eksik

kalan ve açıklık getirilmesi gereken konular ayrıntılı bir şekilde değerlendirildiğinde daha kaliteli sonuçların elde edilmesi kaçınılmazdır [205], [206].

- Sınıf dengesiz büyük veride olası iki senaryo vardır. İlk senaryoda, çoğunluk sınıfın fazla ve azınlık sınıfın az olduğu bir dengesizlik vardır ancak her iki sınıftan kayıtlar bol miktarda bulunur. Önerilen çözümler de büyük ölçekli analitiklere uyarlanmalıdır. İkinci senaryo ise sınıf dengesizliğinin öğrenme güçlüklerinin ana kaynağı olamayabileceği kayıtlarla ilgilidir. Bu durum, azınlık sınıfının yapısının ve örneklerinin derinlemesine analizini gerektirir. Ek olarak, her bir zor bölgenin yerel analizi ve her birine ayrı ayrı çözümlerin üretilmesi gerektiren karmaşık senaryolar da olasıdır.
- Büyük veri analitiği süreci temelde elde edilecek değer ile ilgilidir. Bu sebeple, süreç yönetiminde; dengesizliğin kaynağı nedir, hangi tür nesneler en zor olanlarıdır, üst üste binme ve gürültünün olduğu yerler nerelerdir gibi sorulara cevap aranmalıdır.
- Sınıf dengesizliği problemi, çeşitlilikten kaynaklanan zorluklar ile birleşince büyük veri analiz sürecini kısıtlamaktadır. Çeşitli türdeki verileri sayısal değerlere dönüştürmeye çalışmak yerine, bu karmaşık yapıların doğrudan işlenmesini sağlayacak önişleme ve öğrenme algoritmalarının tasarlanması gerekmektedir.
- SMOTE tabanlı aşırı örnekleme yöntemleri MapReduce gibi dağıtık ortamlarda başarısız olma eğilimindedir. Rastgele parçalara ayrılmış veri üzerinde çalışan fonksiyonlar, uzamsal ilişkileri olmayan gerçek örneklerden yapay örnekler oluşturulmasına sebep olabilir. Bu sebeple ya örnekler arasındaki ilişkileri koruyan veri bölümleme yöntemleri ya da süreci denetleyecek hakemlik birimlerine ihtiyaç duyulmaktadır.
- Büyük verinin kritik avantajlarından biri de, pek çok veri kaynağından alınan kayıtların birleşiminden oluşmasıdır. Fakat her kaynak farklı veri modellerine dayandığı için uyum zorlukları ortaya çıkmaktadır. Dolandırıcılık analiz sistemlerinin başarısının artırılması için, mümkün olduğunca çok faklı kaynaktan elde edilen kayıtların verimli bir şekilde ilişkilendirilmesi gerekmektedir.
- Yalnızca geçmişte meydana gelen vakalardan oluşan eğitim veri kümesine dayalı olarak dolandırıcılık desenlerini algılayan modeller, yeni hileli yolları ve stratejileri tespit edemez. Bu sebeple, değişen dolandırıcılık ortamına ayak uydurabilecek, dinamik modellerin geliştirilmesi gerekmektedir.

4. MAPREDUCE TABANLI ÖNERİLEN YÖNTEMLER

Bu bölümde MapReduce tabanlı iki farklı dolandırıcılık tespiti yöntemi önerilmiştir. İlk yöntem, Dinamik İmzalar ile Dağıtık Dolandırıcılık Tespiti (Distributed Fraud Detection with Dynamic Signatures, DFDDS) olarak adlandırılmıştır. DFDDS, davranış anomalileri doğrultusunda telekom dolandırıcılığı aktivitelerini büyük veri kümelerinde ölçeklenebilir bir şekilde tespit etmeyi amaçlamaktadır. İkinci yöntem ise, Dolandırıcılık Tespiti için Dağıtılmış Küme Tabanlı Yeniden Örnekleme (Distributed Cluster-based Resampling for Fraud Detection, DCRFD) olarak isimlendirilmiştir. DCRFD, veri seviyesinde sınıflar arası ve sınıf içi dengesizlikleri gidererek büyük veri kümelerinde ölçeklenebilir bir sınıflandırma sağlamayı amaçlamaktadır.

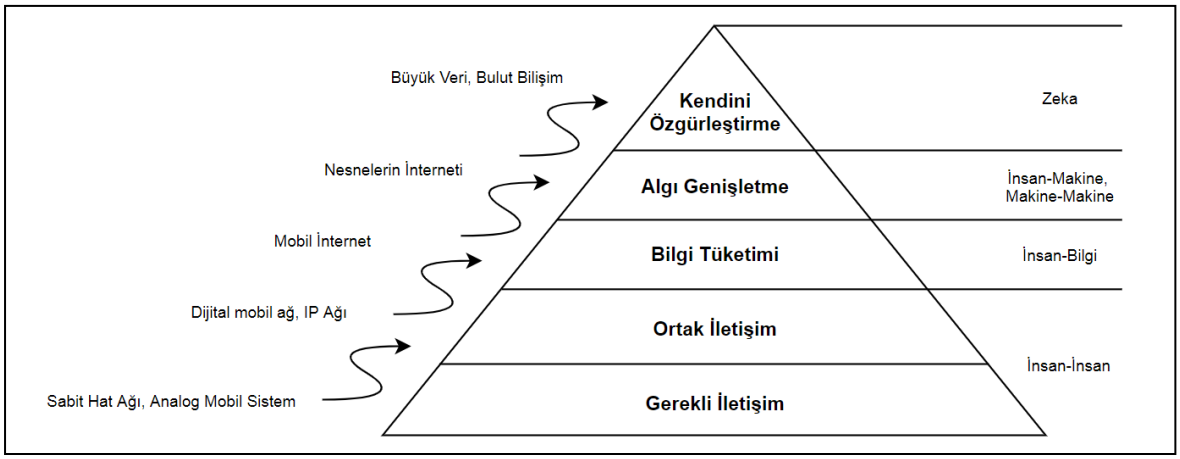
Önerilen yöntemler, Gazi Üniversitesi Büyük Veri ve Bilgi Güvenliği Merkezi (GAZI BIDISEC) altyapısındaki büyük veri teknik ve teknolojileri kullanılarak geliştirilmiştir. Analiz ve test için kullanılan donanım ortamı 6 x Compute / Storage Nodes, 2 x 18-Core 2.3 GHz Intel ® Xeon, 8 x 16GB DDR4 Memory, and 12 x 8TB 7,200 RPM High Capacity SAS Disks'den oluşmaktadır. Önerilen yöntemler, Apache Spark ve Apache Hive kullanılarak geliştirilmiştir. Detaylı açıklamalar ve elde edilen sonuçlar ilerleyen alt bölümlerde sunulmuştur.

4.1. Dinamik İmzalar ile Dağıtık Dolandırıcılık Tespiti

İletişim; bir yerden, kişiden veya gruptan diğerine bilgi aktarma eylemidir ve toplumun hem varlığı hem de gelişimi için temel koşul ve aynı zamanda bireysel bir sosyal ihtiyaçtır [207]. İletişim ihtiyaçları, Maslow'un ihtiyaçlar hiyerarşisinden esinlenerek geliştirilmiş olup Şekil 4.1'de gösterildiği gibi beş seviyeli bir model olarak incelenmektedir [208].

Gerekli iletişim, iletişim araçlarının yetersizliği ve yüksek maliyeti nedeniyle yalnızca en acil ve temel iletişim ihtiyaçlarının karşılanmasıdır. Ortak iletişim, maliyetlerin düşmesi ve iletişim araçlarının artması ile bağlantı kurabilmenin kolaylaşması ve iletişim trafiği miktarı hızla artmasını ifade eder. Bilgi tüketimi seviyesindeki temel ihtiyaç iletişim araçlarının çeşitliliğidir ve bu da internet geniş bant ihtiyacını keskin bir şekilde artırır. Algı genişletme, iletişim deneyiminin daha da genişleyerek dünyanın her köşesini kapsadığı ve insan ile

makine arasındaki bağlantının arttığı seviyedir. Önceki dört seviye, bağlantının ölçeği ve türüne odaklanırken, kendini özgürleştirme seviyesi bağlantının iyileştirilmesiyle üretilen bilgiye daha fazla önem verir. Son seviyede, iletişim sisteminin akıllı hale gelerek insanların kendini özgürleştirmesini sağlayacağı öngörülmektedir.



Şekil 4.1. İletişim ihtiyaçları hiyerarşisi [208]

Etimolojik olarak “uzaktan iletişim” anlamına gelen Telekomünikasyon veya Telekom; telefon, uydular, radyo ve televizyon yayıncılığı, fiber optik ya da internet gibi elektronik teknolojiler yoluyla ses, mesaj, görüntü ve video gibi bilgilerin gönderilmesi ve alınması sürecini ifade eder, kısaca bilginin belirli bir mesafeden elektronik yollarla iletilmesidir [209]. Telekom, çoğunlukla telefon ağını referans alır ve bu ekosistem aşağıda belirtilen beş ana bileşen üzerine kuruludur [210].

- Genel Anahtarlamalı Telefon Ağı (Public Switched Telephone Network, PSTN): Telefon ağlarının tarihi çekirdeği olan bu ağ, analog ses sinyallerini iletmek için devre anahtarlama teknolojisini kullanan bakır telefon kablolarından oluşur. Anahtarlar, arayanın telefonundan arananın telefonuna özel bir fiziksel devre oluşturarak çağrı kurulumunu yönetir.
- Tümleşik Servisler Dijital Ağı (Integrated Services Digital Network, ISDN): ISDN, bakır hatlar üzerinden dijital iletim sağlar ve veri ya da ses iletimi için fiziksel telefon hattını çoğaltır. Dijital ağlar ve bant dışı sinyalleşme çeşitli güvenlik sorunlarını çözerek; telefona sesli posta, çağrı yönlendirme ve arayan kimliği ekranı gibi yeni özellikler getirmiştir.

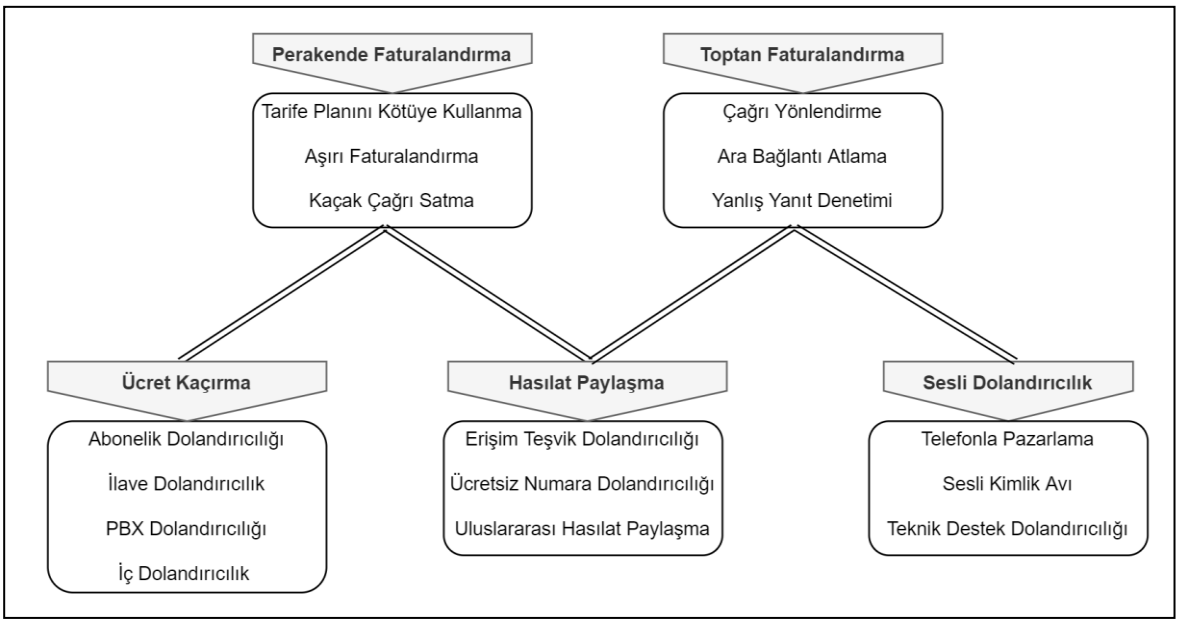
- Mobil Ağlar: Hücre siteleri veya baz istasyonları olarak bilinen sabit konumlardaki alıcılar ve vericiler tarafından kablosuz olarak bağlanan bir iletişim ağıdır. Her mobil iletişim nesli, iletişimi ve müşteri kimliğini işlemek için özel bir tür şifreleme ve ekipman kullanır.
- İnternet Protokolü Üzerinden Ses (Voice over Internet Protocol, VoIP): Analog telefon sinyallerini internet üzerinden gönderilebilecek dijital sinyallere dönüştürerek daha az maliyet ve daha çok işlevsellikle iletişim sağlar. Bu avantajlarıyla VoIP, geleneksel PSTN'ye güçlü bir alternatiftir.
- Özel Santral (Private Branch Exchange, PBX): Bir kurum içindeki tüm hatların belirli sayıda harici telefon hattını paylaşmasını sağlarken, yerel hatlardaki kullanıcılar arasında aramaları değiştiren bir telefon sistemidir. PSTN'nin aksine PBX'in amacı, telefon şirketinden her kullanıcı için bir hat talep etme maliyetinden tasarruf etmektir. Ayrıca IP-PBX, kurum içi IP telefonları internet protokolü üzerinden bağlayabilir.

Telekom, pazara girmeye devam eden yeni teknolojiler ve altyapılar sayesinde sürekli gelişen canlı bir sektördür. 2020 Küresel Telekom İstatistikleri doğrultusunda, dünya çapında yaklaşık 7,7 milyar aktif mobil geniş bant abonesi ve 1,1 milyardan fazla sabit geniş bant abonesi bulunduğu bilinmektedir [211]. Telekom sektöründeki hızlı gelişim, sistemlerin kötüye kullanma biçimleri olan dolandırıcılık faaliyetlerinin de oldukça büyük bir piyasa değerine ulaşmasına sebep olmuştur. İletişim Sahteciliği Kontrol Derneği tarafından yayımlanan 2019 Küresel Telekom Dolandırıcılık Anketi sonuçlarına göre, dolandırıcılık kayıpları toplam gelirlerin %1,74'ü olan 28,3 milyar dolara ulaşmıştır [212]. Telekom dolandırıcılığı temel olarak; çeşitli nedenlerden dolayı sistemdeki zayıflıklardan yararlanmak için karmaşık teknikler kullanılarak gerçekleştirilen faaliyetlerdir. Bu tanım aşağıdaki gibi genişletilebilir [213].

- Neden: Telekom şebekesinin karakteristiği olan eski sistemler, çeşitli operatörler ve birbirine bağlı teknolojiler gibi zayıflıklara neden olan durumlardır.
- Zayıflık: Güvenlik mekanizması eksikliği, uluslararası yasa uygulama güçlüğü, faturalandırmanın geç kullanılabilirliği veya insan ihmali nedeniyle güvenlik açıklarının ortaya çıkmasıdır.
- Teknik: Çağrı yönlendirme, numara elde geçirme, trafik pompalama ya da sosyal mühendislik gibi zayıflıkları kötüye kullanmak için çeşitli mekanizma veya hizmet kullanılmasıdır.

- Yarar: Dolandırıcılık faaliyetinin sonucunda; ödemeden kaçınmak, hizmet satmak, rekabet avantajı elde etmek ya da itibar kazanmak gibi faydalar sağlanabilir.

Hem telekom hizmeti sağlayıcılarını hem de işletmeleri ve müşterileri mağdur eden telekom dolandırıcılığı, genel olarak hizmetlerin çalınmasını veya diğer dolandırıcılık faaliyetlerini gerçekleştirmek için bu hizmetlerin kullanılmasını amaçlamaktadır ve Şekil 4.2'de verildiği üzere çeşitli şekillerde gerçekleştirilmektedir.



Şekil 4.2. Telekom dolandırıcılığı türleri

Ücret kaçırma dolandırıcılığı, arama ücretlerini ödemek zorunda kalmadan arama yapmayı amaçlamaktadır [214]. Abonelik dolandırıcılığı, ücret ödemekten kaçınmak için bir hizmete abone olurken çalınmış kimlik bilgilerinin ya da sahte bilgilerin kullanması işlemidir. Bu durumda gerçekleştirilen tüm faaliyetler gayrimeşru olacaktır. İlave dolandırıcılık, hücresel klonlama veya hırsızlık yoluyla meşru bir müşterinin hesabı üzerinde hakimiyet kurularak çağrı ücretlerinin bu hesaba yüklenmesini amaçlamaktadır. Bu durumda normal ve anormal arama örüntüleri beraber görülmektedir. PBX dolandırıcılığında, güvenliği ihlal edilmiş PBX'ler ücretsiz arama yapmak için kullanılarak arama ücretleri PBX sahibi olan kuruma yüklenir. İç dolandırıcılık ise kullanıcı hesaplarına, tarife planlarına ve faturalandırma sistemine erişimi olan bir telekom şirketi çalışanı tarafından gerçekleştirilen, faturalandırmayı devre dışı bırakma, arama kayıtlarını değiştirme veya tarife planını değiştirme gibi usulsüz eylemlerdir.

Perakende faturalandırma dolandırıcılığı müşteri, operatör veya üçüncü parti servis sağlayıcıları tarafından gerçekleştirilen usulsüzlüklerdir [214]. Tarife planını kötüye kullanma, faturalandırma sistemlerinin karmaşıklığı veya kafa karıştırıcı fiyatlandırma politikaları nedeniyle tarife planlarındaki tutarsızlıkların veya hataların müşteriler tarafından istismar edilmesiyle ortaya çıkan durumdur. Aşırı faturalandırma, operatörlerin gelirlerini gayri meşru bir şekilde artırmak için müşterilerine veya iş ortaklarına çeşitli usulsüzlüklerle uyguladığı yetkisiz ücretlendirmelerdir. Kaçak çağrı satma, ücret kaçırma dolandırıcılığı teknikleriyle elde edilen çağrıların piyasa fiyatlarından daha düşük fiyatlara satıldığı sahtekârlık sürecidir.

Toptan faturalandırma dolandırıcılığı, pazardaki operatörler arası faturalandırma işlemlerinde yapılan sahtekârlıkları kapsamaktadır [215]. Çağrı yönlendirme, telekom şebekelerinde rota saydamlığının eksikliğinden kaynaklanır ve arbitraj fırsatlarından yararlanmak için, çağrının daha düşük maliyetli ara bağlantı anlaşmalarına sahip bir ülkeye gönderilmesiyle gerçekleştirilir. Ara bağlantı atlama veya gri yönlendirme, meşru ağ geçitlerinden ve uluslararası sonlandırma ücretlerinden kaçınmak için SIM (Subscriber Identification Module) kutusu veya PBX kullanarak, gayri meşru ağ geçidi değiş tokuşlarının kullanımı olarak tanımlanabilir. Yanlış yanıt denetimi dolandırıcılığı, operatörlerin hileli süreçler kullanarak çağrılardan elde edilecek geliri artırmalarını amaçlar. Bu süreçler, operatörün; çağrıyı gerçek şebekeye aktarmak yerine kayıtlı bir mesaja yönlendirip ücretlendirmeyi başlatması, arayan gerçekten cevap verene kadar sahte bir zil sesi çalarak çağrı süresini uzatması ya da çağrı bağlantı kesme mesajının arayan tarafa iletimini geciktirmesi şeklinde gerçekleştirilmektedir.

Hasılat paylaşma, bir operatörün veya üçüncü taraf hizmet sağlayıcının başka bir tarafla önceden tanımlanmış gelir paylaşım numaralarına çağrı oluşturacak bir anlaşma yapması durumudur [216]. Genellikle birden fazla dolandırıcılık türünün bir kombinasyonunu kapsar. Erişim teşvik dolandırıcılığı, kırsal alandaki bir operatörün yüksek miktarda gelen çağrıları olan bir şirketle anlaşma yaparak trafiğini artırması ve gelirinin bir kısmını şirket ile paylaşmasıyla gerçekleştirilir. Ücretsiz numara dolandırıcılığında, dolandırıcı bir operatörle anlaşarak, bu operatörü ücretsiz bir numaraya büyük miktarda çağrı başlatmak için kullanır. Ücretsiz numaralar için kullanılan ters ödeme akışı nedeniyle, operatör bu çağrı hacminden gelir elde eder ve bu geliri, çağrıları üreten dolandırıcıyla paylaşır. Uluslararası hasılat paylaşmada, dolandırıcılar genellikle bir operatörün ağına erişmek için çeşitli yasadışı

kaynaklar kullanır, ardından trafiği aşırı derecede yüksek ücretler alan telefon numaralarına yönlendirir. Özel tarifeli ya da uluslararası pahalı numaralar, genellikle Uluslararası Prim Oranı Sağlayıcıları olarak adlandırılan hileli şirketler tarafından satılır. Dolandırıcılar, bu çağrılar sonlandığında elde edilen gelirin bir kısmını alır.

Sesli dolandırıcılık; müşterilerin bildirimi, sızdırılmış veri tabanları ya da satın alma yoluyla elde edilen telefon numarası listelerine, otomatik çeviriciler vasıtasıyla istenmeyen ve gayri meşru çağrıların yapılmasını kapsamaktadır [217]. Telefonla pazarlama, bir satış elemanının müşterileri telefonla ürün veya hizmet satın almaya ikna ettiği doğrudan pazarlama yöntemidir. Sesli kimlik avında, arayan kişi meşru bir kişiyi ya da kuruluşu taklit eder ve sosyal mühendislik kullanarak özel, kişisel ve finansal bilgilere erişmeye çalışır. Teknik destek dolandırıcılığında ise, dolandırıcılar insanları bilgisayarlarına kötü amaçlı yazılım bulaştığına ikna etmeye çalışırlar ve çoğunlukla onları uzaktan erişim araçlarını yüklemek için kandırdıktan sonra sözde sorunu çözmek için ödeme talep ederler.

Telekom dolandırıcılığını önlemek veya durdurmak amacıyla; çağrı akışlarını izleme, çağrı imzalarındaki ani değişimleri tespit etme ve risk tanımlama gibi faaliyetlerden sonra; uyarı oluşturma, çağrıları engelleme, yeni çağrı rotası belirleme veya giden arama planlarını değiştirme gibi aksiyonlar alınmaktadır [218]. Fakat dijital hizmetlerin peşinden giden tüketiciler; işlemlerin artmasına sebep olarak, faturalandırma ve operasyonel destek sistemlerinin ölçeklenebilirlik sınırlarını zorlamaktadır. Bu ve benzer güçlükler doğrultusunda, telekom dolandırıcılığı sorununa çözüm üretmek amacıyla son beş yılda önerilen akademik çalışmalar Çizelge 4.1'de karşılaştırmalı olarak sunulmuştur. yüksek trafik yoğunluğundaki küçük yüzdeye sahip dolandırıcı çağrılarını düşük maliyetle çözmek için çeşitli yaklaşımların geliştirildiği görülmektedir. Çalışmalar, analizlerde kullanılan veri kümeleri açısından değerlendirildiğinde, verinin çoğunlukla mobil iletişim operatörlerine ait ve gizli olduğu görülmektedir. Büyük veri çağında telekom dolandırıcılığı tespit edilirken, karşılaşılan sorunlardan biri arama kayıtlarının meşru ya da dolandırıcı olduğuna dair etiket bilgisinin edinilmesindeki zorluktur. Dolandırıcılık genellikle ay sonunda müşteri şikâyetleri veya fatura ihtilafları ile tesadüfen fark edilir. Bu geri bildirimlerle elde edilen etiketli veri de, tüm analiz uzayında kullanmak için yeterli değildir. Etiketli veri; elde edilmesi zor, hata barındırma ihtimali yüksek ve yeni tür dolandırıcılıkları kapsayamaz olduğu için, dolandırıcılık tespitinde danışmanlı öğrenme yöntemlerinin alternatifleri öne çıkmaktadır.

Çizelge 4.1. Telekom dolandırıcılığı çözümleri

Kaynak	Amaç	Veri Kümesi	Yöntem/ Kısaltması		Sonuç			
					AUC	ACC	TPR	FPR
[219]	Graf madenciliği tabanlı hileli telefon görüşmesi algılama	Whoscall Dataset	Ağırlıklı Köprü Kaynaklı Konu Arama	HITS	82,9	*	*	*
[220]	Dolandırıcılık tespiti için ve modeli	Bir mobil iletişim operatörünün trafik verisi	Davranış Örüntüsü Tanıma	*	*	*	90	1,22
			Arama Hedefi profilleme	*	*	*	95	0,7
[221]	Abone kimlik modülü dolandırıcılığı tespiti	Bir mobil iletişim operatörünün arama detayları kayıtları	Yapay Sinir Ağı	ANN	99,7	98,7	98,7	*
			Destek Vektör Makinesi	SVM	98,5	98,87	98,7	*
[222]	Davranış modellerindeki sapmalara göre dolandırıcılık tespiti	Frıtz!Box Incident Data	İletişim Davranışı Örüntüsü Tanıma	*	*	*	98,4	0,01
[223]	Kullanıcıların arama davranışlarını analiz ederek hileli aramaları bulma	Reality Mining Dataset	Bulanık Kümeleme ve Destek Vektör Makinesi	FCM-SVM	*	94,62	88,89	5,08
				FKM-SVM	*	95,65	89,36	9,73
[224]	Davranışsal profil modelleme ile dolandırıcılık tespiti	Reality Mining Dataset	Olasılıksal Bulanık Kümeleme	PFCM	*	93,09	95,07	9,25
[225]	Kullanıcıların gizliliğinin korunarak dolandırıcılık tespiti	Bir mobil iletişim operatörünün arama detayları kayıtları	Gizli Dirichlet Tahsisi ve Maksimum Ortalama Tutarsızlık	*	98,7	*	*	*
[226]	Makine öğrenmesi modülü ve şablon algılama modülü içeren dolandırıcılık tespiti yöntemi	Bir mobil iletişim operatörünün arama detayları kayıtları	Destek Vektör Makinesi ve Sonlu Durum Makinesi	*	*	93,56	89,22	*
[227]	Mobil iletişimde dolandırıcıları tespit etme	Bir mobil iletişim operatörünün arama detayları kayıtları	Derin Evrişimli Sinir Ağı	DCNN	*	82	*	*
[228]	Yoğun ağ yapısı ve zamansal bilgiler ile dolandırıcılık davranışlarını tespit etme	Bir mobil iletişim operatörünün arama detayları kayıtları	İlişkilendirilmiş İki Taraflı Ağ	*	89,1	*	80,4	*

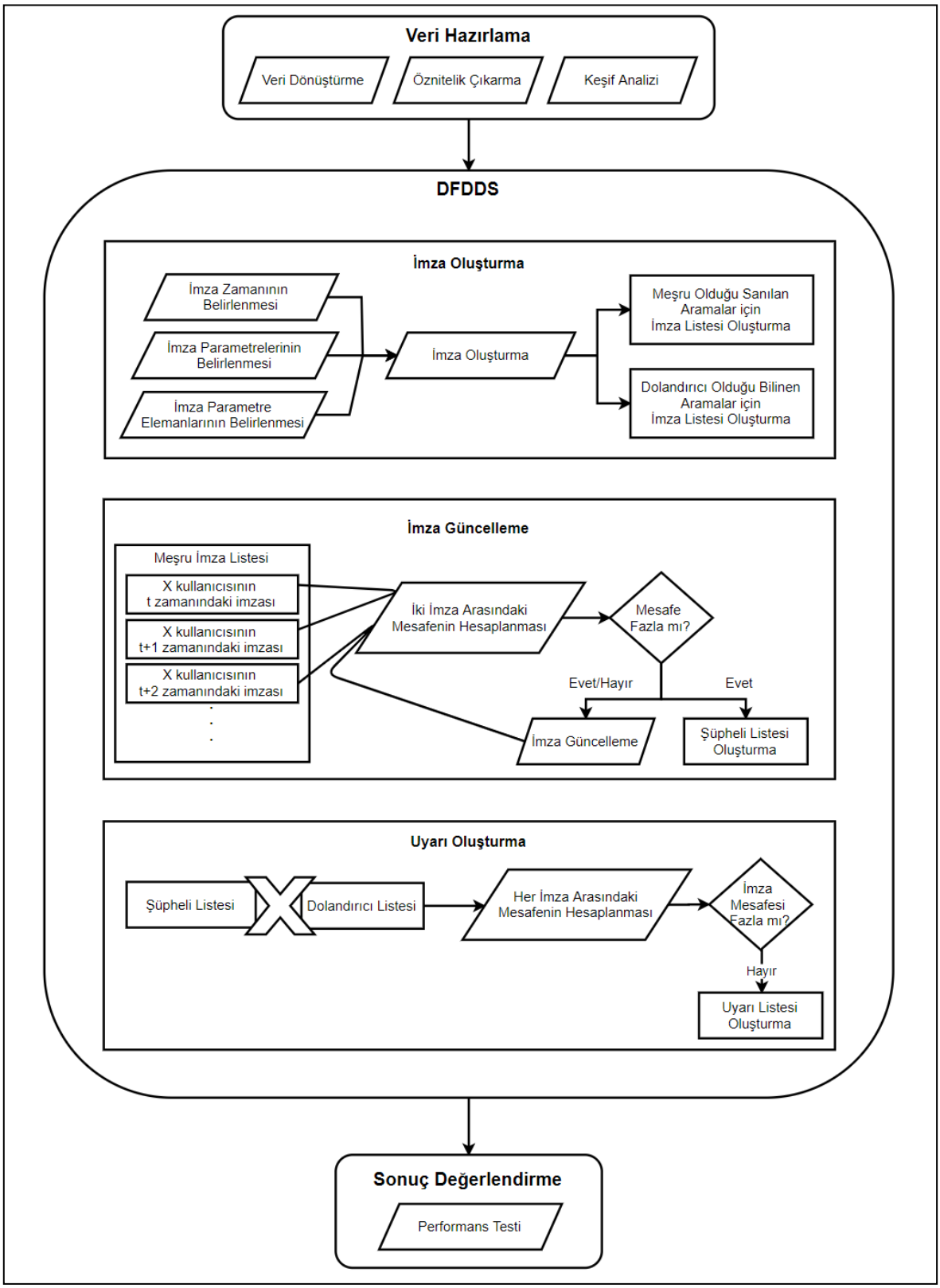
Çizelge 4.1. (devam) Telekom dolandırıcılığı çözümleri

Kaynak	Amaç	Veri Kümesi	Yöntem/ Kısaltması		Sonuç			
					AUC	ACC	TPR	FPR
[229]	Normal veri dağılımına uyan ve uymayan örnekleri ayırt etme	Bir bankanın işlem kayıtları	Üretken Çekişmeli Gürültüden Arındırılmış Otokodlayıcı	GAN-DAE	*	*	88,33	1,02
[230]	Sıralı örüntüleri yakalamak için birbirini izleyen davranışsal dizilerini modelleme	Bir mobil iletişim operatörünün arama detayları kayıtları	Tarihsel Dikkat Tabanlı Ve Etkileşimli Uzun Kısa Süreli Bellek	HAInt-LSTM	97,97	*	*	*
[231]	Gelen aramaları bulanık kümeleme ile etiketledikten sonra grup veri işleme ile dolandırıcılık tespiti	Reality Mining Dataset	Bulanık Kümeleme ve Grup Veri İşleme	GMDH	92,18	*	94,3	*
[232]	Önyargılı veri kümesinde etkili dolandırıcılık tespiti	Bir mobil iletişim operatörünün arama detayları kayıtları	Kolektif Sınıflandırıcılar	Random Forest	92,05	99,75	*	*
				AdaBoost	98,97	99,96	*	*
				XGBoost	99,26	99,97	*	*
[233]	Hızlı güncellenen dolandırıcılık tespiti	Arama konuşmaların metinleri	Geniş Öğrenme Sistemi	BLS	*	92,45	94,07	*
[234]	Anormal örüntülerin tespiti için kullanıcıların sıralı ve ilişkili davranışları aynı anda modelleme	Sentetik Veri	Dinamik Etkileşim Ağları	COSIN-Separate	69,9	*	*	*
				COSIN-Caller	76,34	*	*	*
				COSIN	91,6	*	*	*
[235]	Dengesiz veride dolandırıcılık tespiti için dinamik bir çerçeve	Bir mobil iletişim operatörünün arama detayları kayıtları	Heterojen Kolektif Öğrenme ve Destek Vektör Makinesi	Stacked-SVM	85,33	82,29	93,83	*

*:Belirtilmemiş

Dolandırıcılık tespitinde imza temelli kullanıcı profili oluşturma çözümleri, kullanıcı davranışı sapmalarını algılama yetenekleri nedeniyle büyük önem taşımaktadır. İmzalar, çeşitli çağrı parametrelerinin olasılık dağılımını içeren, zamanla değişen hesap profilleridir [236]. Genellikle; imza oluşturma, güncelleme ve uyarı oluşturma olmak üzere üç aşamada gerçekleştirilir. İmza tabanlı sistem; işlemlerin tanımlanan zaman dilimleri boyunca biriktirilip toplu olarak işlendiği zaman odaklı ya da her yeni veri örneğinin oluştuğu anda işlenen olay odaklı olmak üzere iki işleme modeli üzerinde kurgulanır. Fakat güncel ve

gelişmiş tehditlere karşı koyabilmek için bu aşamaların, hızla artan veri kümeleri üzerinde çalışabilir ve değişimlere duyarlı olarak uyarlanması gerekmektedir.



Şekil 4.3. DFDDS yönteminin metodolojisi

Etiket bağımlılığını azaltmak ve büyük veri kümelerinde ölçeklenebilir bir tespit sağlayabilmek amacıyla, Dinamik İmzalarla Dağıtık Dolandırıcılık Tespiti (Distributed Fraud Detection with Dynamic Signatures, DFDDS) adlı yeni bir çözüm yöntemi önerilmiştir. Şekil 4.3'de metodolojisi verilen DFDDS yönteminde, NETAS firmasına ait arama kayıtlarını içeren veri kümesi ön işlemden geçirildikten sonra, her kullanıcının davranış örüntüsünün bir özeti olan imzaları oluşturulmuştur. Normalden sapan davranışların yani imzalardaki ani değişikliklerin mevcudiyetinde, bu davranışlar dolandırıcı olduğu bilinen imzalar ile karşılaştırılarak dolandırıcı olma ihtimalinin yüksekliği doğrultusunda nihai karar üretilmiştir. Şüpheli aramalar, hem test verisi hem de dolandırıcı uzmanının görüşleri doğrultusunda değerlendirilmiştir.

4.1.1. Veri kümesi

Telekom sistemleri, operatörler tarafından toplanan; Arama Detay Kayıtları (Call Detail Recording, CDR), müşteri profili verileri, ağ verileri, raporlar veya anketler gibi kaynaklar olmak üzere kullanıcılar hakkındaki bilgilerle ilgili olan dört tür büyük veri türünden oluşmaktadır [237]. CDR; bireysel kullanıcı profilleme veya benzer grup analizi ile alışılmadık örüntülerin tespiti, uyum takibi, riskli hedeflerin tanımlanması ve fatura bilgilerinin doğrulanması gibi durumlar için sıklıkla tercih edilen veri kaynağıdır [238]. CDR; arayan ve aranan numaralar, tarih, saat, arama süresi ve bazen de konum gibi temel arama bilgilerini içerir.

Geleneksel telefon sisteminden internet üzerinden ses protokolü yani VoIP teknolojisine geçiş; çok düşük maliyetli veya ücretsiz görüşmeler yapılabilmeyi, iletişimde hareket kabiliyetini artırmayı ve değişen ihtiyaçlara esnek uyum yeteneği sağlamıştır [239]. Bunların yanında kullanımın yaygınlaşmasıyla; hizmet kalitesi, güvenlik ve yanlış kullanım sorunları da ortaya çıkmıştır. Ortaya çıkan sorunlardan bir diğeri de, ücretli iletişim hizmetlerine yetkisiz yöntemler kullanarak erişme ve hizmet kullanımı için son kullanıcıyı veya servis sağlayıcıyı borçlandırma faaliyetleri olan VoIP dolandırıcılığıdır [240]. VoIP dolandırıcılığı tespiti, farklı teknolojilere bağlı olduğu, güçlü yerleşik güvenlik mekanizmalarından yoksun olduğu ve sisteme yeni geliştirilmiş saldırıların entegrasyonu kolay olduğu için diğer telekom ağlarından daha zordur. Bu sistemler; çağrı dolandırıcılığı ve ücret dolandırıcılığına ek olarak hizmet reddi saldırıları, fidye yazılımları, konuşmaların gizlice dinlemesi ve hatta konuşma sırasında yapılan klavye vuruşlarının takip edilmesi gibi tehditlerin altındadır.

Tez kapsamında geliştirilen DFDDS yönteminde kullanılan veri kümesi, NETAŞ firması tarafından sağlanmıştır ve VoIP hizmeti kullanılırken üretilen CDR kayıtlarından oluşmaktadır. NETAŞ, bilgi ve iletişim teknolojileri alanında yenilikçi sistem entegrasyonu ve teknoloji hizmetleri sunmaktadır ve dünyadaki ilk on küresel VoIP ve multimedya laboratuvarlarından birine sahiptir [241]. Tez kapsamında geliştirilen yöntemi, NETAŞ'a ait bir ürün olan NOVA FMS dolandırıcılık yönetim sistemine entegre olarak çalışması hedeflenmiştir.

Çizelge 4.2. NETAŞ veri kümesine ait özet bilgiler

	Meşru Kayıt	Dolandırıcı Kayıt	Toplam Kayıt	Toplam Öznitelik
Arama Sayısı	1582955	72505	1655460	45
Tekil Arayan Sayısı	407476	24	407500	45

2018 yılına ait olan veri kümesi NETAŞ tarafından anonimleştirilmiş olup, Çizelge 4.2'de verildiği üzere 24'ü dolandırıcı olarak bilinen 407500 tekil kullanıcının yapmış olduğu toplam 1655460 arama kaydından oluşmaktadır. Veri kümesine ait IR değeri 21,83'dür. Veri kümesinde dolandırıcı olduğu bilinen kullanıcıların sayısı her ne kadar 24 olsa da, geri kalan kullanıcıların meşru olduğu kesin değildir. Mevcut etiketleme NETAŞ tarafından, dolandırıcı davranışı gösteren bir kullanıcının her araması dolandırıcıdır varsayımına dayanarak gerçekleştirilmiştir. Her kaydı tanımlayan 45 öznitelik bulunmaktadır. Gizlilik sözleşmesi gereğince bu öznitelikler detaylı olarak verilememektedir.

4.1.2. İmza oluşturma

İmza, değerleri belirli bir süre içinde değişen öznitelik değişkenleri vektörünü ifade eder. S bir imza olarak değerlendirilirse, t zaman birimi için f fonksiyonu ile Eş. 4.1'de verildiği şekilde oluşturulmaktadır [242].

$$S = f(t) \tag{4.1}$$

Zaman birimi saat, gün, hafta veya ay olabilir. İmza oluşturma aralığı, dolandırıcılığın hızlı tespiti için kritik bir konu olsa da, bu aralık hizmet kullanım yoğunluğuna göre seçilmektedir. Çağrıların sık olmadığı durumlarda, imza parametrelerinin dağılımı kararsız olmaktadır. Başka bir deyişle, arama yapılmayan aralıklar olasılık dağılımını olumsuz

etkileyen ani değişimler gibi görülerek anomali olarak tanımlanmaktadırlar. NETAŞ veri kümesi çok sık çağrılar içermediğinden tez kapsamında geliştirilen yöntemde imza zaman birimi, gün olarak belirlenmiştir.

İmza oluşturmada, zaman biriminden sonra belirlenmesi gereken diğer husus ise imza parametrelerinin seçimidir. Bu parametreler, zaman birimi ve verideki öznitelikler doğrultusunda çeşitlenmektedir. Parametreler; çağrı süresi, son aranan ülkeler, son fatura bedeli, haftanın günlerindeki veya günün saatlerindeki çağrı ve mesaj sayısı ya da ortalaması olarak seçilebilmektedir [243]. Bu yöntemde oluşturulan imzalarda aşağıda verilen altı parametre kullanılmıştır.

1. Gün bölümü: şafak, sabah, öğle arası, öğleden sonra, akşam, gece
2. Çalışma saati: mesai içi, mesai dışı
3. Aramanın ücret kategorisi: çok ucuz, çok ucuz gsm, ucuz, ucuz gsm, orta, orta gsm, pahalı, pahalı gsm, çok pahalı, çok pahalı gsm
4. Aramanın sonlanma şekli: kesme, iptal, hata, sorunsuz
5. Dakika cinsinden süre aralıkları: 0, 3, 5, 10, 15, 20, 25, 30, 60, 90, 120, 240, diğer
6. Hata kodları: 0, 200, 503, 486, 404, 484, 408, 480, 487, 504, diğer

Bu altı parametrenin elemanlarının, çağrılarda var olma olasılığı doğrultusunda arayana dair ilk imza oluşturulduktan sonra, aynı şekilde oluşturulan gelecek aramalara ait imzaların bir önceki ile karşılaştırıldığı güncelleme aşamasına geçilir.

4.1.3. İmza güncelleme

Belirli bir süre sonra imzayı aniden değiştirmek yerine, imzayı aşamalı olarak güncellemek sadece imzayı dolandırıcılığa karşı güncel tutmakla kalmaz, aynı zamanda hesaplama karmaşıklığı açısından da etkilidir. İmza güncelleme sürecinin gerçekleştirilebilmesi için de, imzalar arasındaki benzerliğin ya da farklılığın yani mesafenin belirlenmesi gerekmektedir.

Literatürde imzalar arasındaki benzerliği ölçmek için; histogramdan elde edilen ortalama ağırlıklı Kullback-Leibler mesafesi [244], ortalamaya olan mesafeyi baz alan z-skor fonksiyonu [243] ve z-skorun iyileştirilmesi ile geliştirilen Poisson fonksiyonu [245] gibi metrikler kullanılmaktadır.

Geliştirilen yöntemde, olasılıksal dağılımlarda daha iyi sonuçlar veren Hellinger Mesafesi [246], birbirini izleyen iki güne ait imzalar arasındaki benzerliği ölçmek için kullanılmıştır. İki olasılık dağılımı tamamen farklıysa Hellinger Mesafesi bir değerini almaktadır. Küçük mesafe ise güçlü benzerliğe işaret etmektedir. $S: s_1, s_2, \dots s_n$ ve $C: c_1, c_2, \dots c_n$ olarak tanımlanmış n değere sahip iki dağılım arasındaki Hellinger Mesafesi Eş. 4.2'de verildiği şekilde hesaplanmaktadır.

$$H^2(S, C) = \frac{1}{2} \sum_{i=1}^{n} \left(\sqrt{s_i} - \sqrt{c_i}\right)^2 \quad (4.2)$$

Yakınlığı değerlendirmek için, bir benzerlik eşik değeri olan θ tanımlanmıştır. İmzadaki altı parametreden herhangi biri için maksimum mesafe θ'dan düşükse, imza güncellenir. Güncelleme yapılırken, yeni eylemin etkisinin, imza değerlerinde ağırlıklandırılması gerekmektedir [245]. S_t t zamanındaki imza, β yeni eylemin ağırlığı, P_C ise t zamanında işlenmiş c CDR kayıtları olarak tanımlandığında, imza Eş. 4.3'de verildiği gibi güncellenmektedir.

$$S_{t+1} = \beta \,.\, S_t + (1 - \beta) \,.\, P_C \quad (4.3)$$

Bir kullanıcı aşırı derecede farklı davrandığında, yani parametrelerden herhangi birinin maksimum mesafesi θ değerinden büyük olduğunda, imza hem uyarı listesine eklenir hem de güncellenir. Böylece, imza üzerinde yenilenme etkisine neden olan bu doğal veya anormal değişiklik yeni örüntülerin adaptasyonunun sağlanmasına yol açmaktadır.

Dolandırıcı olduğu bilinen kullanıcılarda ise, davranışların en basit haliyle izlenmesi hedeflenmektedir. Bu nedenle, dolandırıcılık imzalarından güncelleme yapılmaksızın dolandırıcılık listesi oluşturulmaktadır.

4.1.4. Uyarı oluşturma

Gerçek dünya uygulamalarında dolandırıcılık tespit sistemleri bol miktarda uyarı üretmektedir. Ne yazık ki çoğu da yanlış pozitiftir, yani yanlışlıkla dolandırıcılığı gösteren ve dolandırıcılık uzmanının daha fazla araştırma yapmasını gerektiren uyarılardır. Zorluk,

yanlış pozitiflerin sayısını azaltmak ve uyarıların kalitesini artırmaktır, çünkü analist bilgisini genellemek zordur ve dolandırıcılık örüntüleri oldukça dinamiktir [247].

Uzmanın görüşüne sunulacak uyarılar, üretilen uyarı listesindeki her imzanın dolandırıcılık listesindeki her imza ile olan benzerliği doğrultusunda belirlenmektedir. Gereksiz karşılaştırmalarından kaçınmak için, dolandırıcılık listesindeki benzer imzalar Hellinger Mesafesi kullanılarak elenmektedir. İki dolandırıcılık imzası arasındaki mesafe θ'dan az olduğunda, bunlardan biri dolandırıcılık listesinden kaldırılmaktadır. Bu şekilde, farklı davranışlara sahip benzersiz bir dolandırıcılık listesinin oluşturulması amaçlanmaktadır. Bu noktada, dolandırıcılık listesinin kalitesi ve kapsamı yeni ve farklı dolandırıcılıkların keşfedilmesinde büyük önem taşımaktadır.

Dolandırıcılık dışı sebepler yüzünden de kullanıcı davranışlarında sapmalar meydana gelebilmektedir. Dolandırıcılık listesini bu tür durumlardan arındırabilmek için benzerlik oranı geliştirilmiştir. Her imza parametresi için atanmış olan ağırlık değeri α olarak tanımlanmaktadır ve dolandırıcılık olasılığını arttırdığı düşünülen parametrelerin α değeri daha büyük seçilebilir. Her imza parametresindeki elemanların olasılıksal dağılımı *x* olmak üzere, benzerlik oranı *R*, Eş 4.4'de verildiği şekilde hesaplanmaktadır.

$$R = \sum_{i=1}^{n} \alpha_i * x_i \begin{cases} 1 & if\ x < \theta \\ 0 & diğer \end{cases} \tag{4.4}$$

İmzaların herhangi bir parametresi θ'ye ne kadar yakın olursa, uyarı dolandırıcılığa o kadar benzerdir. Uyarının dolandırıcılığa yüksek oranda benzediği durumlarda, bu aramadan ve benzerlik oranından oluşan bir şüpheli listesi üretilmektedir. Her bir uyarının her bir dolandırıcılıkla karşılaştırılması aynı uyarıların tekrar tekrar bildirilmesine neden olabilir. Bu sebeple, şüpheli listesindeki tekrarlar atılarak benzersiz şüpheli listesi elde edilmektedir. Yöntemin çıktısı olan bu şüpheli listesi, ileri değerlendirme ve inceleme için dolandırıcılık uzmanın görüşüne sunulmaktadır.

4.1.5. MapReduce şeması

DFDDS'yi büyük veri analitiğine uygun olarak gerçekleştirmek için bir MapReduce şeması geliştirilmiştir. Sözde kodu Şekil 4.4'de verildiği üzere, süreç temelde Map ve Reduce

fonksiyonları vasıtasıyla tamamlanmaktadır. Map fonksiyonu, bir satırın bir kayıt içerdiği giriş dosyasının her bir kaydına uygulanırken, Reduce fonksiyonu ise, Map fonksiyonuna ait çıktıları alıp; imza oluşturma, imza güncelleme, benzer imzaları tespit etme ve şüpheli imzaları yakalama adımlarından oluşan tüm işlemleri gerçekleştirerek nihai sonuçları toplar.

Algoritma 1: DFDDS için MapReduce Algoritması

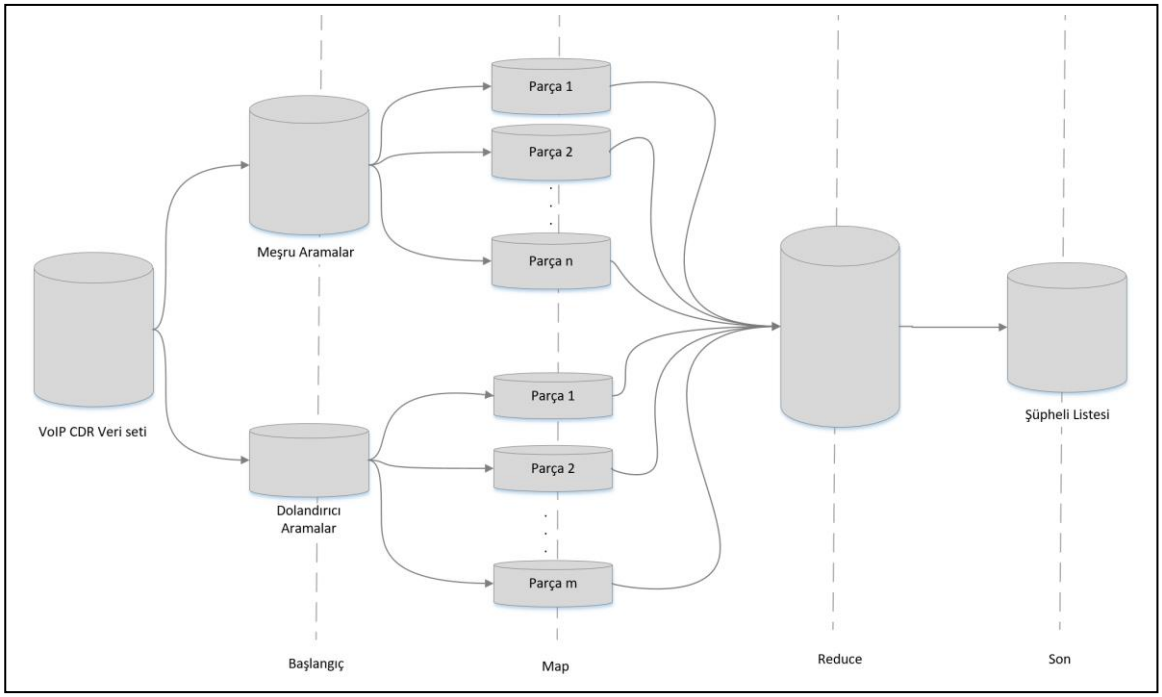
```
Function MAP(anahtar,değer):
    veri⟵ VeriGetir(değer)
    kayıt⟵ veri.KayıtSırala(veri.Kullanıcı, veri.Tarih)
    sınıf⟵ veri.Sınıf
    if sınıf==Meşru then
        meşru⟵ İmzaOlustur(kayıt.Özellikler[], İmzaParametreleri[])
        güncel⟵ mesru.İmzaGüncelle(BenzerlikEsikDeğeri, YeniEylemAğırlığı)
        imzalar.Ekle(güncel)
    else
        dolandırıcı⟵ İmzaOlustur(kayıt.Özellikler[], İmzaParametreleri[])
        benzersiz⟵ dolandırıcı.BenzerDolandırıcılarıÇıkar()
        dolandırıcılar.Ekle(benzersiz)
    end
    Yayınla(imzalar, dolandırıcılar)
end
Function REDUCE(imzalar, dolandırıcılar):
    uyarılar⟵ imzalar.İmzaKarsılastır(BenzerlikOranı, ParametreAğırlıkları)
    şüpheliler⟵ BenzerlikBul(uyarılar, dolandırıcılar)
    son⟵ AynıŞüphelileriÇıkar(süpheliler)
    Yayınla(null,son)
end
```

Şekil 4.4. DFDDS için MapReduce sözde kodu

Bu süreç Şekil 4.5'de verildiği gibi; Başlangıç, Map, Reduce ve final olmak üzere dört aşama çerçevesinde sürdürülmüştür.

1. Başlangıç: Veri dağıtık dosya sistemine yerleştirilir. Dolandırıcı çağrılar ile meşru çağrılar birbirinden ayırılır.
2. Map: Kayıtlar, arayana ve zamana göre sıralanır. Her parçada bir ya da birkaç arayana ait çağrılar bulunacak şekilde aramalar haritalandırılır. Tüm arayanlar için imzalar oluşturulur. Meşru imzalar için, her imza bir sonraki imzayla karşılaştırılır ve son imza güncellenir. Son imza bir önceki imzadan oldukça farklıysa, bir uyarı oluşturulur ve uyarı listesine eklenir. Dolandırıcı imzaları için ise, benzer imzalar ortadan kaldırılarak benzersiz dolandırıcılık listesi oluşturulur.

3. Reduce: Dolandırıcı listesinin uyarı listesi ile Kartezyen çarpımı yapılır ve benzerlik oranları elde edilir. Her parçadaki benzerlik oranı yüksek şüpheli aramalar birleştirilerek, şüpheli listesi elde edilir.
4. Final: Üretilen benzersiz şüpheli listesi, dolandırıcı uzmanına sunulur.

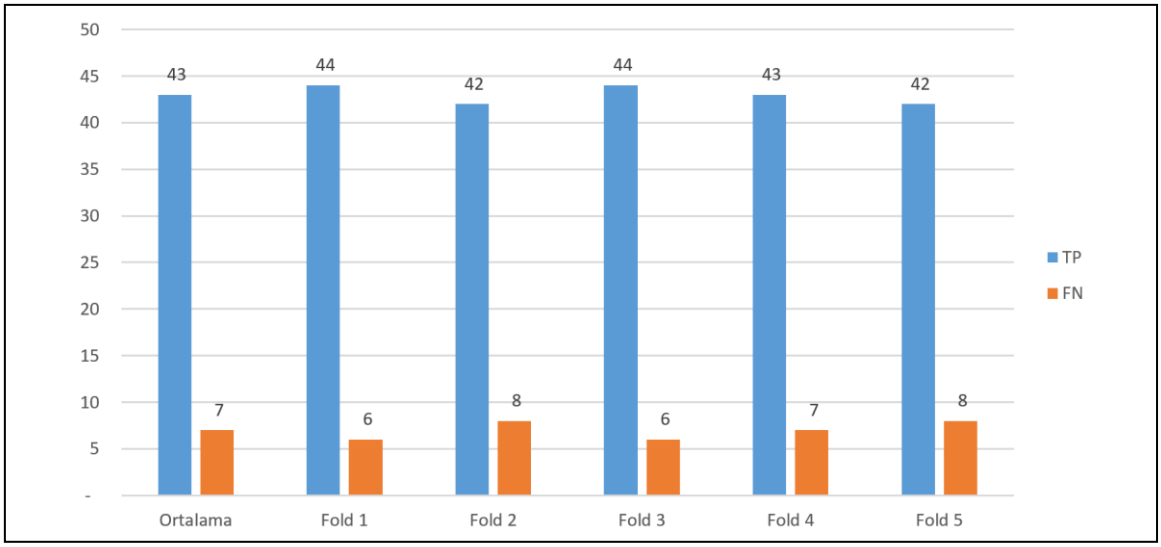


Şekil 4.5. DFDDS yönteminin MapReduce akış şeması

4.1.6. Deneysel sonuçlar

DFDDS'nin test sürecinde, dolandırıcı olduğu bilinen aramaların tespiti birincil öncelik olduğu için değerlendirme yapılırken recall metriği esas alınmıştır. Dolandırıcı etiketinin azlığı sebebiyle, dolandırıcı olduğu bilinen imzaların 50 tanesi ayrıştırılarak test amacıyla kullanılmıştır. Elde edilen sonuçların eğitim kümesi bağımlılığını azaltmak için beş katlı çapraz doğrulama testi gerçekleştirilmiştir. Her katta test için rastgele 50 imza seçildikten sonra geri kalan imzalara eğitim amacıyla kullanılmıştır ve değerlendirme bu katların ortalamasına göre yapılmıştır. Yöntem, bu aramalardan ortalama 43 tanesini dolandırıcı olarak tespit ederek %86 başarı elde etmiştir.

Yöntem dolandırıcı tespitinin yanında ek çıktılara da sahiptir. Bu bağlamda, hedefler doğrultusunda üretilen sonuçlar aşağıda verildiği üzere dört şekilde değerlendirilebilir.



Şekil 4.6. NETAŞ veri kümesindeki beş katlı doğrulama testi sonuçları

1. Uyarıların tespiti: Kullanıcının iki günlük imzası arasında bir fark olduğunda, bir davranış değişikliği tespit edilir ve bu değişiklik, dolandırıcı bir davranışa benziyorsa bir uyarı oluşturulur.
2. Büyük oranda hileli davranışlara benzeyen şüpheli davranışların tespiti: Daha önce tespit edilmeyen ve dolandırıcı davranışlarda bulunma olasılığı yüksek olan kullanıcılar tespit edilir.
3. Kısmen hileli davranışlara benzeyen şüpheli davranışların tespit edilmesi: Tanımlanmış dolandırıcılıklara çok az benzerlik göstermesine rağmen, önceki davranışlarından sapan kullanıcılar şüpheli olarak tespit edilir. Bu şekilde, muhtemelen yeni bir dolandırıcılık davranışına sahip olan kullanıcılar tanımlanır.
4. Birçok hileli davranışa benzeyen şüphelilerin tespiti: Şüpheli bir kullanıcının birden fazla hileli davranışa benzerliği belirlenir. Böylece farklı davranışları barındıran şüpheliler tespit edilir.

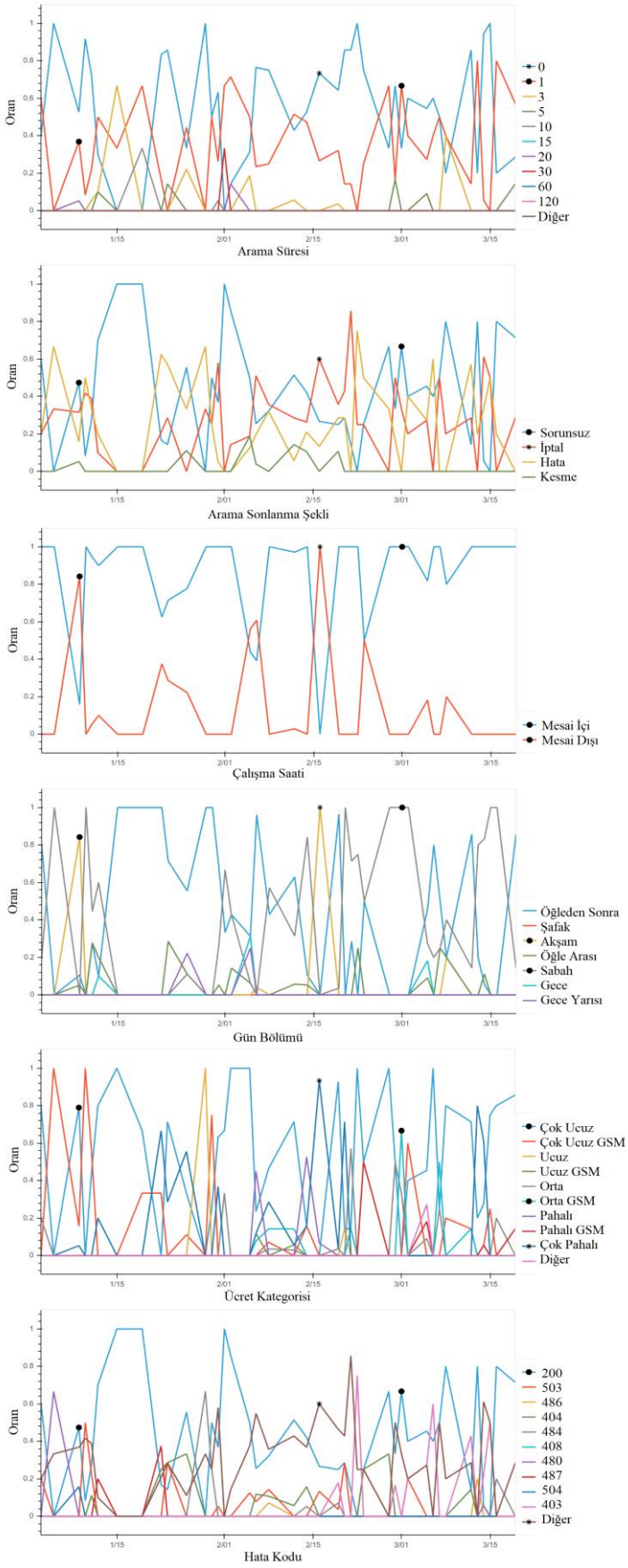
Bir kullanıcıya ait imzaların zaman içerisindeki değişiminin verildiği Şekil 4.7'deki bir örnek özelinde, DFDDS'nin sonuçları görselleştirilmiştir. Şekildeki her bir çizgi grafiği imzanın bir parametresini, her bir renkli çizgi imza parametresine ait elemanları ve çizge grafiğinin yatay ekseni de tarihi ifade etmektedir. Bu kullanıcı, hizmet kullanımı süresi boyunca 1/9, 2/16 ve 3/01 tarihlerinde olmak üzere üç uyarı üretmiştir. Nokta ile işaretlenen uyarılar dolandırıcılık benzerlik oranları düşük olduğu için göz ardı edilmiştir. Bu uyarılar, daha önce kullanıcı örüntüsünde bulunmayan fakat ani bir değişiklik de sayılamayacak

davranış değişikliklerinden kaynaklanmıştır. Yıldız ile işaretlenen uyarının ise, bilinen bir dolandırıcılığa büyük oranda benzediği tespit edildiği için şüpheli olarak değerlendirilmiştir. Bu şüpheli imzaya ait parametre değişimlerine detaylı olarak bakıldığında, şüphe yaratan aramanın; iş saati dışında, akşam, pahalı bir hatta doğru yapılmış ve iptal ile sonlanmış olmasından kaynaklandığı görülmektedir.

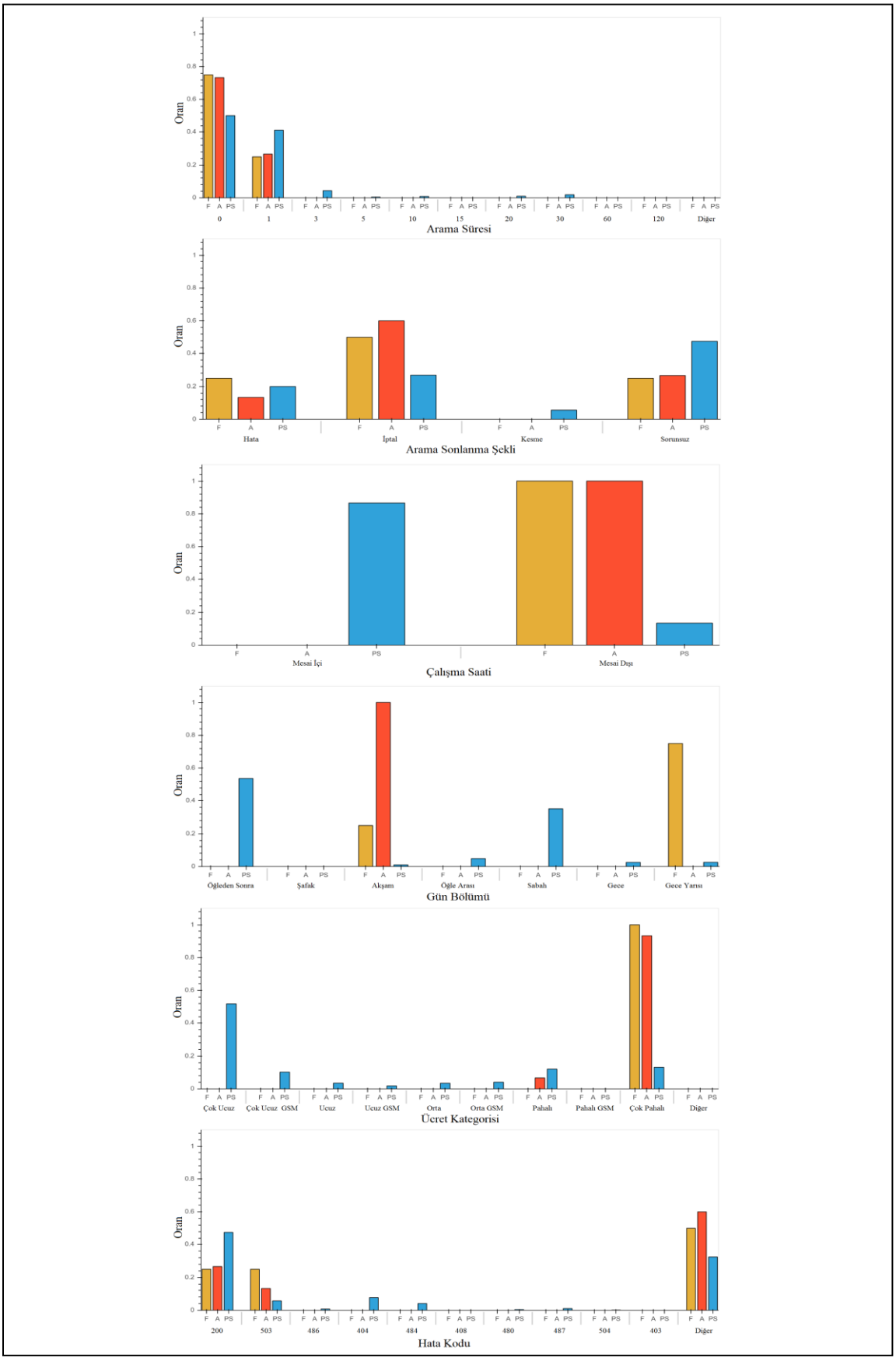
Şekil 4.7'de yıldız ile işaretlenen uyarının, Şekil 4.8'de verilen bilinen bir dolandırıcılık türüne %80 ve Şekil 4.9'da verilen bilinen bir diğer dolandırıcılık türüne ise %64 oranında benzediği tespit edilmiştir. Şekillerdeki her bir çubuk grafiği uyarı, uyarı öncesi ve uyarının benzediği dolandırıcı imzalarının bir parametresini, çubuk diyagramın yatay ekseni imza parametrelerinin elemanlarını ve çubuk diyagramın dikey ekseni ise bu elemanların bu imzada bulunma yüzdesini ifade etmektedir. Şekil 4.8'de verilen büyük orandaki benzerlik; arama süresinden, arama sonlanma biçiminden, aramanın mesai saatleri dışında olmasından ve arama kategorisinden kaynaklanmıştır. Şekil 4.9'da verilen benzerlik ise; arama süresinden, arama sonlanma biçiminden ve aramanın mesai saatleri dışında olmasından kaynaklanmıştır.

DFDDS yönteminde, birçok parametre ve metrik probleme özgü olarak değiştirilebilir ve optimize edilebilir. Çağrı merkezi gibi yoğun çağrılara sahip bir model için imza oluşturma zaman birimi daha küçük seçilebilir ve nispeten daha az çağrı yapılan modeller için daha büyük birimler seçilebilir. Farklı zaman birimi, hem hızlı aksiyon alınması hem de davranış takibi ile ilişkilidir. Gün içindeki saatlerin, haftanın günlerinin veya ayın günlerinin etkilerinin takip edilmesi istendiğinde, imza parametreleri seçilen aralıklara bağlı olarak artabilir veya azalabilir. İmza güncellemesi için kullanılan benzerlik eşiğinin küçük tanımlanması uyarı sayısının artmasına neden olabilecekken, tüm değişikliklerin takibini sağlayabilir. Son olarak, benzerlik oranının hesaplanması sırasında her bir parametrenin önemi yönünde ağırlıklandırılması, uzman tarafından belirtilen bir parametreden kaynaklanan daha ciddi uyarıların üretilmesini sağlayabilir.

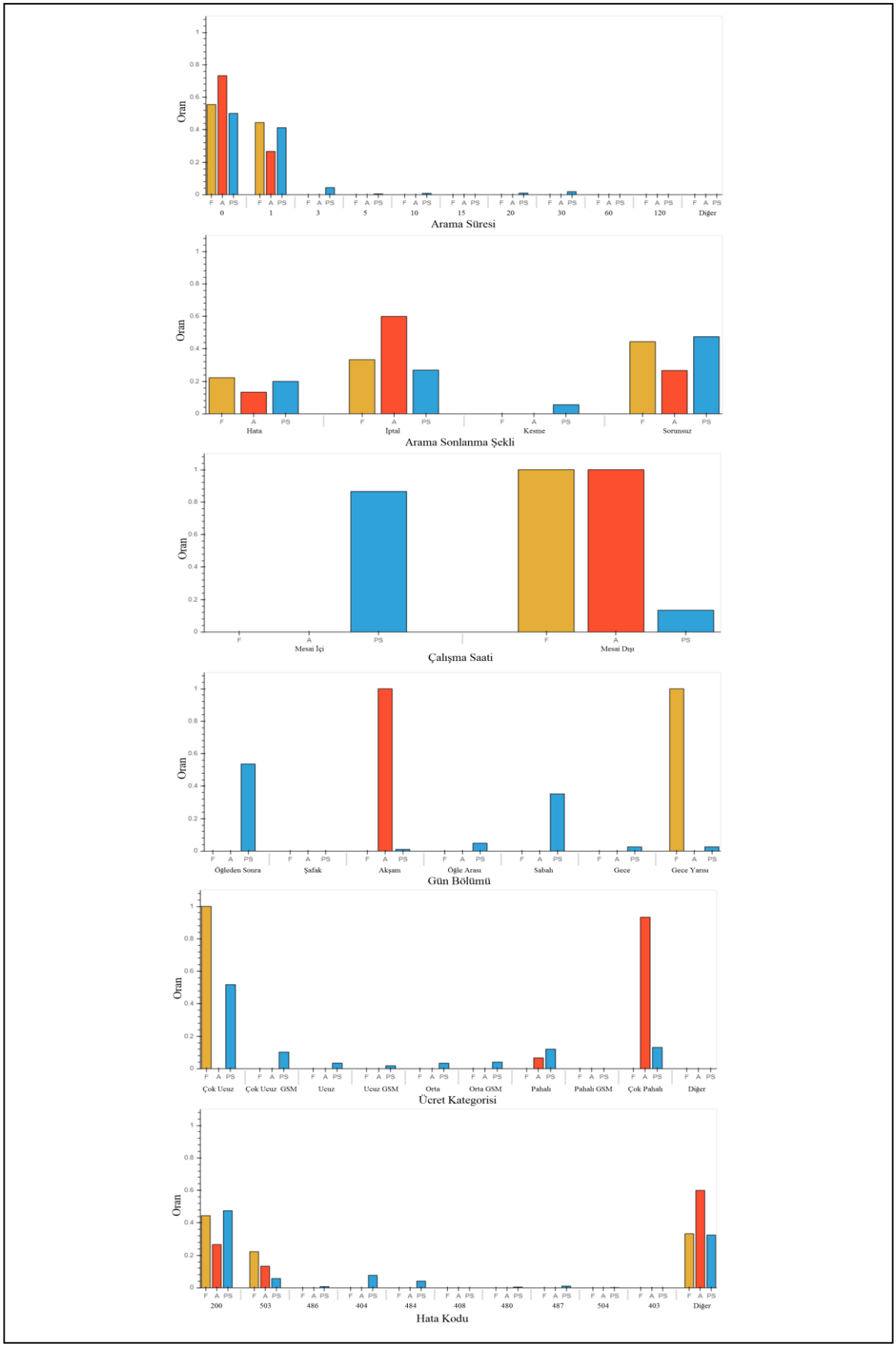
DFDDS, gerçek trafik verileri üzerine uygulandığı için, veri kümesindeki her kullanıcının günlük düzenli aramaların bulunmaması, kimi kullanıcıların günlük bir iki aramasının bulunması ya da dolandırıcıların kısa süreli servis kullanması gibi nedenler davranış örüntülerinin çıkarılmasını zorlaştırmakta ve olasılık dağılımında sıkça değişiklik oluşmasına sebep olmaktadır.



Şekil 4.7. Bir kullanıcıya ait imzaların zaman içerisindeki değişimi



Şekil 4.8. Bir kullanıcıya ait bir imzadan (S) sonra oluşan uyarı imzasının (A) %80 benzediği dolandırıcılık imzası (F)

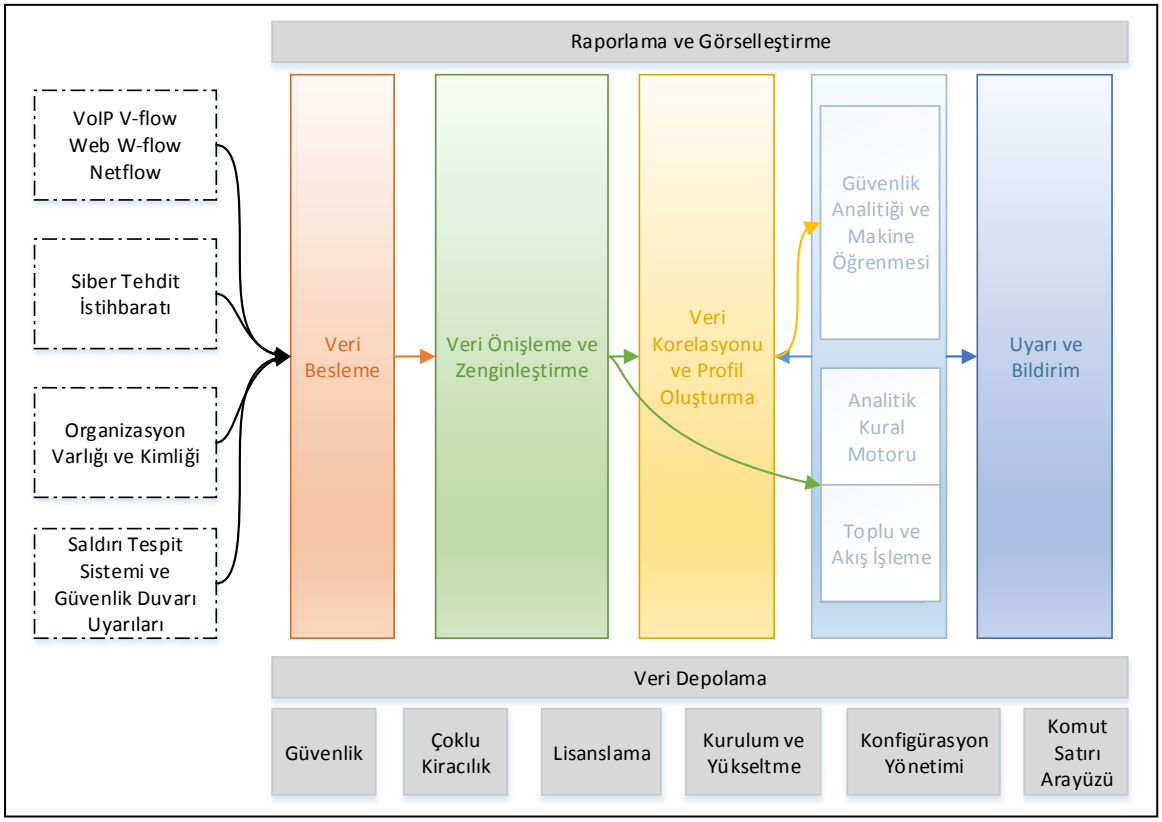


Şekil 4.9. Bir kullanıcıya ait bir imzadan (S) sonra oluşan uyarı imzasının (A) %64 benzediği dolandırıcılık imzası (F)

Tüm bu durumlar, modelin çok sayıda uyarı üretmesine yol açmaktadır. Bu uyarılar son kontrol için uzman görüşüne sunulacağından mümkün olduğunca rafine edilmelidir. Dolandırıcılık listesi bu durumu en aza indirmek için kullanılmıştır. Dolandırıcılık listesi ve uyarı listesi ile elde edilen şüpheli listesi hem uyarı sayısını azaltmakta hem de şüpheli aramanın ciddiyetini ortaya çıkarmaktadır.

4.1.7. Entegrasyon

Elde edilen sonuçlar doğrultusunda NETAŞ, DFDDS yöntemini Şekil 4.10'da anahtar özellikleri verilen NOVA FMS (Fraud Management System) [241] olarak adlandırılan büyük veri analitiği platformunun, sıra dışı kalıpları algılayan Analitik Kural Motoru birimine entegre etmiştir.



Şekil 4.10. DFDDS yönteminin NOVA FMS platformuna entegrasyonu [241]

NOVA FMS, istatistiksel ve makine öğrenimi yaklaşımlarını kullanan derin ve bütünsel bir veri güvenliği analiz platformudur. Gelişmiş ağ izleme sonucunda hem yakın gerçek zamanlı hem de zamanla birikmiş olan yüksek miktardaki veriyi analiz ederek; karmaşık

anormallikleri, siber saldırıları, siber tehditleri ve siber dolandırıcılıkları tespit etmeyi amaçlar. NOVA FMS, bilinen saldırı desenlerini ve anormallikleri kural tabanlı olarak tespit ederken, normal kullanıcı davranışlarını öğrenerek her kullanıcının hesabındaki ve çağrı kullanımındaki değişiklikleri ve olağan dışı faaliyetleri gelişmiş makine öğrenimi kullanarak tespit eder. Kullanıcı profili oluşturma, davranış analizi ve tehdit modelleme yöntemleri daha iyi sonuçlar elde etmek için bu birlikte kullanılır. Son olarak, yanlış pozitifleri azaltmak ve soruşturmayı hızlandırmak için kullanıcılara risklerine göre skorlar atanır ve dolandırıcılık uzmanlarına potansiyel tehditler bildirilir.

4.2. Dolandırıcılık Tespiti için Dağıtılmış Küme Tabanlı Yeniden Örnekleme

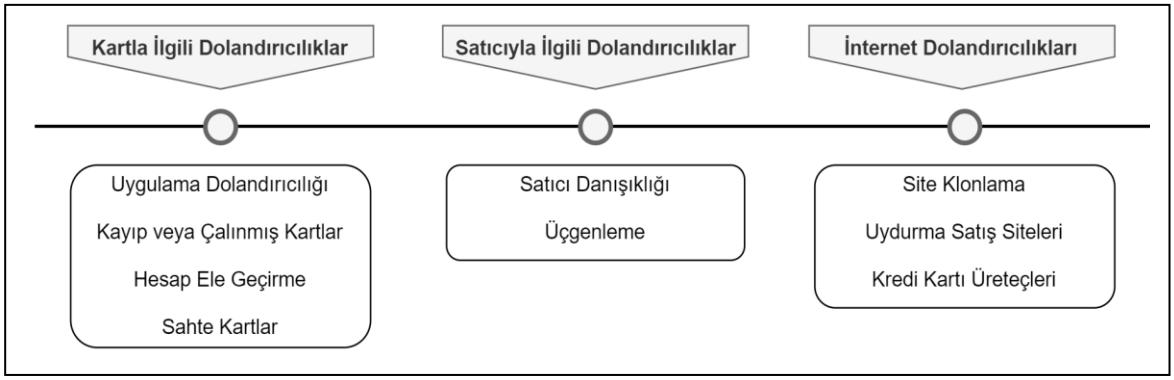
Ödeme endüstrisindeki teknolojik ve finansal yenilikler, daha gelişmiş ödeme yöntemleri sunarak tüketicilerin ödeme aracı seçimlerini etkilemektedir. Ödeme kartları, yüz yüze ve çevrimiçi satın alımlar için öncül ödeme yöntemi haline gelmektedir. Tüketicilerin harcama davranışları geliştikçe, ödeme kartları için küresel pazar da büyük ölçüde büyümüştür. Küresel ölçüde; genel amaçlı ve özel markalı kredi kartları, banka kartları ve ön ödemeli kartlar, 2019 yılında mal ve hizmet alımlarında ve nakit avans ve para çekme işlemlerinde 43,916 trilyon dolar işlem hacmi üretmiştir [248]. Ödeme yöntemlerindeki yenilikler, daha fazla tüketici rahatlığı ve verimliliği sağlarken; yeni ürünlerin, sağlayıcıların ve teknolojilerin oluşturduğu karmaşıklık, yeni risk faktörlerinin de ortaya çıkmasına neden olmuştur. Genellikle bankacılık işlemlerin %0,1'den daha azının hileli olduğu bilinmesine rağmen [249], maddi kayıp bu oranın oldukça üzerindedir. Bir şekilde kartı ele geçirmek suretiyle çevrim dışı olarak ya da kart mevcut olmaksızın çevrimiçi olarak gerçekleştirilen dolandırıcılıklar, 2019 yılında küresel çapta 30,07 milyar dolar kayba sebep olmuştur ve bu rakamın 2023'de 35,67 milyar dolara ulaşacağı tahmin edilmektedir [250].

1950'li yıllarda kullanılmaya başlamasına rağmen 1980'li yıllarda yaygınlaşan ödeme kartları, günümüze ulaşana dek değişen dinamiklere sahip dolandırıcılıklara maruz kalırken bu süreç içerisinde alınan güvenlik önlemleri de Çizelge 4.3'de verildiği üzere evrimsel olarak gelişmiştir [251], [252]. Bu süreç, bireyler tarafından gerçekleştirilen ve hafif önlemlerle önüne geçilebilecek olan fırsatçılığa dayalı eylemlerin, küresel suç şebekeleri tarafından gerçekleştirilen ve derin veri analitiğine dayalı tekniklerle çözülebilecek karmaşık yöntemlere dönüşmesine evrilmiştir.

Çizelge 4.3. Kart dolandırıcılığının ve güvenliğinin evrimi

	1980	1990	2000	2010	2020
Dolandırıcılar	Bireyler	Takımlar	Yerel Suç Şebekeleri	Küresel Suç Şebekeleri	Merkezi Olmayan Organizasyonla Küresel Suç Şebekeleri
Hedef	Tüketiciler	Küçük ve Orta Büyüklükteki İşletmeler	Büyük İşletmeler	Amir Bankalar	Çevrimiçi Ödeme/Satış Servis Sağlayıcıları
Gerekli Kaynaklar	Fırsatçılık	Temel Bilgi	Teknik Bilgi	Teknik Uzmanlık, İçeriden Bilgi, Küresel Bağlantılar	Dijital Kimlikler, Çoklu Temas Noktasından Tüketici Verileri
Önde Gelen Dolandırıcılık Türleri	Kayıp, Çalıntı	Yerel Kalpazanlık	Kimlik Hırsızlığı, Oltalama	Fiziki Kart Olmadan Yapılan Dolandırıcılık, 3D Doğrulama Dolandırıcılığı, Bankamatik Dolandırıcılığı	Hibrit Dolandırıcılıklar
Hedeflenen Kart Türleri	Seyahat/Eğlence Kartları	İkramiye Kredi Kartları	Kitle Pazar Kredi Kartları	Kredi Kartları, Banka Kartları, Ön Ödemeli Kartlar	Banka Hesapları, Mobil Bankacılık Dijital Kimlikleri
Ödeme Güvenliği	İmza paneli, Kart Kabartma, Mikro Baskı, Elektronik Yetkilendirme	Kart Doğrulama Değeri, Risk Puanlama, Çip ve Kişisel Kimlik Numarası Spesifikasyonları	Gelişmiş Yetkilendirme, Yarı Gerçek Zamanlı Uyarılar, Kişisel Kimlik Numarası Şifreleme	Risk Puanlama İyileştirmeleri, Belirteç Hizmeti, Kimlik Doğrulama, Biyometrik Spesifikasyonlar	Tehdit İstihbaratı, Güçlü Kimlik Doğrulama, Derin Veri Analitiği

Kredi kartı; kart sahibinin adı, kart numarası, hesap numarası, son kullanma tarihi, kuruluş logosu, hologram, gömülü çip gibi unsurları barındıran ve bir banka veya bankacılık dışı finans şirketi tarafından müşterilerine kredi vermeye hazır bir plastik karttır [253]. Nakit ödemeye güçlü bir alternatif olan kredi kartının, kart sahiplerinin ve kartı sağlayan kuruluşun haberi olmaksızın dolandırıcılar tarafından kişisel çıkarlar doğrultusunda sömürülmesiyle kredi kartı dolandırıcılığı faaliyetleri meydana gelmektedir [254]. Bu faaliyetler genellikle, kart sahibinin haberi olmaksızın; otomatik vezne makinesinden para çekme ve bir iş yerinden, çevrim içi platformlardan veya telefon bankacılığı kullanılarak satın alma ya da işlem yapma şeklinde gerçekleştirilmektedir.



Şekil 4.11. Kredi kartı dolandırıcılığı türleri

Teknoloji geliştikçe ve değiştikçe, dolandırıcıların teknikleri ve dolayısıyla hileli faaliyetleri yürütme biçimleri de değişmektedir. Bu bağlamda gerçekleştirilen eylemlerin, hem kart sahiplerine hem de kart sağlayıcılarına ciddi maddi yükler yüklemesinin yanında, uyuşturucu kaçakçılığı ve organize suç gibi yasadışı faaliyetler için finansman sağlamak gibi sektör üzerinde daha geniş çaplı etkileri de bulunmaktadır. Kredi kartı dolandırıcılığı, Şekil 4.11'de verildiği üzere; kartla ilgili dolandırıcılıklar, satıcıyla ilgili dolandırıcılıklar ve internet dolandırıcılıkları olmak üzere üç şekilde gerçekleştirilmektedir.

Kartla ilgili dolandırıcılıklar, geleneksel yöntemleri kapsar [255]. Uygulama dolandırıcılığı, kredi kartı başvurusu tahrif ettiğinde ortaya çıkar ve üç şekilde gerçekleştirilir: bir bireyin yasadışı bir şekilde başka bir bireyin kişisel bilgilerini ele geçirdikten sonra kendi adına hesaplar açmasıyla, bireyin kredi elde etmek için finansal durumu hakkında yanlış bilgiler sağlamasıyla, bir kartın sahibine ulaşmadan posta hizmetinden çalınmasıyla. Kayıp ve çalınmış kartlar, bir dolandırıcının teknolojiye yatırım yapmaksızın diğer bireyin bilgilerini ele geçirmesinin en kolay yoludur ve de üstesinden gelinmesi en zor olan dolandırıcılık türlerindendir. Hesap ele geçirme, dolandırıcının meşru bir müşteriye ait hesabın kontrolünü devralıp orijinal kart sahibi gibi göründüğü durumdur. Sahte kart dolandırıcılığı ise; kayıp veya çalınmış kartın metalik şeridini güçlü bir elektromıknatısla silerek, sıfırdan sahte kart oluşturarak, kart ayrıntılarını değiştirerek ya da kartların manyetik şeritlerindeki bilgileri kopyalayarak gerçekleştirilmektedir.

Satıcıyla ilgili dolandırıcılıklar, ticari kuruluşunun sahipleri veya çalışanları tarafından başlatılır [256]. Satıcı danışıklığı, ticari kurumların kart sahibi olan müşterilerinin kişisel bilgilerini dolandırıcılara iletmeleriyle ortaya çıkmaktadır. Üçgenleme ise, dolandırıcılar

tarafından oluşturulan bir web sitesinden alışveriş yapan kurbanın bilgilerinin ele geçirilerek, çalıntı kredi kartıyla başka ürünlerin satın alınması ile gerçekleşmektedir.

İnternet dolandırıcılıkları; sınır ötesi sosyal, ekonomik ve politik alanların genişlemesiyle oluşan ideal zeminde kolay bir şekilde gerçekleştirilmektedir [257]. Site klonlama, dolandırıcıların tüm bir siteyi veya yalnızca siparişin verildiği sayfaları klonlayarak müşteri bilgilerini ele geçirmesiyle ortaya çıkmaktadır. Uydurma satış siteleri yönteminde, dolandırıcılar tarafından hazırlanan siteler müşteriye son derece ucuz bir hizmet sunarken erişim karşılığında sadece ad, adres ve kredi kartı gibi bilgileri talep etmektedirler. Bu siteler genellikle, gelirlerini artırmak için topladığı ayrıntıları kullanan daha büyük bir suç ağının parçasıdır ya da küçük dolandırıcılara geçerli kredi kartı bilgilerini satarlar. Kredi kartı üreteçleri ile gerçekleştirilen dolandırıcılıklar ise, yasadışı yollarla geçerli kredi kartı numaraları ve son kullanma tarihleri üreten yazılımların kullanıldığı faaliyetlerdir.

Bir ödeme sistemi içinde, kredi kartı dolandırıcılığının önlenmesinin temel alanları: gerçek sorunların anlaşılması, dolandırıcılık önleme politikası, dolandırıcılık bilinci, dolandırıcılık önleme teknolojisi, kimlik yönetimi ve yasal caydırıcılık iken bu alanların destekleyicileri: kullanıcılar, kurumlar, ağ, hükümet ve sanayidir [258]. Gerçek sorunları anlama sürecindeki önemli unsurlar arasında dolandırıcılık verilerinin toplanması ve yönetimi bulunmaktadır. Dolandırıcılık önleme politikası; bankalar ve diğer finansal kurumlar, hükümet organları ve sanayi organları tarafından güncel tehditleri ve tedbirleri kapsayacak şekilde oluşturulmaktadır. Dolandırıcılık bilinci, farkındalığını artırmak ve suç fırsatlarını en aza indirmek için gerçekleştirilen eğitim ve bilgilendirme faaliyetlerini kapsamaktadır. Dolandırıcılık önleme teknolojisi, oldukça fazla seçenek sağlamasına rağmen her biri uygulanabilir değildir çünkü her seçeneğin, maliyeti ve faydası doğrultusunda detaylı olarak değerlendirilmesi gerekmektedir. Kimlik yönetimi, kişisel kimlik bilgilerinin sahipliğini, kullanımını ve korunmasını tanımlayan teknik sistem, kural ve prosedürlerin birleşimi ile soruşturma ve kovuşturma süreci için de gereklidir. Son olarak yasal caydırıcılık, önleme tedbirlerine rağmen dolandırıcıların faaliyetlerini engellemek için cezai yaptırımların güçlendirilmesidir.

Kredi kartı dolandırıcılığını önlemek, durdurmak veya baskılamak adına uygulanan; kimlik ve erişim yönetimi, farkındalık kampanyaları ve eğitimi, güvenlik yönetimi ya da olay

Çizelge 4.4. Kredi kartı dolandırıcılığı çözümleri

Kaynak	Amaç	Veri Kümesi	Yöntem/ Kısaltması		Sonuç				
					AUC	ACC	Recall	Prec	F-mea
[259]	Farklı özellik çıkarım yöntemleri ile kredi kartı işlem dolandırıcılığı tespiti	Bir Belçika kredi kartı düzenleyicisine ait veri	Yenilik-Sıklık-Maliyet Çerçevesi	RFM	98,6	98,77	*	*	*
[202]	Büyük veri analizi ile kredi kartı dolandırıcılığı sınıflandırma	ccFraud	Parçacık sürü optimizasyonu ve otomatik ilişkisel sinir ağı	PSOAANN	*	*	88,64	*	*
[260]	Sınıf dengesizliği ile başa çıkarak dolandırıcılık tespiti	Bir Tayvan kredi kartı şirketine ait veri	Veri örnekleme ile melez topluluk modeli	RUSMRN	*	79,73	53,36	*	*
[261]	Kredi kartı dolandırıcılığında kullanılan sınıflandırıcıları karşılaştırma	UCSD-FICO	Birden fazla öğrenciyi birleştirme	Rotation Forest	*	98,25	98	98	98
			Karar Ağacı	REPTree	*	98,07	98	98	98
			Yapay sinir ağı	MLP Classifier	*	97,98	98	98	97
			Tembel öğrenci	LWL	*	97,89	98	98	97
			Bayes	A2DE	*	97,90	98	98	98
			Kural Seti	Decision Table	*	97,93	98	98	98
[262]	Manuel ve otomatik dolandırıcılık sınıflandırma kombinasyonu	Bir çevrimiçi moda perakendecisine ait veri	Skorlama ve sınıflandırma		*	*	58,7	40,7	*
[263]	Kredi kartı dolandırıcılığında çeşitli teknikleri karşılaştırma	German (Statlog)	Yapay Sinir Ağları	BR	98,74	*	*	*	*
				GDA	98,63	*	*	*	*
				LM	95,57	*	*	*	*
		Australian Credit Approval	Yapay Sinir Ağları	BR	99,02	*	*	*	*
				GDA	96,29	*	*	*	*
				LM	93,86	*	*	*	*
[264]	Hibrit yaklaşımları farklı stratejiler ile skorlayarak dolandırıcılık tespiti	Bir Türk bankasına ait veri	Modeller topluluğunda, iyimser kötümser ve ağırlıklı oylama	OPT	*	96,60	31,59	94,65	*
				PES	*	86,65	93,92	26,05	*
				WGT	*	97,55	64,02	82,12	*
[265]	Derin otomatik kodlayıcı tabanlı dolandırıcılık tespiti	German (Statlog)	Uyumsuz yöntem	UDAE	*	81,6	*	*	*
			Aşırı uyumlu yöntem	ODAE	*	84,1	*	*	*

Çizelge 4.4. (devam) Kredi kartı dolandırıcılığı çözümleri

Kaynak	Amaç	Veri Kümesi	Yöntem/ Kısaltması		Sonuç				
					AUC	ACC	Recall	Prec	F-mea
[266]	Şüpheli ödeme örüntülerinden dolandırıcılık tespiti	Worldline	Alt çizgelerden yeni öznitelikler elde ederek sınıflandırma	RF	97,1	99,9	44,4	37,1	*
				LR	94,4	99,6	40,6	9,5	*
[267]	Farklı sınıf dengesizliği oranlarında dolandırıcılık tespiti	Bir bankaya ait veri (3:97)	Karar Ağaçları	J48	97,58	*	*	*	*
				RF	96,97	*	*	*	*
		Bir bankaya ait veri (25:75)	Karar Ağaçları	J48	93,5	*	*	*	*
				RF	94,32	*	*	*	*
[268]	Proaktif olay güdümlü karar verme prototipi	SPEEDD	Olay Tanımlarının Çevrimiçi Öğrenilmesi	OLED	*	*	77,6	89,4	83
			Endüktif Mantık Programlama	SC	*	*	87,4	91,2	89,2
[269]	Sınıf dengesiz verideki dolandırıcılıkları gerçek zamanlı tespit etme	UCSD-FICO	Ağaç tabanlı meta sınıflandırıcı	TBMC	86	*	81	*	*
[270]	Aykırı değer tespiti tabanlı ve değişen örüntülere dayanıklı dolandırıcılık tahmini	European	Tutarlılık Tahmin Yöntemi		89.37	*	*	*	*
[271]	Kötüye kullanıma dayalı dolandırıcılık tespiti	Bir Çin e-ticaret şirketine ait veri	Sınıflandırma ve regresyon ağaçları tabanlı rastgele orman algoritması		*	98,67	59,62	32,68	*
[272]	Dolandırıcılık sınıfını aşırı örnekleme ile sınıflandırma	European	Çekişmeli Üretici Ağ	GAN	*	99,963	73,282	93,204	82,051
			Sentetik Azınlık Aşırı Örnekleme Tekniği	SMOTE	*	99,962	69,466	96,809	80,889
[273]	Kredi kartı işlem dolandırıcılığı tahmini	Bir bankaya ait veri	Kümeleme ve destek vektör makinesi	KSVM	79,49	*	96	96,1	95,6
			Kümeleme ve destek vektör makinesi ile uyarlamalı artırma	KASVM	98,72	*	98,3	98,3	98,2
[188]	Dolandırıcılık tespit seviyesini artırma	Brazilian	Paralel ve artımlı öğrenme topluluğu	TWB	77,38	*	69,32	*	*

*:Belirtilmemiş

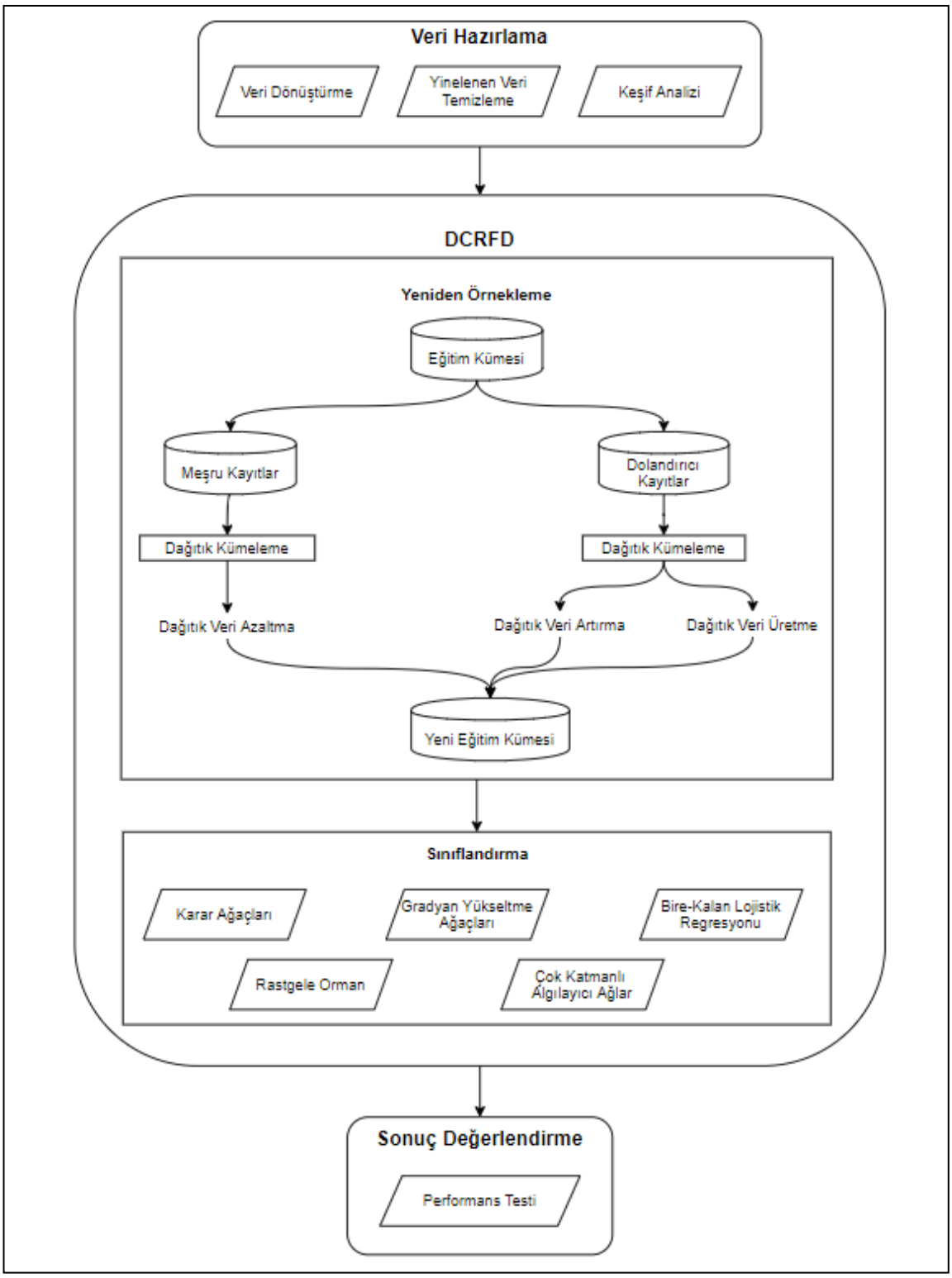
müdahale planlaması gibi eylemler her ne kadar riski sınırlandırsa da, veri sızıntısı konusu finans kurumlarının karşı karşıya olduğu en önemli problemlerden biridir [274]. 2019 yılında sızdırılan kredi kartı verilerinin 2018'e göre %212 oranında artmasıyla, tarihteki en büyük küresel veri sızıntısına ulaşılmıştır [275]. Bu artış ise, daha önce bağlantısı kesilmiş bilgileri tek ve zengin bir profilde birleştirmek için aralarında çapraz referanslar kurarak daha derin analizleri mümkün kılan *"mozaik etkisi"* durumunun ortaya çıkmasına sebep olmaktadır [276]. Hem mozaik etkisi hem de finans verisinin dinamik ve mahrem doğası gereği, kurumlar dolandırıcılıkla ilgili bilgilerini paylaşma konusunda genellikle isteksizdirler [277]. Bu durum, analiz edilen veri kümelerinin gerçekleri yansıtmamasına ek olarak dolandırıcılık ilişkilerinin de doğru tespit edilememesine yol açmaktadır.

Bu ve benzer güçlükler doğrultusunda, kredi kartı dolandırıcılığı sorununa çözüm üretmek amacıyla son beş yılda önerilen akademik çalışmalar Çizelge 4.4'de karşılaştırmalı olarak sunulmuştur. Kredi kartı dolandırıcılığı tespiti çözümleri, yüksek hacimli ticari işlemler içerisinde oldukça düşük yüzdeye sahip dolandırıcılık işlemlerini çoğunlukla geleneksel yöntemler kullanarak tespit etmeye odaklanmaktadır. Geleneksel yöntemler sadece zaman alıcı, pahalı ve yanlış alarm sayısı yüksek olmakla kalmaz, aynı zamanda büyük veri çağında pratik değillerdir.

Azınlık sınıfı, doğal olarak büyük veride çok daha düşük bir orana sahiptir ve genellikle pek çok alt kavramı barındırır. Fakat bu alt kavramlar, büyük bölmeler yaratmaya meyilli sınıflandırıcılar tarafından göz ardı edilerek iyi bir şekilde temsil edilemezler. Yeniden örnekleme yöntemleri, harici ve taşınabilir olmaları nedeniyle sınıf dengesizliği problemi ile başa çıkmak için yaygın olarak kullanılmaktadır. Bu yöntemlerin uygulanması çok basit olsa da, en etkili şekilde ayarlanmaları kolay bir iş değildir. Özellikle, veri artırmanın veri azaltmadan daha etkili olup olmadığı ve hangi örnekleme oranının kullanılması gerektiği belirsizdir ve bu durum her uygulama alanı için farklılık göstermektedir [278]. Büyük veri analitiği söz konusu olduğunda ise, kimi yeniden örnekleme yöntemleri paralel ve dağıtık yapılar üzerinde uygulanabilir değildir.

Alt kavramları tespit etmedeki güçlüğü gidermek ve büyük veri kümelerinde ölçeklenebilir bir sınıflandırma sağlayabilmek amacıyla, Dolandırıcılık Tespiti için Dağıtılmış Küme Tabanlı Yeniden Örnekleme (Distributed Cluster-based Resampling for Fraud Detection, DCRFD) adlı yeni bir çözüm yöntemi önerilmiştir. Şekil 4.12'de metodolojisi verilen

DCRFD sayesinde, yeniden örnekleme sınıflardaki alt davranışlara göre yapılırken, problem paralel analiz için alt problemlere dönüştürülür ve böylelikle daha başarılı sınıflandırma modelleri üretilir.



Şekil 4.12. DCRFD yönteminin metodolojisi

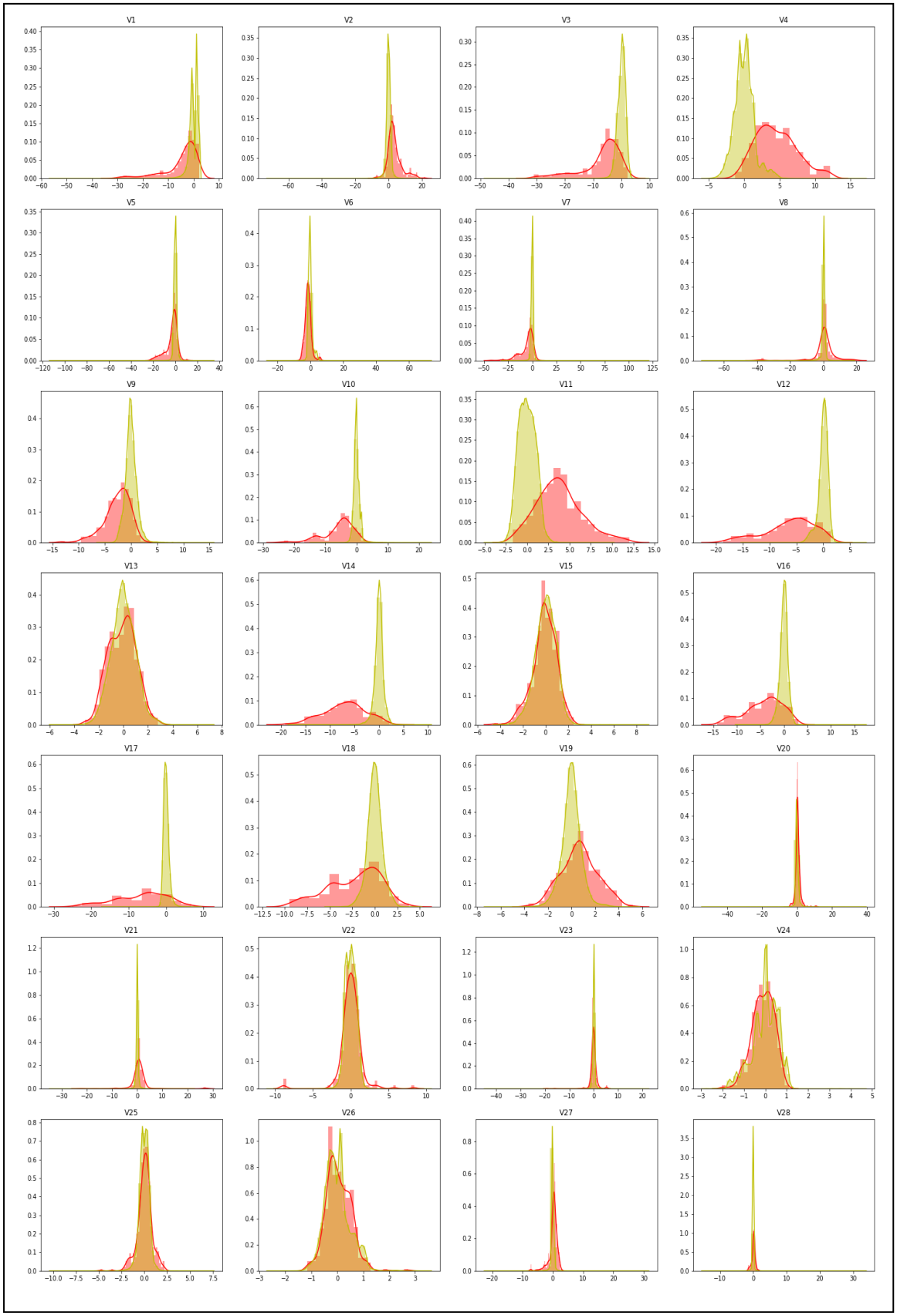
DCRFD, aşağıda verilen yöntemlerin bir arada kullanılması ve bu yöntemlerin iyileştirilmesiyle oluşturulmuştur.

1. Veri kümesi, yeniden örnekleme işleminden önce eğitim kümesi ve test kümesi olarak ikiye ayrılmıştır. Her ne kadar eğitim verilerindeki temsili azaltmış ve sınıflandırıcının tahminlerini zayıflatmış gibi görünse de [279], bu sayede bilgi sızıntısının önüne geçilmesi amaçlanmıştır.
2. Azınlık ve çoğunluk sınıfları, dengesiz büyük verilerdeki farklı davranışları belirlemek için ayrı ayrı kümelenmiştir [280].
3. Sınıflar arasındaki dengesizlik oranını azaltmak için azınlık sınıfları aşırı örneklenirken çoğunluk sınıfından örneklem alınmıştır [281].
4. Alt kavramlardaki örneklerin kalitesini sağlamak için çoğunluk sınıfı için küme temelli örnekleme [282] ve azınlık sınıfı için küme temelli aşırı örnekleme [283] uygulanmıştır.
5. Azınlık sınıf ile çoğunluk sınıfın örnekleri eşitleyen yöntemlerin [284] tersine, bu modelde örnekleme oranları kullanılmıştır. Bu sayede aşırı öğrenmeden ve önemli örneklerin atılmasından kaçınılmıştır.
6. Performansı artırmak ve ağ trafiğini azaltmak için dağıtılmış veri işleme paralel hale getirilerek yeniden örnekleme yöntemleri bölümlerde yürütülmüştür [285].

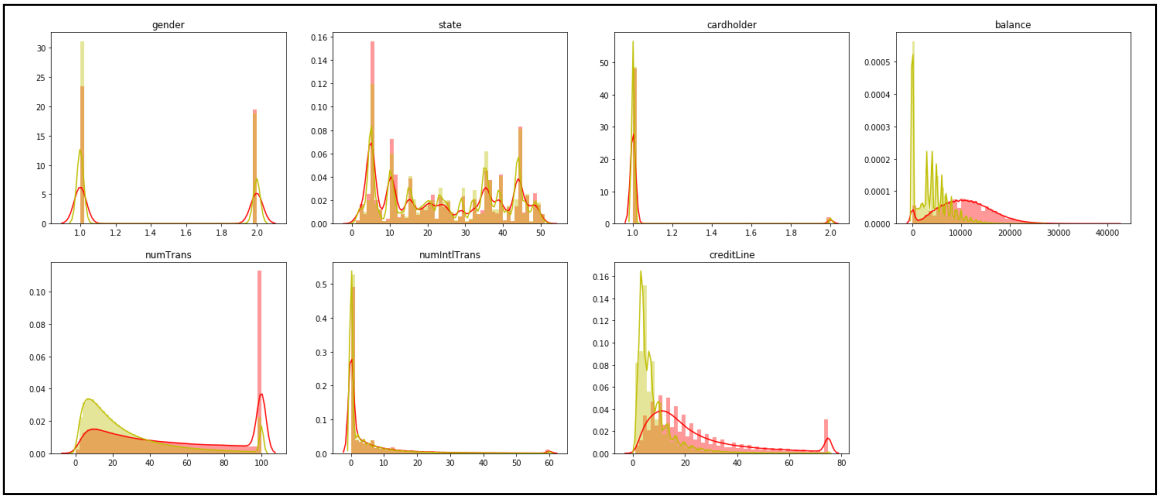
4.2.1. Veri kümeleri

Önerilen kredi kartı dolandırıcılığı tespiti yöntemlerinin başarım değerlendirmesinde kullanılmak üzere iki farklı açık veri kümesi kullanılmıştır; European ve ccFraud. European, Eylül 2013'te Avrupalı kart sahipleri tarafından kredi kartları ile yapılan iki günlük işlemleri içermektedir [286]. Veri kümesi, gizlilik sorunları nedeniyle orijinal özellikler ve veriler hakkında fazla arka plan bilgisi sağlanamamaktadır. Çok boyutlu uzaydaki bir verinin daha düşük boyutlu bir uzaya izdüşümünü ifade eden PCA sonucu olan sayısal giriş değişkenlerinden oluşmaktadır. ccFraud ise, bir Microsoft ürünü olan RevoScaleR paketi için hazırlanmış örnek bir veri kümesidir [287].

Veri kümelerindeki özniteliklerin, sınıf etiketi bağlamında dağılımları Şekil 4.13'de ve Şekil 4.14'de verilmiştir. Yeşil ve kırmızı kutucuklar sırasıyla dolandırıcı ve meşru veriye ait histogramı yani verinin dağılımını göstermektedir. Histogramın x ekseni tüm değer aralığındaki bölme sayısını belirtirken y ekseni belirli bir bölmenin frekansını temsil



Şekil 4.13. European veri kümesinin öznitelik yoğunluk grafiği



Şekil 4.14. ccFraud veri kümesinin öznitelik yoğunluk grafiği

etmektedir. Çizgiler ise, popülasyonla çıkarımın yapıldığı temel bir veri düzeltme olan çekirdek yoğunluğu tahminini göstermektedir. Şekillerden de görülebileceği üzere, veri kümelerine ait dağılımların çoğu normal olarak kabul edilen çan şekline benzemeyen çarpık yapılara sahiptirler. Bu durum, ortalamaya odaklanmak yerine veri kümesinin uçlarının daha etkili olduğunu göstermektedir.

Çizelge 4.5. Dolandırıcılık veri kümelerine ait özet bilgiler

	Meşru Kayıt	Dolandırıcı Kayıt	Toplam Kayıt	Dengesizlik Oranı	Toplam Öznitelik
European	283253	473	283726	598,84	28
ccFraud	9403986	596014	10000000	15,77	7

European ve ccFraud veri kümelerine ait bilgiler Çizelge 4.5'de özetlenmiştir. Her iki veri kümesinin de yüksek dengesizlik oranına sahip olduğu görülmektedir. Görünmeyen veriler üzerinde model performansının daha iyi bir göstergesini elde edebilmek amacıyla ve sınıf dengesizlik oranı aynı kalacak şekilde, veri kümelerindeki örneklerin %70'i eğitim kümesi, geri kalanı ise test kümesi olarak ayırılmıştır.

Kümeleme, yeniden örnekleme ve sınıflandırma, özelliklerin ölçeklerinden oldukça etkilendiği ve beklenmeyen sonuçlar ürettiği için eğitim kümesine normalizasyon uygulandıktan sonra bu analizler gerçekleştirilmiş ve modellerin test verisi üzerinde elde ettiği sonuçlar üzerinden karşılaştırmalar yapılmıştır.

4.2.2. Yeniden örnekleme

Yeniden örnekleme yöntemleri, büyük veri kümeleri üzerinde uygulandığında hem zaman alıcıdır hem de dağıtık ortamlarda başarısız olma eğilimindedir. Dağıtık dosya sisteminde rastgele parçalara ayrılmış veri üzerinde çalışan fonksiyonlar, uzamsal ilişkileri olmayan örneklerden yapay ilişkiler elde edilmesine yol açabilmektedir. Bu durumun önüne geçmek için, meşru ve dolandırıcı kayıtlar ayrıştırılarak hem alt kavramlarının tespiti hem de problemin dağıtık analiz için alt problemlere bölünmesi amacıyla kümeleme yöntemi uygulanmıştır. Kolay uygulanabilir olan ve yoğunlaşmış bölgelerin etkili bir şekilde algılanabildiği *k*-means kümeleme algoritması kullanılmıştır. Yeniden örnekleme, tespit edilen bu kümeler üzerinde gerçekleştirilmiştir.

Çizelge 4.6. Örnekleme stratejileri

Örnekleme Kısaltması	Örnekleme Ayrıntıları
NUNO	Veri azaltma yok, veri artırma yok
NUO	Veri azaltma yok, azınlık sınıf ROS ile %100 artırılmış
NUO2	Veri azaltma yok, azınlık sınıf ROS ile %200 artırılmış
NUS	Veri azaltma yok, azınlık sınıf SMOTE ile %100 artırılmış
NUS2	Veri azaltma yok, azınlık sınıf SMOTE ile %200 artırılmış
UNO	Çoğunluk sınıfın %90'ı RUS ile alınmış, veri artırma yok
U2NO	Çoğunluk sınıfın %80'i RUS ile alınmış, veri artırma yok
UO	Çoğunluk sınıfın %90'ı RUS ile alınmış, azınlık sınıf ROS ile %100 artırılmış
US	Çoğunluk sınıfın %90'ı RUS ile alınmış, azınlık sınıf SMOTE ile %100 artırılmış
UO2	Çoğunluk sınıfın %90'ı RUS ile alınmış, azınlık sınıf ROS ile %200 artırılmış
US2	Çoğunluk sınıfın %90'ı RUS ile alınmış, azınlık sınıf SMOTE ile %200 artırılmış
U2O	Çoğunluk sınıfın %80'i RUS ile alınmış, azınlık sınıf ROS ile %100 artırılmış
U2S	Çoğunluk sınıfın %80'i RUS ile alınmış, azınlık sınıf SMOTE ile %100 artırılmış
U2O2	Çoğunluk sınıfın %80'i RUS ile alınmış, azınlık sınıf ROS ile %200 artırılmış
U2S2	Çoğunluk sınıfın %80'i RUS ile alınmış, azınlık sınıf SMOTE ile %200 artırılmış

Sınıf yapılarının farklılığı, sınıflar arasındaki mesafelerin değişkenliği ve sınıf içi dengesizliklerin varlığı sebebiyle yeniden örnekleme tekniklerinin performansı her veri kümesi ve yeniden örnekleme oranı için farklılık göstermektedir. Oluşturulan kümeler bu nedenle, Çizelge 4.6'da verildiği üzere on beş farklı strateji dâhilinde RUS, ROS ve SMOTE teknikleri ile yeniden örneklendirilmiştir.

N boyutlu bir sınıf dengesiz veri kümesinin, *MA* olarak ifade edilen çoğunluk sınıfı örneklerinden ve *MI* olarak ifade edilen azınlık sınıfı örneklerinden oluştuğu ve *k* kümeye bölündüğü varsayıldığında, $1 \leq i \leq k$ olmak üzere *i* kümesindeki yeni örnek sayısı *r* yeniden örnekleme oranı doğrultusunda RUS ile Eş. 5.1'de ve ROS ile Eş. 5.2'de verildiği şekilde oluşturmaktadır. RUS ve ROS'dan farklı olarak SMOTE, *MI* sınıfının bir örneği olan *x*'in en yakın komşularından *n* tane seçerek, yeni örnekleri Eş. 5.3'de verildiği gibi enterpolasyon ile üretmektedir.

$$NN_{MA}^{i} = N_{MA}^{i} - N_{MA}^{i} * r \quad (5.1)$$

$$NN_{MI}^{i} = N_{MI}^{i} + N_{MI}^{i} * r \quad (5.2)$$

$$NN_{MI}^{i} = N_{MI}^{i} + \left(x + rand(0,1) * \left(x_j - x\right)\right) * r \, , j = 1,2,3 \ldots n \quad (5.3)$$

4.2.3. Sınıflandırma yöntemleri

Sınıflandırma, bilgisayarların verileri önceden belirlenmiş özelliklere göre gruplandırdığı bir süreçtir. Danışmanlı makine öğrenmesi yaklaşımlarından biri olan sınıflandırma, bir dizi bilinen girdi ile eşleşen çıktının oluşturduğu eğitim kümesi doğrultusunda, daha önce görülmeyen girdi verilerinin çıktısını tahmin etmeyi amaçlar. Sınıflandırma algoritmaları, karar ağaçları, Bayes sınıflaması, kural tabanlı yöntemler, sinir ağları ve destek vektör makineleri gibi geniş bir öğrenme yöntemi alanını kapsamaktadır. Tüm bu algoritmalar pek çok defa çeşitli kıstas veri kümeleri ve koşullarda kullanılmış ve performansları açısından yetenekleri kanıtlanmıştır [288].

Büyük verinin ortaya çıkmasıyla daha ölçeklenebilir makine öğrenme paketlerine ihtiyaç duyulmuştur. Apache Spark, çeşitli makine öğrenimi algoritmalarının dağıtılmış ve ölçeklenebilir uygulamalarını sağlaması ve sürekli büyüyen veri kümelerini işlemeyi mümkün kılması ile öne çıkan dağıtık makine öğrenmesi paketi çerçevelerinden biri haline gelmiştir [289]. Spark, en önemli ve en sık kullanılan makine öğrenme algoritmalarının dağıtık uygulamalarını sunmaktadır. Spark tarafından sınıflandırma amacıyla sunulan ve tez kapsamında kullanılan bu algoritmalar: Karar Ağacı (Decision Tree, DT), Rastgele Orman (Random Forest, RF), Gradyan Yükseltme Ağaçları (Gradient-Boosted Trees, GBT), Çok Katmanlı Algılayıcı Ağlar (Multilayer Perceptron, MLP) ve Bire-Kalan Lojistik Regresyonu (One-vs-Rest Logistic Regression, OVR)'dur.

Karar Ağacı; düğümler, dallar ve yapraklar içeren ağaç yapısı biçimindeki bir sınıflandırıcıdır. Düğümler, belirli bir özniteliğin değerinin test edilmesini; dallar, testin sonucunu; yapraklar ise sonucun tahminini oluşturan sınıf etiketlerini temsil etmektedir [290]. Karar ağacı veri normalleştirmesi gerektirmez, sayısal ve kategorik verileri işleyebilir ve eksik değerlerle çalışabilir.

Rastgele Orman, daha yüksek bir tahmin doğruluğu elde etmek amacıyla sınıflandırma sürecinde bir topluluk olarak çalışan bir dizi karar ağacı kullanılır. Komite olarak çalışan çok sayıda görece ilişkisiz ağaç, herhangi bir kurucu modelden daha iyi performans göstermektedir [291].

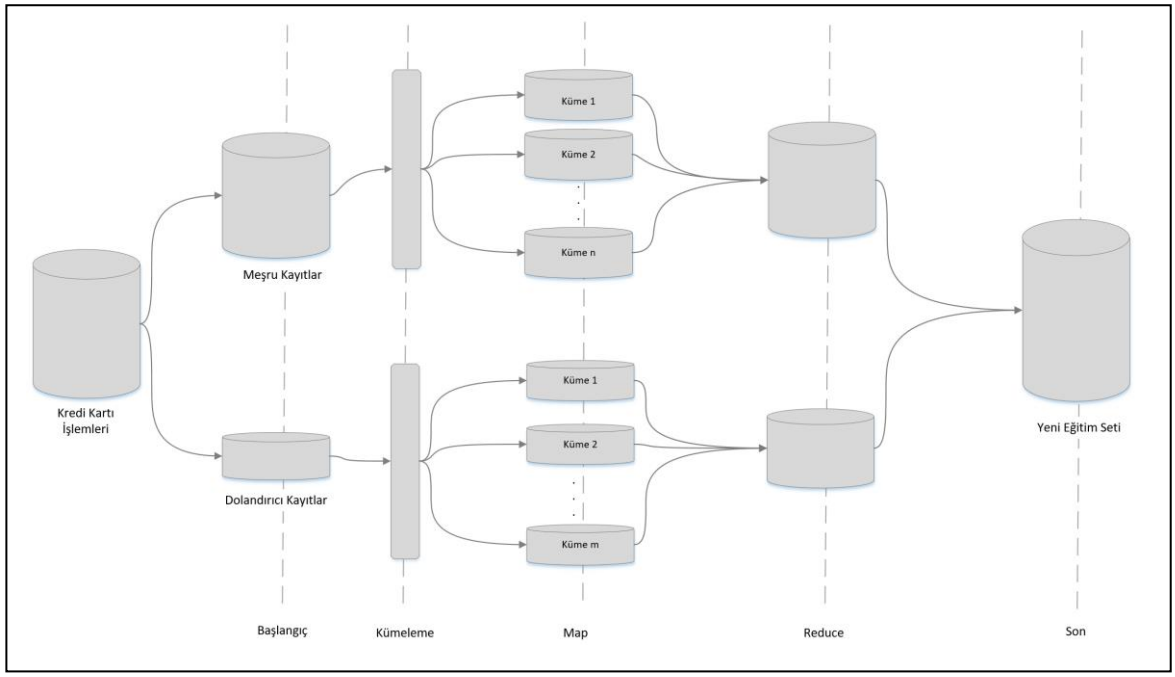
Gradyan Yükseltme Ağaçları, zayıf tahmin modellerinin bir topluluğu şeklinde bir sınıflandırma veya regresyon modeli üretir. Model, isteğe bağlı farklılaşabilir bir kayıp fonksiyonunun optimizasyonuna izin veren bir arttırıcı yöntem ile genelleştirme sağlar. Böylelikle, zayıf öğrenenler tek bir güçlü öğrenende yinelemeli bir şekilde birleştirilir [292]. GBT ve RF, ağaçların inşa edilme ve sonuçların birleştirilme biçiminde farklılık gösterir. GBT, her yeni ağacın önceden eğitilmiş ağaç tarafından yapılan hataları düzeltmeye yardımcı olduğu tek bir ağaç oluştururken, RF verilerin rastgele bir örneğini kullanarak her ağacı bağımsız olarak eğitir. RF sürecin sonunda sonuçları birleştirirken, GBT sonuçları süreç boyunca birleştirir.

Çok Katmanlı Algılayıcı Ağlar, doğrusal olarak ayrılamayan veri kümelerini gürbüz olarak sınıflandırabilen ileri beslemeli bir yapay sinir ağıdır. Sinyali almak için bir giriş katmanı, giriş hakkında bir karar veya tahmin yapan bir çıkış katmanı ve bu ikisi arasında esas hesaplamaları gerçekleştiren rastgele sayıda gizli katmandan oluşur [293]. MLP'nin tahmin yeteneği, ağların hiyerarşik yapısından gelir. Bu yapı, farklı ölçeklerde veya çözünürlüklerde özellikleri seçebilir ve bunları daha üst düzey özelliklerde birleştirebilir.

Bire-Kalan Lojistik Regresyonu, sınıf başına tek bir sınıflandırıcıyı eğitmeyi içerir. Her seferinde sadece bir sınıfın örneklerini alarak sınıfların her biri için bir ikili sınıflandırma problemi yaratarak tahminler üretir. Tahminler, her bir ikili sınıflandırıcı değerlendirilerek yapılır ve en güvenli sınıflandırıcının indisi, etiket olarak verilir [294].

4.2.4. MapReduce şeması

Büyük ölçekli uygulamaları kullanmak için basitlik ve esneklik sağlayan MapReduce paralel programlama paradigmasının yardımıyla, DCRFD prosedürü Şekil 4.15'de verildiği gibi; Başlangıç, Kümeleme, Map, Reduce ve Son olmak üzere beş adımda geliştirilmiştir.



Şekil 4.15. DCRFD MapReduce akış şeması

1. Başlangıç: Bağımsız HDFS bloklarındaki parçalara ayrılmış orijinal veri, eğitim ve test amacıyla ikiye ayrılır. Daha sonra, eğitim verisindeki dolandırıcı kayıtları yani azınlık sınıfı ile meşru kayıtlar yani çoğunluk sınıfı ayrıştırılır.
2. Kümeleme: Alt kavramların tespiti için çoğunluk sınıfı ve azınlık sınıfı ayrı ayrı dağıtık olarak kümelenir. Dağıtık olarak gerçekleştirilen kümeleme işlemi, Şekil 4.16'da sözde kodu verildiği şekilde gerçekleştirilmiştir.
3. Map: İşlemler, elemanı oldukları kümelere göre sıralanırlar. Her parçada bir ya da birkaç kümeye ait elemanlar bulunacak şekilde veri haritalandırılır. Yeniden örnekleme stratejileri doğrultusunda çoğunluk sınıfından örneklem alma, azınlık sınıf için aşırı örnekleme ve veri artırma yöntemleri uygulanır. Bu yöntemler Şekil 4.17'de verilen sözde kod doğrultusunda gerçekleştirilir.
4. Reduce: Yeniden örnekleme stratejileri sonucunda elde edilen azaltılmış çoğunluk sınıfı ile artırılmış azınlık sınıfı birleştirilir.

5. Son: Bu aşamada, sınıflandırma için eğitim kümesi olarak kullanılmak üzere nihai veri elde edilir. Bu verinin boyutu, örnekleme oranlarına göre değişiklik göstermektedir.

Algoritma 1: Kümeleme için MapReduce Algoritması

```
Function MAP(anahtar,değer):
    if anahtar==ÇoğunlukSınıf then
        çoğunluk⟵ VeriGetir(değer)
        k_çoğunluk⟵ KümeSayısıHesapla(çoğunluk)
        c_çoğunluk⟵ KümeMerkeziHesapla(k_çoğunluk)
        son.Ekle(çoğunluk.KümeElemanlarınıAta(k_çoğunluk,c_çoğunluk))
    else
        azınlık⟵ VeriGetir(değer)
        k_azınlık⟵ KümeSayısıHesapla(azınlık)
        c_azınlık⟵ KümeMerkeziHesapla(k_azınlık)
        son.Ekle(azınlık.KümeElemanlarınıAta(k_azınlık,c_azınlık))
    end
    merkezler⟵ Birleştir(c_çoğunluk, c_azınlık)
    Yayınla(merkezler,son)
end
Function REDUCE(merkezler,son):
    güncel⟵ KümeMerkezleriniGüncelle(KayıtGetir(merkezler), KayıtGetir(son))
    Yayınla(güncel,son)
end
```

Şekil 4.16. Kümeleme için MapReduce sözde kodu

4.2.5. Deneysel sonuçlar

DCRFD yöntemi, verinin dağıtık olarak kümelenerek yeniden örneklenmesine dayanmaktadır. Kümeleme aşamada, birçok kümeleme yaklaşımı tercih edilebilir, fakat daha önemli olan unsur küme sayısı yani *k*'nın belirlenmesidir. İdeal *k*, genellikle SSE (Sum of Squared Errors) yani küme elemanlarının en yakın merkezlerine olan kare mesafelerinin toplamı grafiğindeki yerel bir minimumdur. Uygun değeri belirlemek amacıyla, *k*-means algoritması her eğitim kümesinde çalıştırılırken *"elbow tekniği"* kullanılmıştır. Elbow tekniğinde, *k*=2 ile başlanıp her adımda *k* bir artırılarak, her *k* değerinde tüm kümeler için ortalama bir puan hesaplar. Genellikle her bir noktadan kendisine atanan merkeze olan kare mesafelerin toplamı olarak karesel hataların toplamı (Sum of Squared Errors, SSE) hesaplanır. SSE'nin, önemli ölçüde düşüp sonraki artışlarda düzlüğün oluştuğu nokta ideal *k* değeri olarak belirlenir [295]. Şekil 4.18'de ccFraud veri kümesinin azınlık sınıfı için küme sayısı seçimi örnek olarak sunulmuştur. Bu durum için *k*=10 olarak belirlenmiştir.

Algoritma 1: DCRFD için MapReduce Algoritması

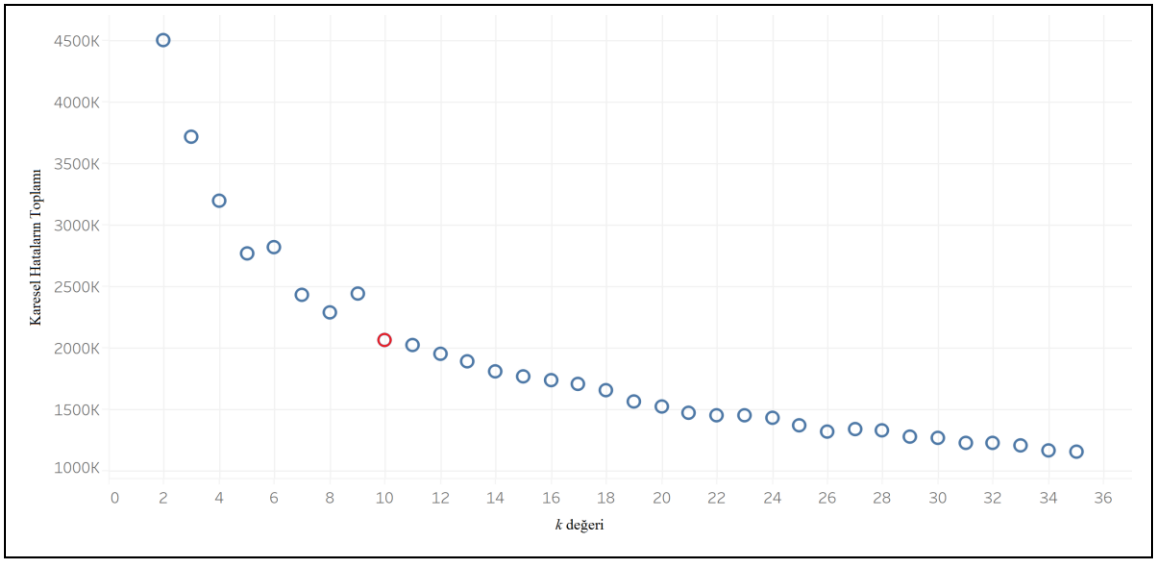
```
Function MAP(küme, değer):
    kayıt⟵ VeriGetir(değer)
    kayıt⟵ kayıt.KayıtSırala(küme)
    sınıf⟵ kayıt.SınıfGetir()
    if senaryo==RUS then
        örnekSayısı⟵ OranHesapla(çoğunluk.ElemanSayısı, ÖrneklemeOranı)
        yeni⟵ çoğunluk.AltListeOlustur(örnekSayısı)
        son.Ekle(yeni)
    else if senaryo==ROS then
        son.Ekle(azınlık)
        örnekSayısı⟵ OranHesapla(azınlık.ElemanSayısı, ÖrneklemeOranı)
        yeni⟵ azınlık.TekrarOlustur(örnekSayısı)
        son.Ekle(yeni)
    else if senaryo==SMOTE then
        son.Ekle(azınlık)
        örnekSayısı⟵ OranHesapla(azınlık.ElemanSayısı, ÖrneklemeOranı)
        komşular⟵ azınlık.YakınKomsuBul(KomsuSayısı)
        yeni⟵ azınlık.TekrarOlustur(komsular, örnekSayısı)
        son.Ekle(yeni)
    Yayınla(sınıf, son)
end
Function REDUCE(sınıf, son):
    if senaryo==RUS then
        eğitim⟵ Birlestir(son, azınlık)
    else if senaryo==ROS then
        eğitim⟵ Birlestir(çoğunluk, son)
    else if senaryo==SMOTE then
        eğitim⟵ Birlestir(çoğunluk, son)
    Yayınla(sınıf,eğitim)
end
```

Şekil 4.17. DCRFD için MapReduce sözde kodu

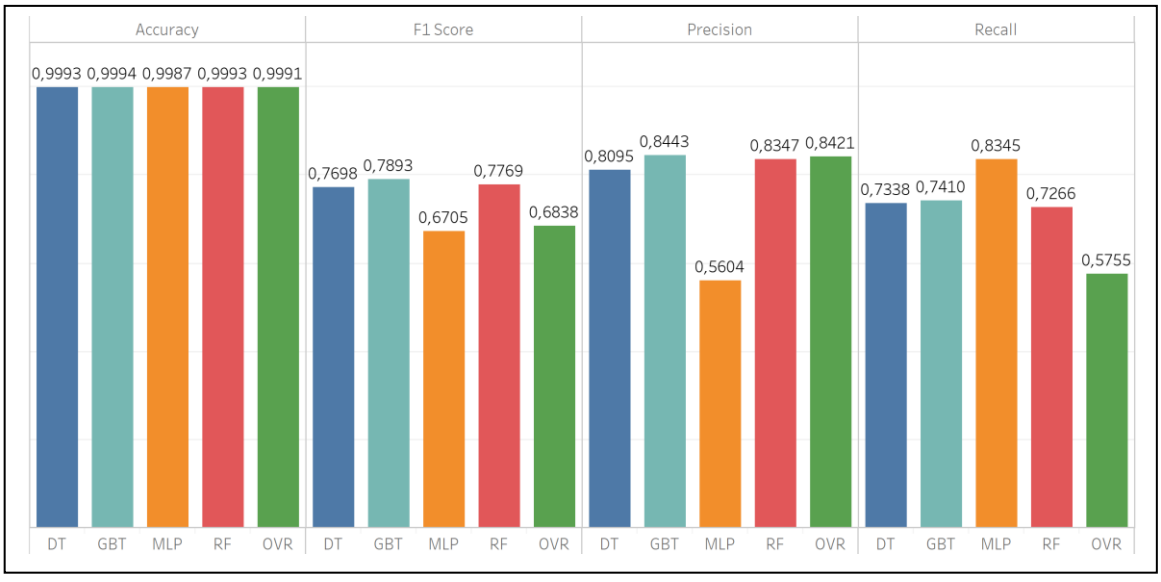
Yeniden örneklemelerin sınıflandırma başarısına etkisini ölçmede bir baz sınıflandırıcı belirlemek amacıyla eğitim kümesine ayrı ayrı RF, GBT, DT, MLP ve OVR sınıflandırıcıları uygulanmıştır. Kredi kartı dolandırıcılığı tespiti yöntemlerinin değerlendirilmesi için en sık tercih edilen metriklerin sırasıyla recall, precision, F_1-score ve accuracy olduğu göz önünde bulundurularak [296], sınıflandırma başarıları bu metrikler ile değerlendirilmiştir.

Şekil 4.19'da European ve Şekil 4.22'de ccFraud veri kümesine ait sınıflandırma başarıları verilmiştir. Her iki veri kümesi de sınıf dengesiz olduğu için accuracy metriğinin güvenilir olmadığı görülmektedir. Daha güvenilir bir metrik olan F_1-score doğrultusunda sınıflandırıcılar değerlendirildiğinde, çeşitli sınıflandırma yöntemlerinin karşılaştırıldığı

[297]'da da kanıtlandığı üzere, GBT en ideal sınıflandırıcı olarak belirlenmiştir. Yeniden örnekleme stratejileri doğrultusunda üretilen yeni eğitim kümesi üzerinde bu sınıflandırıcılar kullanılarak yeniden örneklemelerin etkisi karşılaştırılmıştır.



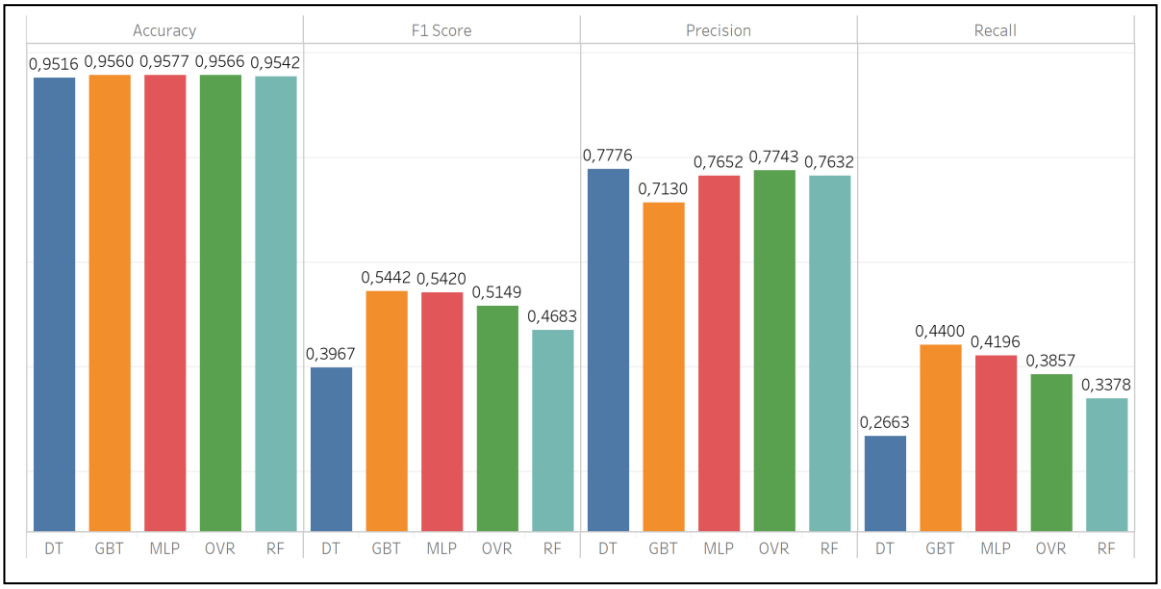
Şekil 4.18. ccFraud veri kümesinin azınlık sınıfı için küme sayısı maliyet ilişkisi



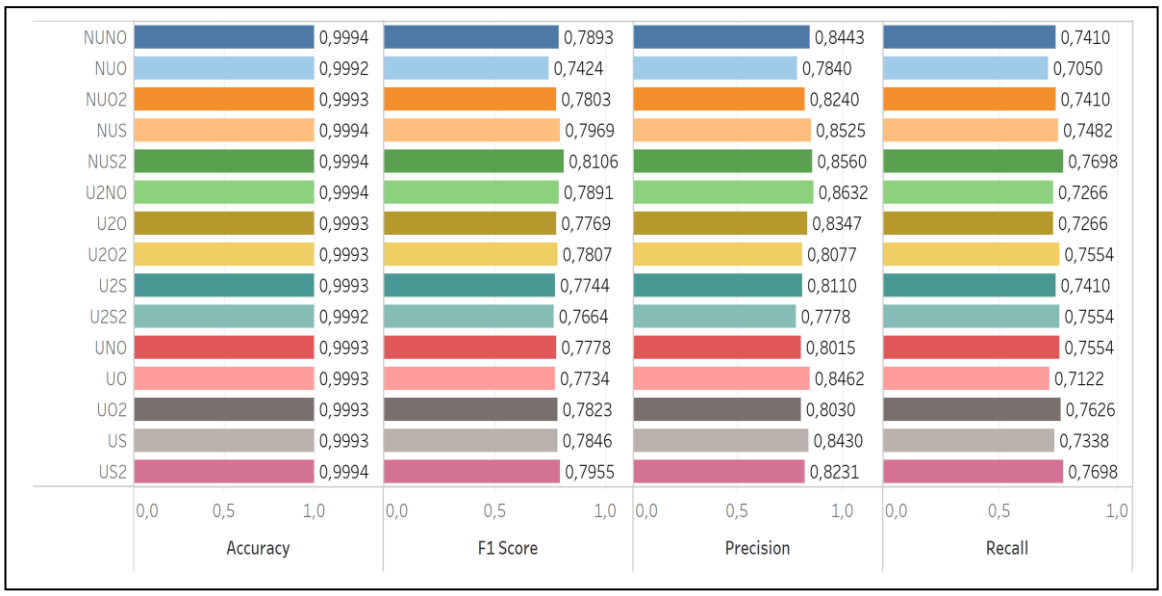
Şekil 4.19. European veri kümesinde sınıflandırıcı türüne göre başarı değerlendirmesi

Şekil 4.21'de verildiği üzere, F_1-score bağlamında European veri kümesi için en etkili örnekleme stratejisi 0,8106 ile NUS2 olmuştur. Daha az örnek hacmine sahip bu veri kümesinde daha başarılı bir sınıflandırma gerçekleştirmek ise, çoğunluk sınıfını korurken azınlık sınıfından büyük oranda sentetik veri üretilmesiyle mümkün olmuştur. Şekil 4.22'de

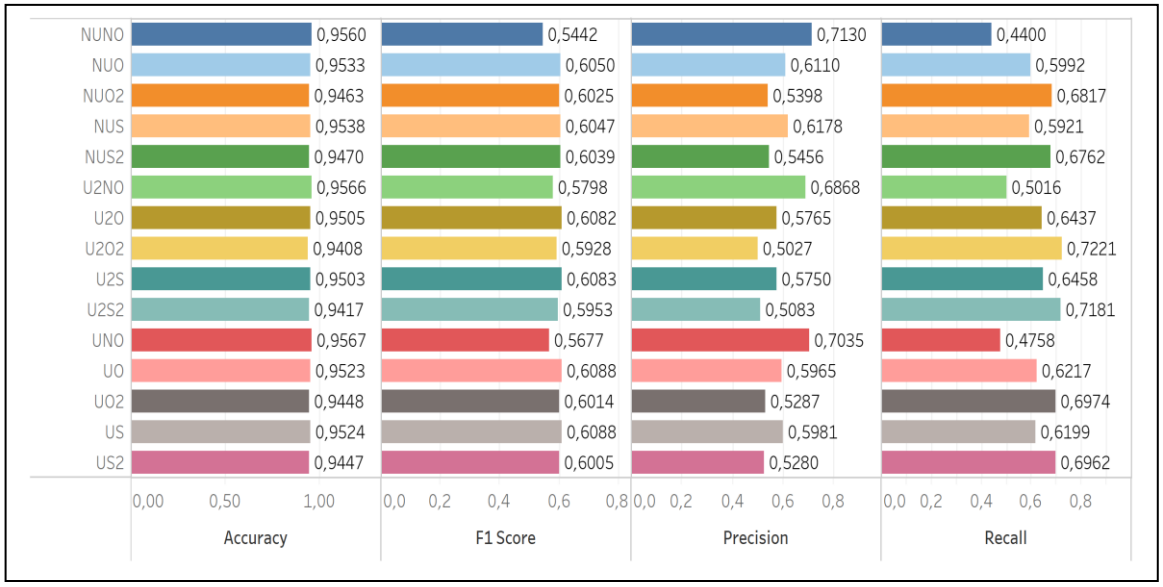
verildiği üzere, F_1-score bağlamında ccFraud veri kümesi için en etkili örnekleme stratejisi 0,6088 ile UO ve US olmuştur. Dolayısıyla, örnek hacmi fazla olan bu veri kümesinde daha başarılı bir sınıflandırma gerçekleştirmek, çoğunluk sınıfından veri azaltma ile azınlık sınıfından veri artırma beraber uygulandığında sağlanmaktadır. Bu veride sentetik veri üretme ile veriyi kopyalamanın yaklaşık olarak aynı etkiyi sağladığı görülmektedir.



Şekil 4.20. ccFraud veri kümesinde sınıflandırıcı türüne göre başarı değerlendirmesi



Şekil 4.21. European veri kümesinde örnekleme stratejilerinin etkisi



Şekil 4.22. ccFraud veri kümesinde örnekleme stratejilerinin etkisi

F_1-score üzerindeki iyileşmelere bakıldığında, ccFraud veri kümesinde yaklaşık 0,07 iken European veri kümesinde yaklaşık 0,03 olduğu görülmektedir. Bu durum, DCRFD yönteminin veri kümesinin boyutu arttıkça daha iyi çalıştığının göstergesidir. DCRFD, yüksek hacimli veride, sınıfların yakınlaştığı sınır noktalarında daha iyi örneklemenin yapılmasını sağlamaktadır.

DCRFD; sınıflandırıcının, örnek sayısı diğer sınıfın örnek sayısından oldukça fazla olan sınıfa doğru geliştirdiği önyargıya karşı koymak ve bu önyargıyı yeterince temsil edilmeyen sınıfa yönlendirmek amacıyla geliştirilmiştir. DCRFD deneyleri senaryolar dahilinde gerçekleştirilirken, çeşitli kısıtlamalar ve iyileştirici koşullar ortaya çıkmıştır. Bu durum aşağıdaki gibi özetlenebilir.

1. Veriler Arası Mesafe Hesaplanması: Yüksek boyutlu veri; kümeleme yapmak ve komşuluk bulmak için kullanılan mesafe fonksiyonunda değişiklik yapılmasını gerektirir.
2. Küme Sayısı Seçimi: k sayısı tüm küçük bölenleri tespit edebilecek kadar büyük olmalıdır. Bununla birlikte, büyük k değerlerinde ortaya çıkan düşük SSE'ler, her sınıf elemanının bir kümeye yerleşmesinden kaynaklanabilir. Diğer bir konu, bölütlemenin küme sayısına göre gerçekleştirilmesinden kaynaklanmaktadır. Bir küme en fazla bir

bölüte uyacak şekilde boyutlandırılmalıdır, böylece komşuluklar başka bir bölüte yerleşmez.

3. Bölüt Sayısı Seçimi: Her bölütteki iş yükü yaklaşık olarak eşit olmalıdır. Optimizasyon için, küçük kümeler aynı bölüte yerleştirilebilir.
4. Aykırı Değer Tespiti: İdeal kümelemede bile tek elemanlı kümeler oluşabilir. Bu durum aykırı değer olarak kabul edilerek bu kümeler filtrelenmiş ve yeniden örneklemeye dahil edilmemiştir.
5. Hafıza Aşımı: Çoğu hesaplamanın doğası gereği, analizler CPU veya hafıza tarafından engellenebilir. Karmaşık parametre uzayı ve parametreler arası etkileşimim ayarlamak zor olmasına rağmen, büyük miktardaki veriyi verimli bir şekilde işlemek gerekmektedir.
6. IR Büyüklüğü: Sınıflardaki örneklem sayısı birbirine yakın olduğu durumlarda, yeniden örnekleme yapmak çoğunluk sınıfı ile azınlık sınıfının yer değiştirmesine sebep olabilir. Aşırı artırma ve azaltmadan kaçınmak için yöntem yüksek IR değerine sahip veri kümeleri üzerinde uygulanmalıdır.

Geliştirilen stratejiler ile precision ve recall arasında iyi bir denge kurularak F_1-score değerlerine iyileştirme elde edilmiştir. Dolandırıcılık bakış açısından bakıldığında, recall ile dolandırıcı olanları doğru tespit etme oranında ve precision ile dolandırıcı olarak tespit edilenlerin gerçekten dolandırıcı olanlarının tespiti oranında iyileştirme gerçekleştirilmiştir. Böylelikle, sınıf dengesizliğine rağmen sınıflandırıcının dolandırıcı etiketli örnekleri göz ardı etmeden çalışması sağlanmıştır.

"Bedava öğle yemeği yok" teorisinde de belirtildiği üzere, genel amaçlı evrensel bir optimizasyon stratejisi imkansızdır, yani hiçbir algoritma diğerlerinden daha iyi değildir [298]. Bir stratejinin diğerinden daha iyi performans göstermesinin tek yolu, çözümün problemin yapısına göre özelleştirilmiş olmasıdır. Bu yöntem için gerçekleştirilen deneysel sonuçlarda da görüldüğü üzere, problemin çözümü veri kümesine bağlıdır. Büyük veri kümeleri üzerinde gerçekleştirilen yeniden örnekleme stratejileri, daha iyi performans üretilmesini sağlamıştır.

5. DERİN ÖĞRENME TABANLI ÖNERİLEN YÖNTEM

Bu bölümde dolandırıcılık tespiti problemi farklı bir perspektifle ele alınarak derin öğrenme tabanlı bir yöntem önerilmiştir. Bu yöntem, Görüntüye Dönüştürme ile Dolandırıcılık Tespiti (Fraud Detection with Image Conversion, FDIC) olarak adlandırılmıştır. FDIC, özniteliklerin birbirleri ile olan ilişkilerinin iki boyutlu olarak ifade edildiği görüntülere dönüştürülerek derin sinir ağları ile sınıflandırılmasını ve klasik yaklaşımlarla fark edilemeyen dolandırıcılık örüntülerinin ortaya çıkarılmasını amaçlamaktadır.

Önerilen yöntem, NVIDIA TESLA V100 Tensor Core 32GB GPU üzerinde gerçekleştirilmiştir. Analizlerde, sayısal hesaplama ve büyük ölçekli makine öğrenimi için açık kaynaklı bir yazılım kütüphanesi olan TensorFlow kullanılmıştır. Elde edilen sonuçlar ilerleyen alt bölümlerde detaylıca sunulmuştur.

5.1. Görüntüye Dönüştürme ile Dolandırıcılık Tespiti

Her kart sahibi için harcama davranışında her zaman güçlü bir periyodik düzen vardır ve bireysel kart sahibine ait geçmiş işlemlerin analiz edilmesiyle işlemlerdeki anormalliklerin tespiti mümkün kılınabilir. Bu gerçeğe dayanarak, dolandırıcılık tespiti bir zaman serisi problemi olarak ele alınabilir [299].

Zaman serisi, zaman içindeki sıralı ölçümlerden elde edilen verilerin koleksiyonunu temsil etmektedir [300]. Bu ölçümler zaman içinde sürekli olarak yapılabilir veya ayrık bir zaman noktasında alınabilir. Kesintisiz olarak kaydedilebilen verilere sahip seriler sürekli zaman serisi, düzenli veya düzensiz aralıklarda elde edilebilen verilere sahip seriler ise kesikli zaman serisi olarak adlandırılmaktadır [301].

Zaman serileri; genel eğilimler, mevsimsel hareketler, döngüsel hareketler ve düzensiz dalgalanmalar olmak üzere dört bileşenden oluşmaktadır [302]. Genel eğilim, uzun vadede zaman serisindeki artışı veya düşüşü göstermektedir. Mevsimsel hareketler, kısa vadeli döngülerin zaman serisinde meydana getirdiği değişikliklerdir. Döngüsel hareketler uzun sürelidir ve genellikle iş döngüsündeki periyodik olmayan dalgalanmalara karşılık gelir. Son olarak, düzensiz dalgalanmalar ise tekrarlanması muhtemel olmayan ani değişikliklerdir.

Herhangi bir zaman serisi analizinde ilk adım, gözlemlerin zamana karşı çizilmesi yani verinin zaman çizelgesinin oluşturulmasıdır. Zaman çizelgesi; genel eğilim, mevsimsel hareketler, aykırı değerler, yumuşak değişiklikler veya dönüm noktaları gibi önemli özellikleri göstererek, hem verinin tanımlanmasında hem de uygun modelin formüle edilmesinde önemli rol oynamaktadır [303].

Benzersiz yapıları nedeniyle zaman serileri, klasik veri madenciliği görevleri için zorluk teşkil etmektedir. Çok boyutluluk, yüksek özellik korelasyonu ve gürültü gibi sorunlarla başa çıkabilmek için analizlerde özel olarak zaman serisi madenciliği kullanılmaktadır. Zaman serisi madenciliği, gözlemlenen zaman serilerine ait bileşenlerin altında yatan nedenleri en iyi betimleyen modelleri tanımlamayı amaçlamaktadır ve aşağıda verildiği üzere, yedi başlık altında gerçekleştirilmektedir [304].

- Dizinleme: Verilen bir zaman serisi ve benzerlik ya da farklılık ölçüsü doğrultusunda benzer zaman serilerinin bulunması.
- Kümeleme: Zaman serisindeki doğal gruplaşmaların tespit edilmesi.
- Sınıflandırma: Etiketlenmemiş bir zaman serisinin, önceden tanımlanmış iki veya daha fazla sınıftan birine atanması.
- Tahmin: Bir zaman serisinin bir sonraki zaman aralığındaki değerinin öngörülmesi.
- Özetleme: Çok fazla veri noktası içeren bir zaman serisinin temel özelliklerinin korunarak küçültülmesi.
- Anomali Tespiti: Normal olduğu kabul edilen bir zaman serisi ve açıklanmamış bir zaman serisi verildiğinde, açıklanmamış üzerindeki beklenmedik oluşumları içeren bölümlerin bulunması.
- Bölütleme: Verilen bir zaman serisinin homojen bölümlere ayrılması.

Dijital bilgi kaynaklarının hızla artması, zaman serisi madenciliğinin çok çeşitli uygulama alanında yer almasına sebep olmuştur. Bazı örnekler arasında ekonomik tahmin, saldırı tespiti, süreç ve kalite kontrol, gen ekspresyon analizi ve tıbbi gözetim verilebilir [305]. Uygulama alanlarının ve veri hacminin artması, zaman serisi madenciliği algoritmaları için çeşitli zorluklara yol açmıştır. Bu zorluklar arasından; veri temsili, benzerlik ölçümü, dizinleme yöntemi ve görselleştirme olmak üzere üç ana sorun öne çıkmaktadır [306].

- Veri Temsili: Zaman serisinin asıl şekil özelliklerinin nasıl temsil edileceği ile ilgili sorundur. Bir temsil tekniği, temel özellikleri korurken verilerin boyutsallığını azaltan bir şekil kavramı türetmelidir.
- Benzerlik Ölçümü: Herhangi bir zaman serisi çiftinin nasıl ayırt edileceği veya eşleştirileceği problemidir. İki seri arasındaki mesafenin algısal ölçütlere dayanması gerekmektedir.
- Dizinleme Yöntemi: Hızlı sorgulamayı sağlamak için zaman serisinin nasıl düzenlenmesi ile ilgili sorundur. Dizinleme yöntemi, minimum alan tüketimi ve hesaplama karmaşıklığı sağlamalıdır.

Çözüm yaklaşımları; özellik tabanlı, model tabanlı ve mesafe tabanlı olarak üç kategoriye ayrılabilir. Özellik tabanlı yaklaşımlar zaman serilerini özellik vektörlerine dönüştürür. Ayrık Fourier dönüşümü (Discrete Fourier Transform, DFT) [307], tekil değer ayrışımı (Singular Value Decomposition, SVD) [308] veya ayrık dalgacık dönüşümü (Discrete Wavelet Transform, DWT) [309] gibi teknikler sıklıkla kullanılmaktadır. Model tabanlı yaklaşımlarda, verileri daha iyi anlamak veya gelecekteki noktaları tahmin etmek için Otoregresif Hareketli Ortalama (AutoRegressive–moving-average, ARMA), Otoregresif Entegre Hareketli Ortalama (AutoRegressive Integrated Moving Average, ARIMA) ve Mevsimsel Otoregresif Entegre Hareketli Ortalama (Seasonal ARIMA, SARIMA); Çok değişkenli zaman serilerinin analizi için Vektör Oto Regresyonu (VAR) ve Yapısal Vektör Oto Regresyonu (Structural Vector autoregression, SVAR); davranışsal varsayımlar yapmak için Hata Düzeltme Modeli (Error Correction Model, ECM) ve Vektör Hata Düzeltme Modeli (Vector ECM, VECM); ve zaman serisi kümelemesi için Gizli Markov Modeli (Hidden Markov Models, HMM) popüler örneklerden bazılarıdır [310]. Son olarak, mesafe tabanlı yaklaşımlar Dinamik Zaman Atlama (Dynamic Time Warping, DTW) veya Minimum Açıklama Uzunluğu (Minimum Description Length, MDL) gibi tekniklerle küresel, yerel veya gömülü özellikler ile zaman serileri arasında benzerlik bulmayı amaçlamaktadır [311].

Zaman serisi madenciliğindeki bir diğer sorun ise veri temsilidir. Zaman serisinin sahip olduğu; yüksek boyutluluk, aşırı gürültü varlığı, veri elemanlarının doğrusal olmayan ilişkisi ve verilerin sık güncellenmesi gibi özelliklerin de etkisiyle veri temsili sorunu giderek karmaşık bir hal almaktadır [312]. Bu nedenle, herhangi bir veri temsil yöntemi, orijinal verinin önemli özelliklerini korurken rastgele gürültülere karşı gürbüz olmayı ve veriyi

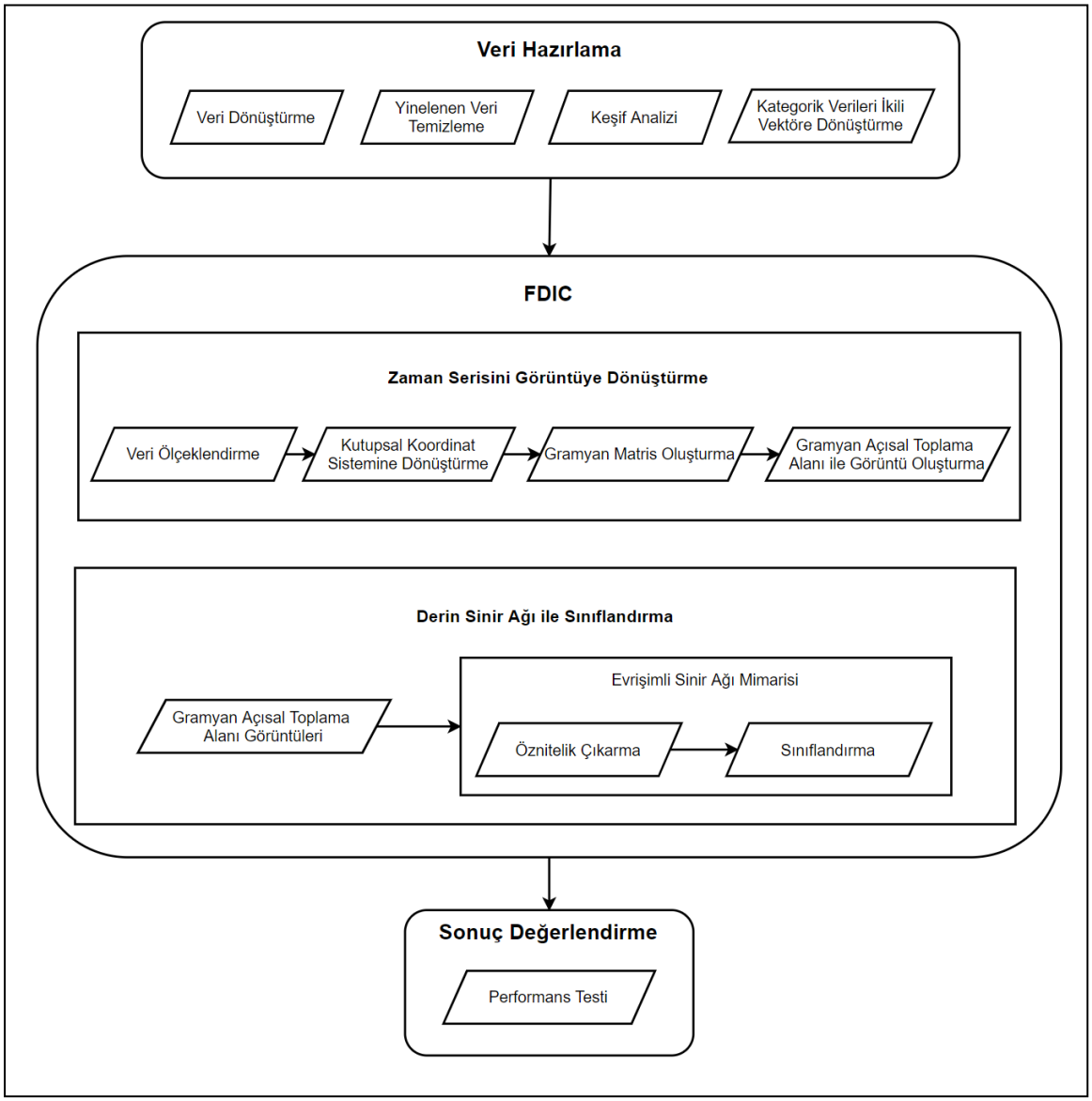
yönetilebilir bir boyuta getirmesi gerekmektedir. Zaman serilerinin başka bir alana dönüştürülmesine yönelik yaklaşımlar; bölütleme, yakınlaştırma, boyut küçültme veya indeksleme kapsamında gerçekleştirilmektedir. Parçalı Kümeleşme Yaklaşımı (Piecewise Aggregate Approximation, PAA), Sembolik Kümeleşme Yaklaşımı (Symbolic Aggregate Approximation, SAX), Sembolik Fourier Yaklaşım Sembolleri Çantası (Bag-of-SFA-Symbols, BOSS), Rastgele Evrişimli Çekirdek Dönüşümü (RandOm Convolutional KErnel Transform, ROCKET) veya Zaman Serisi Sınıflandırması için Kelime Çıkarma (Word ExtrAction for time SEries cLassification, WEASEL) dönüştürme tekniklerinin bazı örnekleridir [313].

Son yıllarda, zaman serilerini görüntüye dönüştüren ve daha sonra derin sinir ağı mimarileri kullanarak sınıflandıran çalışmalar önem kazanmaya başlamıştır. Çizelge 5.1'de özetlendiği üzere, Gramyan Açısal Fark Alanları (Gramian Angular Difference Fields, GADF), Gramyan Açısal Toplama Alanları (Gramian Angular Summation Fields, GASF), Markov Geçiş Alanları (Markov Transition Fields, MTF) ve Yineleme Grafiği (Recurrence Plot, RP) sinyal dönüştürme tekniği olarak kullanıldıktan sonra elde edilen görüntüler çoğunlukla CNN modelleri kullanılarak sınıflandırılmıştır.

Çizelge 5.1. Zaman serilerini görüntüye dönüştürüp sınıflandıran çalışmalar

	Amaç	Dönüştürme Tekniği				Derin Öğrenme Yaklaşımı
		GADF	GASF	MTF	RP	
[314]	Çevrimiçi dolandırıcılık davranışlarını tahmin etme	-	-	+	-	RNN ve CNN kombinasyonu
[315]	Tek değişkenli ve çok değişkenli zaman serilerini sınıflandırma	+	-	-	-	DNN
[316]	Elektrokardiyogram sinyal sınıflandırması	-	+	-	-	FFNN, CNN
[317]	Asenkron motorlarda arıza tespiti	-	-	-	+	CNN
[318]	Kalabalık kaynaklı hizmetlerin arz-talep boşluğunu tahmin etme	+	+	-	+	ResNet
[319]	Yarı iletken anormalliklerinin sınıflandırılması	+	+	-	-	CNN
[320]	Giyilebilir sensör tabanlı insan etkinliği tanıma	+	+	-	-	ResNet
[321]	Gün öncesi güneş ışınımının tahmini	-	+	-	-	Evrişimli LSTM
[322]	Karayolu trafik olayları tespiti	+	-	-	-	CNN
[323]	Sensör sınıflandırması	+	+	+	-	CNN
[324]	Konut seviyesinde enerji yükü tahminleme	+	+	+	+	CNN

Bu bağlamda, geleneksel dolandırıcılık tespiti yaklaşımlarının sınırlandırmalarını ortadan kaldırarak ve zaman serisi analizi için farklı bir yaklaşımda bulunarak gizli örüntüleri ortaya çıkarmayı amaçlayan, Görüntüye Dönüştürme ile Dolandırıcılık Tespiti (Fraud Detection with Image Conversion, FDIC) adı verilen bir yöntem önerilmiştir. Şekil 5.1'de metodolojisi verildiği üzere, kredi kartı işlemleri zaman serisi olarak kabul edilmiş [314] ve özellikler arası ikili ilişkilerin çıkarılması için Gramyan Açısal Alanı tekniği kullanılarak işlemler görüntülere dönüştürülmüştür [325]. Daha sonra, bir evrişimli sinir ağı mimarisi tasarlanarak bu görüntüler dolandırıcı veya meşru olarak sınıflandırılmıştır. Elde edilen sonuçlar hem performans açısından literatürdeki benzer çalışmalarla karşılaştırmalı olarak değerlendirilmiştir.



Şekil 5.1. FDIC yönteminin metodolojisi

5.1.1. Veri kümeleri

FDIC yönteminin performans değerlendirmesi sürecinde, önceki bölümde de analiz edilen ve kredi kartı dolandırıcılıklarını barındıran ccFraud ve European veri kümeleri kullanılmıştır. Fakat yöntemin işleyiş akışı içerisinde bu verideki öznitelikler iki boyutlu görüntülere dönüştürülmüştür. 28 özelliği bulunan European veri kümesinden elde edilen 28x28 boyutundaki görüntüler kullanılmıştır. 7 özelliği bulunan ccFraud veri kümesi ise, kategorik özellikler barındırdığı için bu özellikler ikili vektörlere dönüştürülerek nümerik hale getirilmiş ve yeni oluşturulan özellikler sonucunda 58x58 boyutundaki görüntüler kullanılmıştır.

Çizelge 5.2. Veri kümelerine ait özet bilgiler

	Eğitim Kümesi	Test Kümesi	Toplam Kayıt	Görüntü Boyutu
European	226980	56746	283726	28x28
ccFraud (örneklem alınmış)	240000	60000	300000	58x58

Yeni ccFraud veri kümesinin boyutunun atması analiz sürecinin zorlaşmasına sebep olmuştur. Bu nedenle, konsept olarak analizin gerçekleştirilmesi ve alternatif sonuçların değerlendirilmesi için bu veri kümesinden sınıf dengesizlik oranı aynı kalacak şekilde örneklem alınmıştır. Her iki veri kümesi için örneklerin %70'i eğitim kümesi ve geri kalanı test kümesi olarak ayırılmıştır. Çizelge 5.2'de veri kümelerine ait bilgiler özetlenmiştir.

5.1.2. Zaman serisini görüntüye dönüştürme

Zaman serisini görüntüye dönüştüren ilk yöntem olan Gramyan Açısal Alanları, zaman serilerini kutupsal koordinat sisteminde temsil eden ve bu açıları simetri matrisi olarak ifade eden bir zaman serisi dönüştürme yöntemidir. Gramyan Açısal Toplama Alanları, açıların toplamının kosinüsünü kullanan bir GAF türüdür.

Zaman serisi sıralı gerçek değerler kümesidir. *X* olarak tanımlanan bir zaman serisi, *n* sayıda değerlerden oluşan $X = \{ x_1, x_2, \dots x_n \}$ ile gösterilmektedir. *X*, Eş. 5.4'de verildiği üzere [0,1] aralığında yeniden ölçeklendirildiğinde, $\tilde{X}$ değeri Eş. 5.5 ile açısal kosinüs olarak kodlanarak kutupsal koordinatlarda temsil edilir. Burada *r* yarıçapı, $\emptyset$ kutupsal koordinat açısını, t_i zaman damgasını ve *N* normalleştirme için kullanılan sabit bir faktörü işaret

etmektedir. Yeniden ölçekleme ve zaman serilerinin kutupsal koordinat sistemine dönüştürülmesinden sonra, zaman serisi trigonometrik toplam veya fark yoluyla yeniden tanımlanabilmektedir. I bir birim satır vektörü olmak üzere, GASF Eş. 5.6 ve Eş. 5.7 doğrultusunda elde edilmektedir [326].

$$\tilde{x}_0^i = \frac{x_i - \min(X)}{\max(X) - \min(X)} \tag{5.4}$$

$$\begin{cases} \emptyset = \arccos(\tilde{x}_i), -1 \leq \tilde{x}_i \leq 1, \tilde{x}_i \in \tilde{X} \\ r = \frac{t_i}{N}, t_i \in \mathbb{N} \end{cases} \tag{5.5}$$

$$GASF = \begin{bmatrix} \cos(\emptyset_1 + \emptyset_1) & \cdots & \cos(\emptyset_1 + \emptyset_n) \\ \vdots & \ddots & \vdots \\ \cos(\emptyset_n + \emptyset_1) & \cdots & \cos(\emptyset_n + \emptyset_n) \end{bmatrix} \tag{5.6}$$

$$= \tilde{X}' . \tilde{X} - \sqrt{I - \tilde{X}^2}' . \sqrt{I - \tilde{X}^2} \tag{5.7}$$

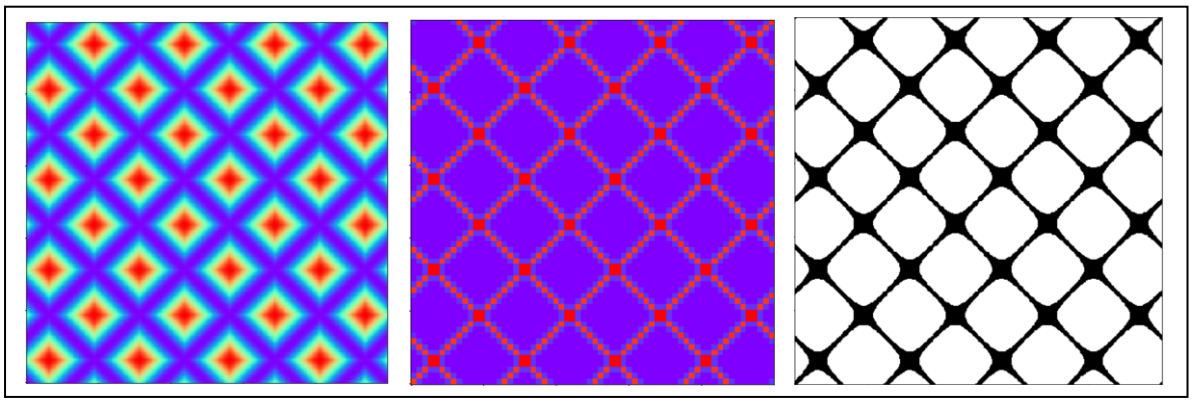
Zaman serisini görüntüye dönüştürme yöntemlerinden bir diğeri olan Markov Geçiş Alanı, ayrıklaştırılmış bir zaman serisi için bir geçiş olasılıkları alanını temsil eder. Bir X zaman serisi verildiğinde, Q olarak adlandırılan dağılım bölmeleri oluşturulur ve her x_i, karşılık gelen q_j ($j \in [1, Q]$) bölmesine atanır. Zaman ekseni boyunca birinci dereceden Markov zinciri şeklinde dağılım bölmeleri arasındaki geçişleri sayıldığında, $Q \times Q$ ağırlıklı bitişiklik matrisi W oluşturulur. $w_{i,j}$, q_i'deki bir noktanın ardından q_j'deki bir noktanın gelmesinin frekansı ile elde edilir. Böylelikle, hem zaman serisinin dağılımından hem de zaman adımlarındaki zamansal bağımlılıktan etkilenmeyen ve zaman serilerinin çok açıklıklı geçiş olasılıklarını kodlayan MTF, Eş. 5.8 doğrultusunda oluşturulur [325].

$$MTF = \begin{bmatrix} w_{ij|x_1 \in q_i, x_1 \in q_j} & \cdots & w_{ij|x_1 \in q_i, x_n \in q_j} \\ \vdots & \ddots & \vdots \\ w_{ij|x_n \in q_i, x_1 \in q_j} & \cdots & w_{ij|x_n \in q_i, x_n \in q_j} \end{bmatrix} \tag{5.8}$$

Zaman serisini görüntüye dönüştüren sonuncu yöntem olan Yineleme Grafiği ise, orijinal zaman serisinden çıkarılan yörüngeler arasındaki mesafeleri temsil eder. Kısaca, iki nokta arasındaki mesafe belirlenen eşik değerinden az ise bu nokta yineleme olarak tanımlanır ve görüntüde siyah bir piksel olarak ifade edilir. Aksi durumda ise, beyaz renkli piksel ile

gösterim yapılır. Bir X zaman serisi verildiğinde; $\vec{x}_i$ bir yörünge, N bu yörüngedeki ölçülen noktaların sayısı, ε bir eşik mesafesi ve θ Heaviside fonksiyonu ($\theta(x) = 0,\ if\ x < 0\ ve\ \theta(x) = 1,\ diğer$) olmak üzere, RP Eş. 5.8 doğrultusunda elde edilir [317].

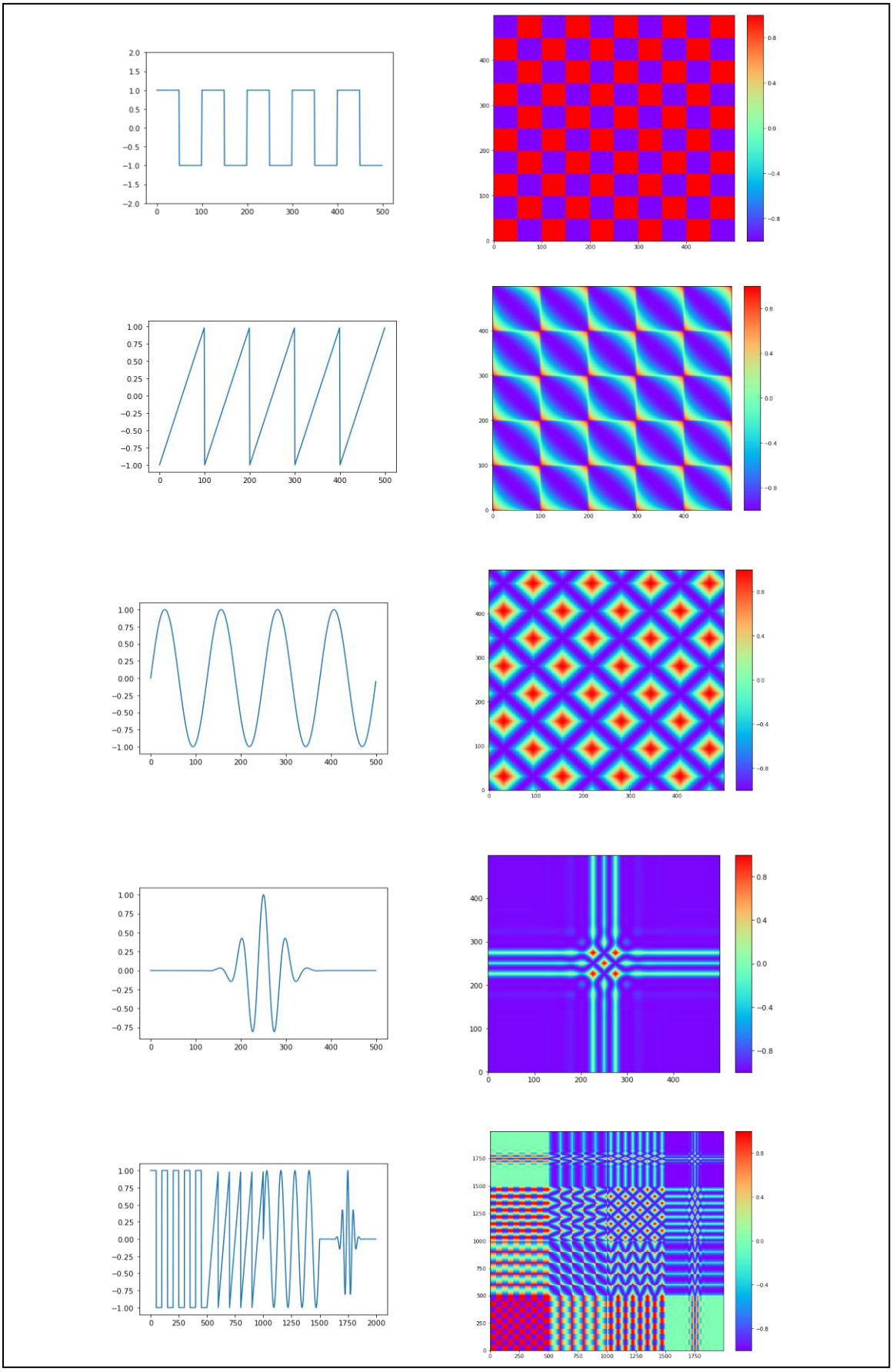
$$R_{i,j} = \theta\left(\varepsilon - \left\|\vec{x}_i - \vec{x}_j\right\|\right), \qquad i,j = 1,\dots,N \tag{5.9}$$



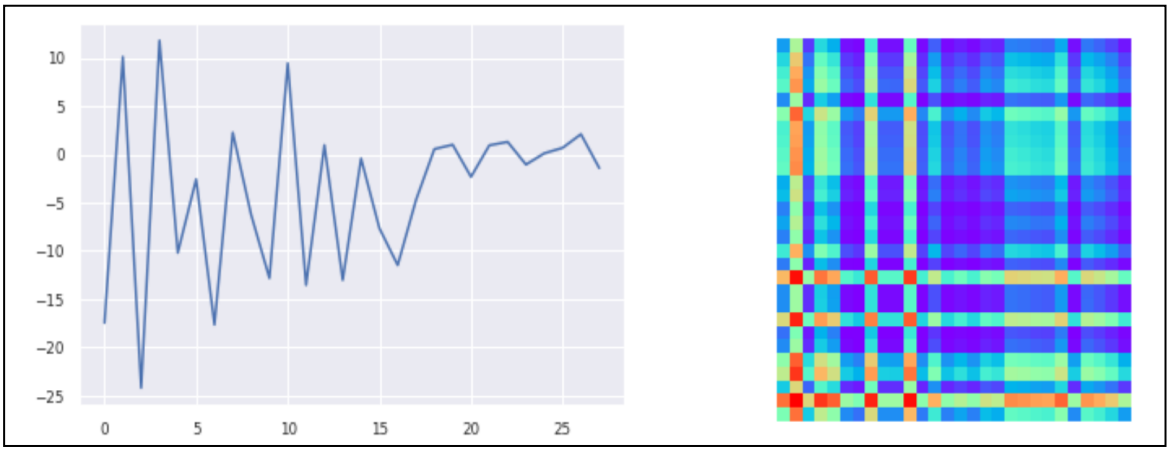
Şekil 5.2. Örnek bir sinüs sinyalinden elde edilen sırasıyla GASF, MTF ve RP görüntüleri

Zaman serisini görüntüye dönüştüren yöntemlerin farklılıklarını ortaya koyabilmek adına, Şekil 5.2'de 500 birimlik örnek bir sinüs sinyali için oluşturulan 500x500 boyutlarındaki GASF, MTF ve RP görüntüleri verilmiştir.

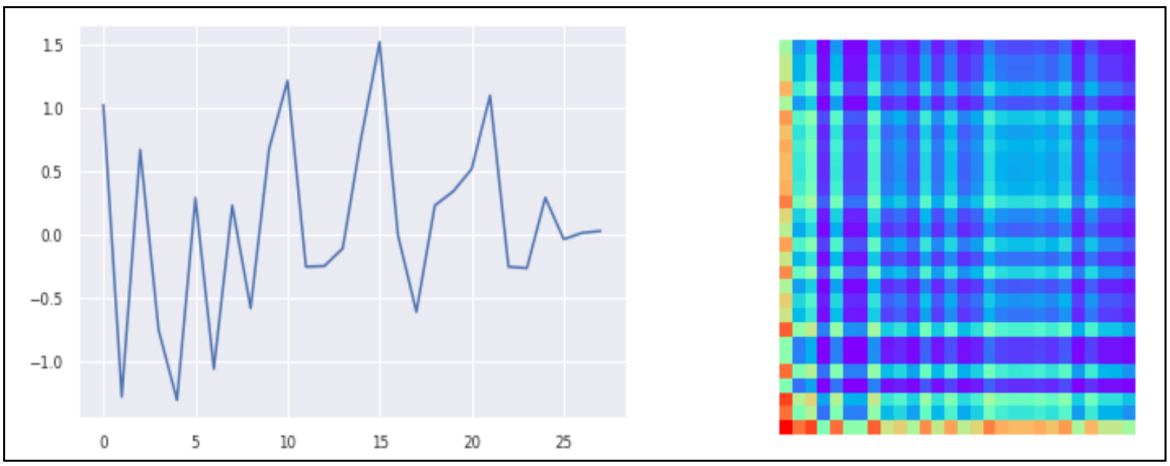
Zaman sinyalinden elde edilen görüntünün okunmasını kolaylaştırmak ve bir referans noktası oluşturmak adına, Şekil 5.3'de sırasıyla; kare, testere dişi, sinüs, gauss vuruşu ve tüm sinyallerin birleşiminden elde edilen GASF görüntüleri verilmiştir. GASF görüntülerinin ifade ettiği ilişkileri netleştirmek açısından, kare sinyal dikkatli bir şekilde incelendiğinde değer aralığının [-1,1] aralığında değiştiği görülmektedir. Üretilen 500 birimlik kare sinyal için 500x500 boyutlarında bir GASF görüntüsü oluşmaktadır. GASF görseline bakıldığında 0 ile 50 arasındaki sinyallerinin birbirleri ile olan ilişkisi grafiğin sol alt köşesindeki 50x50 boyutlarındaki alana denk gelmektedir ve bu alandaki bütün değerler bir olarak yani kırmızı renkte gözlemlenmektedir. Bu durum, bu iki sinyalin kutupsal koordinatlarında oluşan açının arccos değerlerinin toplamının sıfır ya da 2π olmasından kaynaklanmaktadır, çünkü bu durumda toplamın cos değeri bir olacaktır. Tersi bir örnekte ise 0 ile 50 arasındaki sinyallerinin 50 ile 100 arasındaki sinyalleri ile olan ilişkisinde bütün alan eksi bir değerinden yani mavi renkten oluşmaktadır.



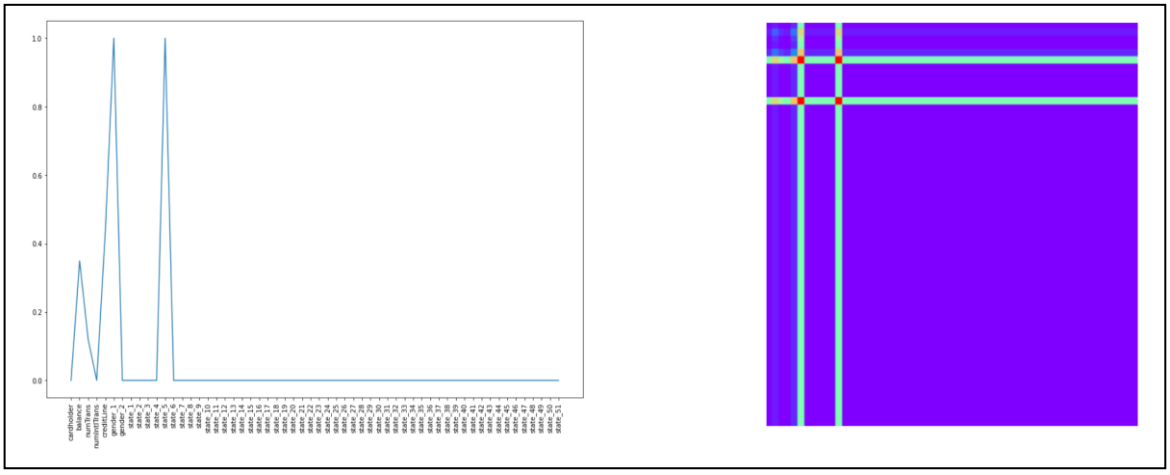
Şekil 5.3. Çeşitli sinyallerden elde edilen GASF görüntü örnekleri



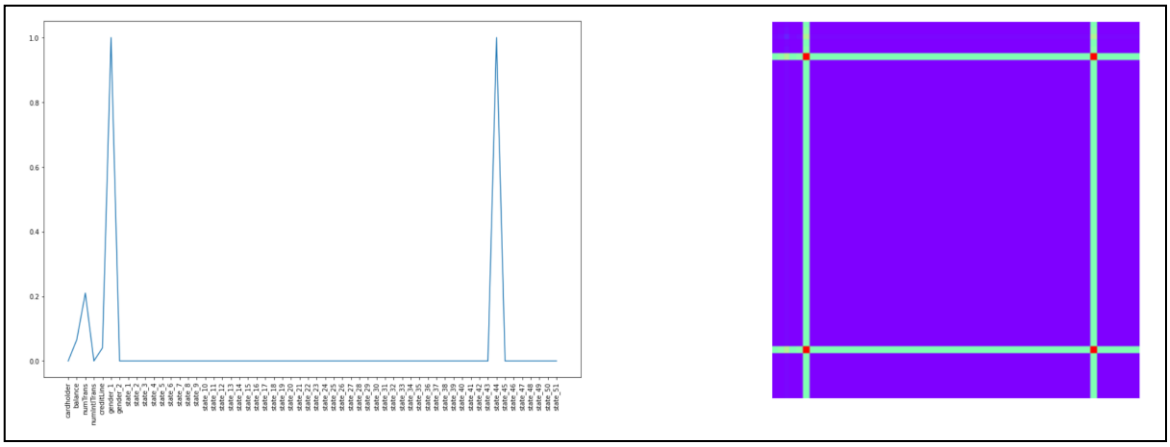
Şekil 5.4. European veri kümesindeki dolandırıcı olduğu bilinen bir sinyal ve GASF görüntüsü



Şekil 5.5. European veri kümesindeki meşru olduğu bilinen bir sinyal ve GASF görüntüsü



Şekil 5.6. ccFraud veri kümesindeki dolandırıcı olduğu bilinen bir sinyal ve GASF görüntüsü



Şekil 5.7. ccFraud veri kümesindeki meşru olduğu bilinen bir sinyal ve GASF görüntüsü

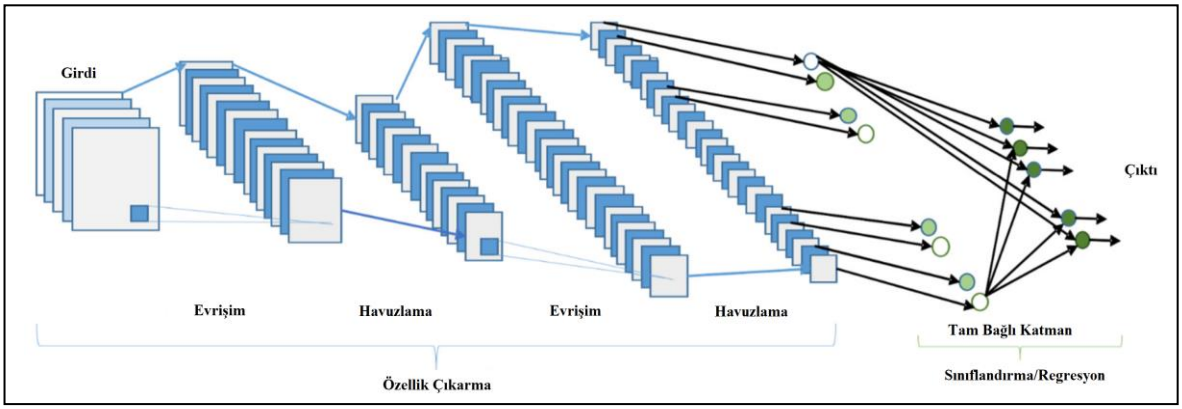
Analizler için kullanılan ccFraud ve European veri kümelerine ait sinyaller rastgele örnek alınarak değerlendirildiğinde, hem zamansal bağımlılığın korunduğu hem de değişkenler arasındaki bütün ikili ilişkilerinin ortaya çıkarıldığı görülmektedir. European veri kümesinden alınan, dolandırıcı olduğu bilinen bir kayıt Şekil 5.4'de ve meşru olduğu bilinen bir kayıt Şekil 5.5'de sinyal ve GASF görüntüsü olarak sunulmuştur. Benzer şekilde, ccFraud veri kümesinden alınan, dolandırıcı olduğu bilinen bir kayıt Şekil 5.6'da ve meşru olduğu bilinen bir kayıt Şekil 5.7'de sinyal ve GAF görüntüsü olarak sunulmuştur. ccFraud veri kümesindeki kategorik özelliklerin ikili vektörlere dönüştürülmesinden dolayı, European veri kümesindeki görüntüler ccFraud veri kümesindeki görüntülerden daha çeşitli piksel değişimlerine sahiptir. Her iki veri kümesi için de zaman serisinin karakteristikleri; renkler, çizgiler veya noktalar gibi farklı özellikler ile iki boyutlu görüntülere aktarıldıktan sonra dolandırıcılık tespiti için sınıflandırma işlemi gerçekleştirilir.

5.1.3. Derin sinir ağı ile dolandırıcılık sınıflandırma

Zaman serisi sınıflandırması, veri noktalarını zaman içinde davranışlarına göre sınıflandırmakla ilgilidir. n gözlemden oluşan X zaman serisi ve Y ayrık sınıf etiketi ikililerinden (X_n, Y_n) oluşan $D = \{ (X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n) \}$ veri kümesi üzerinde zaman serisi sınıflandırıcısı, olası girdilerden sınıf değişken değerleri üzerindeki olasılık dağılımına kadar olan bir eşleştirmedir [327].

Derin öğrenme çözümleri; eksik değerler ve gürültü ile başa çıkma, çok değişkenli girdilere uyumlu olma, doğrusal olmayan fonksiyonları ele alma ve çok adımlı tahminler yapma

yetenekleri nedeniyle zaman serisi analizi için önem kazanmaktadır [328]. Zaman serilerinin sınıflandırması için derin öğrenme yaklaşımları üretken ve ayrımcı olmak üzere iki kategoriye ayrılmaktadır [329]. Üretken yaklaşımlar eğitim aşamasından önce zaman serilerinin iyi bir temsilini bulmayı amaçlarken, ayrımcı yaklaşımlar doğrudan ham zaman serilerinden veya çeşitli dönüşümler sonucunda elle tasarlanmış özelliklerden öğrenmektedir.



Şekil 5.8. Evrişimli sinir ağının genel mimarisi [330]

CNN, çok sayıda tabakanın oldukça etkili bir şekilde eğitildiği bilgisayarla görme uygulamalarında en çok tercih edilen derin öğrenme modellerinden biridir [330]. Şekil 5.8'de örneklendirildiği üzere, temel bir CNN genel olarak; evrişim katmanı, havuzlama (alt örnekleme) katmanı ve tam bağlı katman olmak üzere üç ana nöral katman tipinden oluşur [331]. Evrişim katmanı, farklı özellik haritalarını hesaplamak için kullanılan evrişim çekirdeklerinden oluşur. Havuzlama katmanı, özellik haritalarının ve ağ parametrelerinin boyutlarını azaltmak için kullanılır. Tam bağlı katman ise, geleneksel bir sinir ağı gibi davranır ve önceki katmanlar tarafından çıkarılan özelliklere dayanarak sınıflandırma veya regresyon uygular ya da başka ağlara girdi olarak kullanılır. CNN çoğunlukla soruna göre özelleştirilmiş belirli bileşenler grubundan oluşur ve bu bileşenler bahsi geçen tüm katmanları içermeyebilir.

CNN katmanları arasındaki işletim sürecini matematiksel olarak ifade etmek gerekirse [332]; ağırlık paylaşımını sağlamak ve modelin karmaşıklığı azaltmak için, l katmanında bulunan k özellik haritasının i,j konumundaki özellik değeri olan $z_{i,j,k}^{l}$, Eş. 5.10'da verildiği

şekilde hesaplanmaktadır. Burada w_k^l ağırlık vektörünü, $x_{i,j}^l$ giriş değerini ve b_k^l önyargı terimini işaret etmektedir.

$$z_{i,j,k}^l = {w_k^l}^T x_{i,j}^l + b_k^l \tag{5.10}$$

Aktivasyon fonksiyonu CNN'ye doğrusal olmayan özellikleri tespit etme yeteneği vermektedir. $z_{i,j,k}^l$ evrişimsel özelliğin aktivasyon değeri olan $a_{i,j,k}^l$ Eş. 5.11 ile hesaplanmaktadır.

$$a_{i,j,k}^l = a(\, z_{i,j,k}^l \,) \tag{5.11}$$

Havuzlama Eş. 5.12 ile hesaplanmaktadır. Burada y_{ij} yerel bir komşuluktur ve $\forall\, (m,n) \in y_{ij}$.

$$y_{i,j,k}^l = p(\, a_{m,n,k}^l \,) \tag{5.12}$$

CNN eğitimi, kayıp fonksiyonunu en aza indirerek en uygun parametre kümesinin elde edildiği küresel bir optimizasyon problemidir. *θ,* tüm CNN parametreleri, $x^{(n)}$ n. giriş, $y^{(n)}$ onun etiketi ve $o^{(n)}$ CNN çıktısı olmak üzere bir *N* giriş-çıkış ilişkisi $\{(x^{(n)}, y^{(n)}); n \in [1, \dots, N]\}$ için kayıp fonksiyonu Eş. 5.13 ile hesaplanmaktadır [333].

$$L = \frac{1}{N} \sum_{n=1}^{N} \ell\, (\theta;\, y^{(n)}, o^{(n)}) \tag{5.13}$$

5.1.4. Deneysel sonuçlar

Deneysel çalışmalar sırasında öncelikle, dolandırıcılık tespiti için zaman serisini görüntüye dönüştüren en uygun teknik belirlenmiştir. GASF, MTF ve RP ile elde edilen görüntüler, temel bir CNN mimarisi kullanılarak sınıflandırılmıştır. Çizelge 5.3'de verildiği üzere, European veri kümesi için en iyi sonuç GASF görüntüleri kullanıldığında elde edilmiştir.

GASF görüntülerinin dolandırıcılık tespitinde daha etkili olduğu görüldükten sonra, bu görüntüler üzerinde en iyi sınıflandırmayı yapabilecek CNN mimarisi belirlenmiştir. Bu

süreçte pek çok farklı katman yapısının sınıflandırma başarısı karşılaştırılmıştır. Çizelge 5.4'de üç farklı modelin sonuçları verilmiştir. En iyi sonucu üreten iki numaralı model üzerinde ise çeşitli parametre ayarlamaları sonucunda elde edilen başarı değerleri karşılaştırılmıştır. Çizelge 5.5'de verildiği üzere en iyi sonucu üreten bir numaralı model, sonraki analizlerde temel sınıflandırıcı olarak kullanılmıştır.

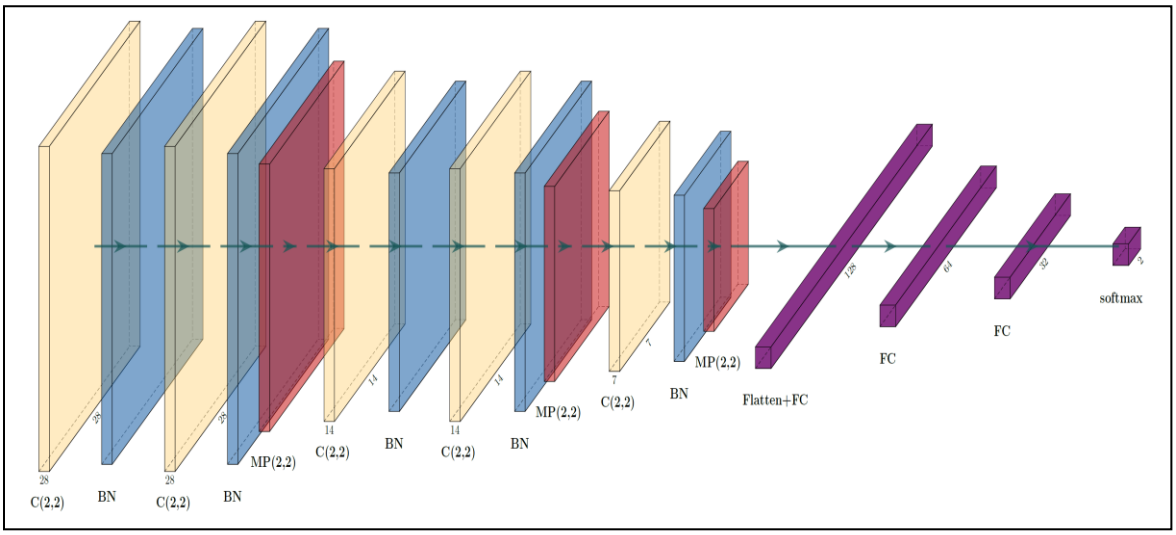
Çizelge 5.3. European veri kümesi için farklı görüntü dönüştürme tekniklerinin uygulanması

Zaman Serisini Görüntüye Dönüştürme Tekniği	ACC	Precision	Recall	F_1-score
GASF	0,9960	0,9294	0,8404	0,8827
MTF	0,9988	0,8181	0,3829	0,5217
RP	0,9992	0,8472	0,6489	0,7349

Çizelge 5.4. European veri kümesi için farklı CNN mimarilerinin uygulanması

Model	Mimari	ACC	Precision	Recall	F_1-score
1	C+BN+C+BN+MP+C+BN+C+BN+MP	0,9995	0,9499	0,7553	0,8402
2	C+BN+C+BN+MP+C+BN+C+BN+MP+C+BN+MP	0,9960	0,9294	0,8404	0,8827
3	C+BN+C+BN+MP+C+BN+C+BN+MP+C+BN+MP+C+BN	0,9995	0,9240	0,7766	0,8439

*C: Convolution, BN: Batch Normalization, MP: Max Pooling



Şekil 5.9. Analizlerde kullanılan CNN mimarisi

Şekil 5.9'da verildiği üzere analizlerde kullanılan CNN mimarisi; beş evrişim katmanı, beş toplu normalleştirme katmanı ve üç havuzlama katmanından oluşmaktadır. Evrişim katmanını girdi olarak alan havuzlama katmanında 2x2 çekirdek boyutunda maksimum havuzlama kullanılmıştır. Tensörü düz tabaka halinde yeniden şekillendirdikten sonra aşırı

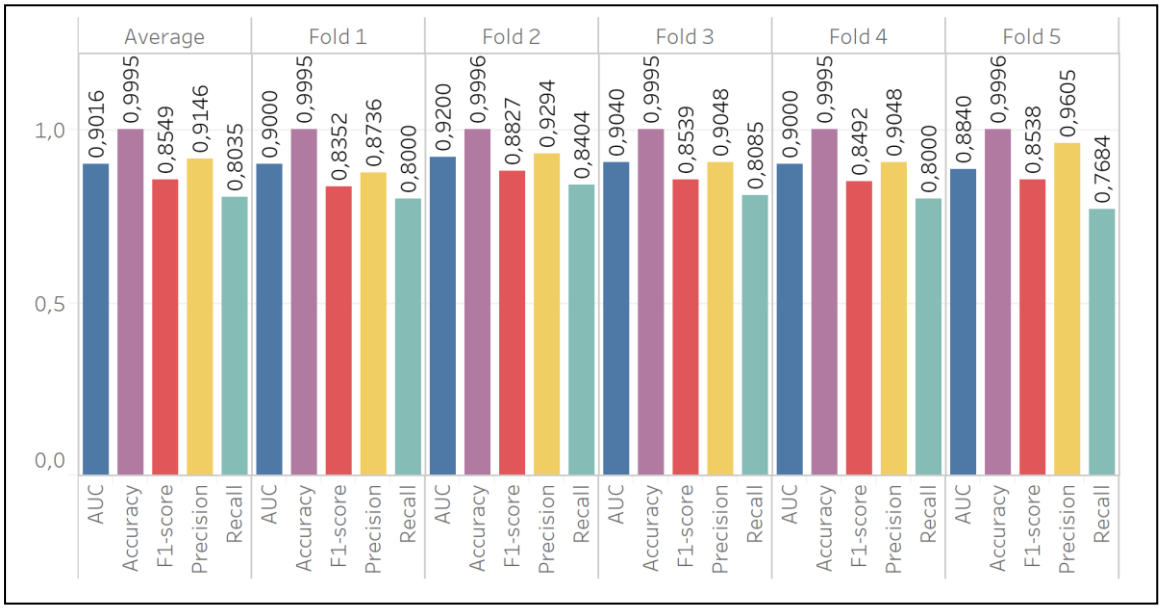
öğrenmeyi önlemek için 0,1 oranında bir düşüş uygulanmıştır. İlk üç tam bağlı katmanda, negatif girdiye izin vererek maliyeti sıfıra daha hızlı yakınsama eğiliminde olan ve *x* bir giriş vektörü olmak üzere Eş. 5.14'de $f(x)$ olarak verilen üstel doğrusal birim kullanılmıştır. Son katmanda ise, her hedef sınıfın olası tüm hedef sınıflar üzerindeki olasılıklarını hesaplayan ve Eş. 5.15'de σ olarak verilen softmax fonksiyonu kullanılmıştır. Son olarak, kayıp fonksiyonu, uyarlanabilir moment tahmini kullanılarak optimize edilen kategorik çapraz entropidir. Gerçek bir y etiketi ve tahmin edilen bir p etiketi verildiğinde, ikili sınıflar için kategorik çapraz entropi olan $L(y,p)$ Eş. 5.16'da verildiği gibi tanımlanmaktadır.

$$f(x) = \begin{cases} x, & x > 0 \\ \alpha(e^x - 1), & x \leq 0 \end{cases} \tag{5.14}$$

$$\sigma(x)_i = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_i}} \tag{5.15}$$

$$L(y,p) = -\sum_{i=1}^{2} y_i \log p_i \tag{5.16}$$

En uygun parametrelere sahip model belirlendikten sonra beş katlı çapraz doğrulama testi gerçekleştirilmiş ve değerlendirme bu katların ortalamasına göre yapılmıştır. Şekil 5.10'da verildiği üzere European veri kümesi için AUC, accuracy, F_1-score, precision ve recall değerleri sırasıyla %90,16, %99,95, %85,49, %91,46 ve %80,35 olarak hesaplanmıştır.



Şekil 5.10. European veri kümesindeki beş katlı doğrulama testi sonuçları

Çizelge 5.5. European veri kümesi için farklı parametre ayarlamaları

Model	Parametreler				ACC	Precision	Recall	F_1-score
	En İyileyici	Öğrenme Oranı	Çekirdek Boyutu	Yığın Boyutu				
1	Adam	1e-4	2x2	16	0,9960	0,9294	0,8404	0,8827
2	Adam	1e-4	2x2	32	0,9994	0,8750	0,7446	0,8046
3	Adam	1e-4	3x3	16	0,9995	0,9359	0,7766	0,8488
4	Adam	1e-4	3x3	32	0,9994	0,9102	0,7553	0,8255
5	Adam	1e-3	2x2	16	0,9994	0,9571	0,7127	0,8170
6	Adam	1e-3	2x2	32	0,9995	0,9487	0,7872	0,8604
7	Adam	1e-3	3x3	16	0,9995	0,9487	0,7872	0,8604
8	Adam	1e-3	3x3	32	0,9994	0,9571	0,7127	0,8170
9	SGD	1e-4	2x2	16	0,9995	0,9480	0,7766	0,8538
10	SGD	1e-4	2x2	32	0,9995	0,9473	0,7659	0,8470
11	SGD	1e-4	3x3	16	0,9996	0,9390	0,8191	0,8750
12	SGD	1e-4	3x3	32	0,9995	0,9487	0,7872	0,8604
13	SGD	1e-3	2x2	16	0,9995	0,9493	0,7978	0,8670
14	SGD	1e-3	2x2	32	0,9995	0,9480	0,7766	0,8538
15	SGD	1e-3	3x3	16	0,9995	0,9493	0,7978	0,8670
16	SGD	1e-3	3x3	32	0,9995	0,9493	0,7978	0,8670
17	RMSProp	1e-4	2x2	16	0,9995	0,9268	0,8085	0,8636
18	RMSProp	1e-4	2x2	32	0,9995	0,9493	0,7978	0,8670
19	RMSProp	1e-4	3x3	16	0,9993	0,8918	0,7021	0,7857
20	RMSProp	1e-4	3x3	32	0,9996	0,9620	0,8085	0,8786
21	RMSProp	1e-3	2x2	16	0,9994	0,9552	0,6808	0,7950
22	RMSProp	1e-3	2x2	32	0,9994	0,9701	0,6914	0,8074
23	RMSProp	1e-3	3x3	16	0,9995	0,9058	0,8191	0,8603
24	RMSProp	1e-3	3x3	32	0,9995	0,9375	0,7978	0,8620

*Adam: Adaptive Moment Estimation, SGD: Stochastic Gradient Descent, RMSProp: Root Mean Square Propagation

Elde edilen sonuçları değerlendirmek için, European ve ccFraud veri kümelerini kullanmış olan diğer çalışmalar ile önerilen yöntem karşılaştırılmıştır. Çizelge 5.6'da verildiği üzere, FDIC %85,9 ile en yüksek F_1-score üretmiştir. İlgili çalışmaların, uygulama detaylarında çeşitli farklılıklar olduğu görülmüştür. FDIC eğitim sürecinde, [334] ve [335] aksine, yanlılığı önlemek için tekrar eden kayıtlar çıkarılmıştır. Bir diğer farklılık ise, sınıf dengesizliği ile başa çıkmak için aşırı örnekleme [334] ile veya sentetik veri oluşturarak [271] azınlık sınıfı örneklerini arttırmayı amaçlayan çalışmaların aksine FDIC orijinal veri üzerinde herhangi bir örnek artışı veya azaltması gerçekleştirmemiştir. Sadece orijinal veriyi

kullanmasına rağmen FDIC, recall ve F_1-score metriklerinde daha iyi sonuçlar üretmiştir. FDIC'nin, dolandırıcı örneklerini sınıflandırmasındaki doğruluğu veren recall ve dolandırıcılık olarak sınıflandırılan durumlarda doğruluğu veren precision arasında iyi bir denge oluşturduğu görülmektedir. Bu durum da, FDIC'nin gerçek pozitifi yani yüksek maliyetle ilişkili olan dolandırıcı örnekleri tespit etmede daha iyi olduğunu göstermektedir.

Çizelge 5.6. European veri kümesi üzerinde analiz edilmiş çalışmaların karşılaştırılması

Kaynak	Yaklaşım	ACC	Precision	Recall	F_1-score
[334]	ROS+Linear SVM Regression	0,9801	0,1391	0,593	0,2253
[334]	ROS+Logistic Regression	0,9868	0,1897	0,528	0,2791
[334]	ROS+NN Based Classification	0,9794	0,1470	0,6598	0,2404
[334]	ROS+Non-Linear Auto-regression	0,9659	0,1000	0,7401	0,1762
[334]	ROS+Deep NN Autoencoders	0,9855	0,1972	0,5821	0,2947
[271]	GAN+DNN (N=630)	0,9996	0,9320	0,7328	0,8205
[271]	SMOTE+DNN (N=630)	0,9996	0,9680	0,6946	0,8088
[335]	Logistic Regression	0,9991	0,9300	0,5700	0,7100
[335]	Autoencoder (Threshold=5)	0,9870	0,0110	0,6400	0,1900
[336]	SMOTE+Balanced Bagging Ensemble	*	0,9412	0,7273	0,8205
[336]	SMOTE+RF	*	0,4324	0,7273	0,5424
[336]	SMOTE+Gaussian Naive Bayes	*	0,0786	0,5000	0,1358
Önerilen Yöntem	FDIC	0,9995	0,9146	0,8035	0,8549

*:Belirtilmemiş

Çizelge 5.7. ccFraud veri kümesi üzerinde analiz edilmiş çalışmaların karşılaştırılması

Kaynak	Yaklaşım	ACC	Precision	Recall	F_1-score
[336]	SMOTE+RF	*	0,3634	0,8500	0,5091
[336]	SMOTE+Gaussian Naive Bayes	*	0,3001	0,8588	0,4448
[336]	ROS+Balanced Bagging Ensemble	*	0,3913	0,6380	0,4851
[337]	Cost-Sensitive Cosine Similarity K-NN	0,9550	*	0,5100	0,5600
[337]	Cost Sensitive C5.0 DT	0,9330	*	0,6500	0,5300
[337]	Distance Based Cost Sensitive K-NN	0,9450	*	0,5300	0,5200
Önerilen Yöntem	FDIC	0,9522	0,6017	0,5878	0,5947

*:Belirtilmemiş

FDIC, aynı süreçler kullanılarak ccFraud veri kümesi üzerinde elde ettiği sonuçlar değerlendirilmiştir. Çizelge 5.7'de sunulan karşılaştırmada verildiği üzere, FDIC %59,47 ile F_1-score bağlamında en iyi sonucu üretmiştir. İlgili çalışmalar karşılaştırıldığında, uygulama detaylarında çeşitli farklılıklar olduğu görülmüştür. [336]'de veri kümesi, %90 eğitim ve

%10 test olarak ayrılmıştır. Bu durumun, nispeten daha üstün sonuçlar üretmesi beklenilmesine rağmen, FDIC F_1-score bağlamında daha iyi performans göstermiştir. [337]'de analizler veri kümesi dengesizlik oranını korunduğu 6000 örnek üzerinde gerçekleştirilmiştir. FDIC uygulamasından daha az sayıda örneğin kullanıldığı bu karşılaştırmada da, FDIC daha iyi performans göstermiştir.

Önerilen yöntem, dolandırıcılık tespiti için birden fazla özellik gerektiren ve her özelliğin hedef değişken üzerindeki ilişkisini modelleyen yöntemlere kıyasla tek bir görüntüyü bir girdi özelliği olarak ele alarak ve ikili özellikler arasındaki ilişkileri modelleyerek daha umut verici sonuçlar ortaya koymaktadır. Böylelikle, özelliklerin farklı bir ilişki uzayına taşınmasının, model performansını artırabileceği görülmektedir.

6. ÖNERİLEN YÖNTEMLERDE AÇIKLANABİLİR YAPAY ZEKÂ

Tez kapsamında geliştirilen üç dolandırıcılık tespiti yöntemi, bu bölümde açıklanabilir yapay zekâ bakış açısıyla ele alınmıştır. Önerilen yöntemlerin, tez akışında yer alma sırası ile açıklanabilir sistemlerin gelişim sırası eşzamanlılık göstermektedir. Bu bağlamda yöntemlerin açıklanabilirliğinde; Dinamik İmzalar İle Dağıtık Dolandırıcılık Tespiti için uzman görüşü alma ve kural çıkarma, Dağıtık Yeniden Örnekleme İle Dolandırıcılık Tespiti için kural çıkarmaya ek olarak özellik önemlerinin çıkarılması ve son olarak Görüntüye Dönüştürme İle Dolandırıcılık Tespiti için ısı haritalarının oluşturulması stratejileri kullanılmıştır. Ek olarak, önerilen son yöntemin barındırdığı görsel anlam zayıflığı sebebiyle, yeni bir açıklanabilirlik yaklaşımı önerilmiştir.

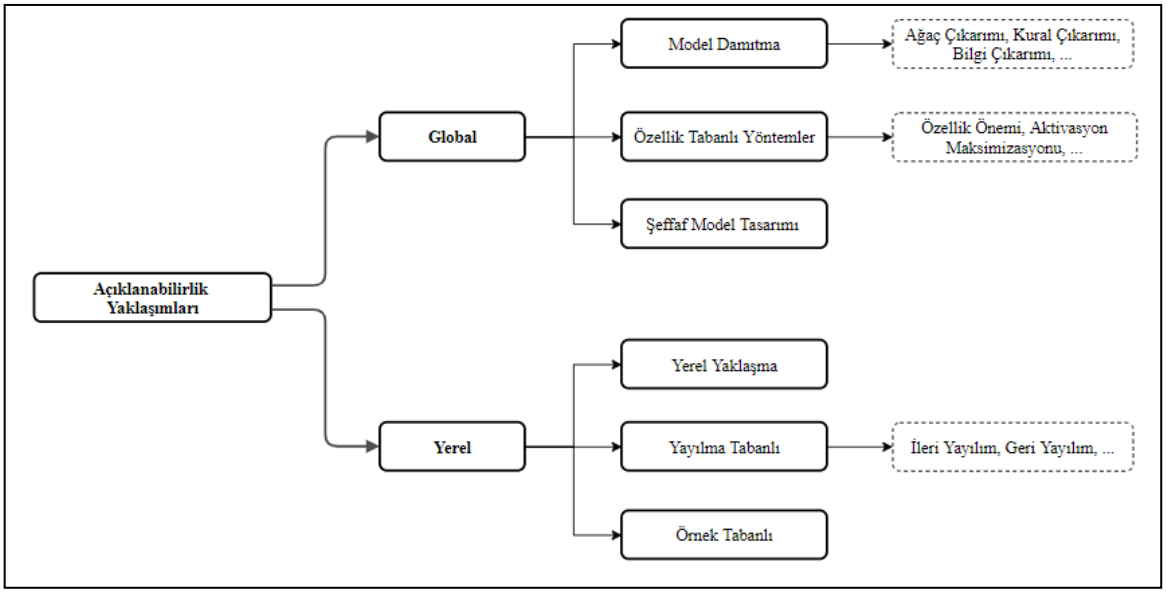
Bu bölüm iki alt bölümde organize edilmiştir. İlk bölüm geleneksel makine öğrenmesi yöntemlerinin açıklanabilirliğini sağlayan yaklaşımların özetlenerek, tez kapsamında geliştirilen ilk iki yöntemin bu doğrultuda değerlendirilmesini kapsamaktadır. İkinci bölümde ise, açıklanabilir yapay zekâ kavramının önem kazanmasını büyük oranda ivmelendiren derin öğrenme tabanlı yöntemlerin açıklanabilirliği için geliştirilen yaklaşımlar göz önünde bulundurularak, tez kapsamında önerilen son yöntemin açıklanabilirliğini de sağlayan yeni bir yaklaşım detaylıca sunulmuştur.

6.1. Açıklanabilirlik Yaklaşımları

Bir modelinin nasıl işlediği ve tahminleri nasıl yaptığı hakkında literatürde pek çok soru bulunmaktadır. Model, verideki hangi özelliklerin en önemli olduğunu bulur? Verideki her bir özellik, modelin gerçekleştirdiği herhangi bir tahmini nasıl etkiler? Bütün özellikler, modelin tahminlerini büyük resim kapsamında nasıl etkiler? Bu soru işaretlerinin giderilmesini sağlayacak birçok faktör vardır. Dolayısıyla bu faktörler ile açıklama yapmanın da pek çok yolu ortaya çıkmaktadır. Yapay zekâ çözümleri geliştirmek için uygulanabilecek açıklanabilirlik yaklaşımları Şekil 6.1'de özetlenmiştir.

Global yöntemler; özelliklerin, öğrenilmiş bileşenler ve yapıların bütünsel bir görünümüne dayanarak, modelin mantığının anlaşılmasını ve tüm tahminler için işlemlerin tam olarak nasıl yapıldığını açıklar [338]. Model damıtma, özellik tabanlı yöntemler ve şeffaf model

tasarımı olarak üç başlık altında değerlendirilebilirler. Model damıtma, orijinal kara kutu modelinden ağaç, kural veya bilgi çıkarımı yaparak yorumlanabilir bir model oluşturup açıklanabilirlik sağlar. Özellik tabanlı yöntemler, genellikle model karmaşıklığının basitleştirilmesine dayanır ve girdi özelliklerinin alaka düzeyini ölçmeyi amaçlar. Her özelliğin model çıktısına göre önemi hakkında bir ölçü sağlayan özellik önemi ve çıktı için tercih edilen girdi üretimi yapan aktivasyon maksimizasyonu gibi teknikler kullanılır. Son olarak şeffaf model tasarımı ise, yorumlanabilirliği artırmak için kara kutu modelini değiştirmeyi veya yeniden tasarlamayı amaçlar.



Şekil 6.1. Yapay zekâ tekniklerinde açıklanabilirlik yaklaşımları

Yerel yöntemler; tüm girdi dağılımından genel belirginlik modeli çıkarılmasının pratikte zorlayıcı olduğu ya da tek bir gözlem için açıklama gerektiği durumlarda kullanılır [339]. Yerel yöntemler, model davranışını tek bir örnek veya bir dizi örnek için doğrulamaya çalışır. Bu kategoride değerlendirilen yerel yaklaşma yöntemi, kara kutu modelinin davranışlarını simüle etmek için anlaşılır vekil modeller üretir. Yayılma tabanlı yöntemler, ek model üretmek yerine çıktıdan girdi özelliğine olan katkıyı ilişkilendiren geri yayılım ya da özelliği bozduktan sonra çıktı farkı yoluyla ilgi düzeyini nicelleştiren ileri yayılım tekniklerini kullanır. Örnek tabanlı yöntemler ise, veri kümesine ilişkin iç görü sağlamak veya kara kutu modelinin davranışını yorumlamak için belirli veri kümesi örneklerini seçer veya üretir. Bu yöntemler, özellikle veri öğeleri görüntüler gibi insan tarafından anlaşılabilir olduğunda etkilidir.

Tez kapsamında geliştirilen yöntemler bu bağlamda incelendiğinde, dinamik imzalar ile dağıtık telekom dolandırıcılığı tespiti modelinin yapısı gereği, açıklanabilirliği için model damıtma yöntemlerinden biri olan kural çıkarımı uygulanmıştır. Yöntemin çıktısı olarak üretilen şüpheli listesindeki ilk 50 arama, NETAŞ uzmanlarına iletilmiştir. Çizelge 6.1'de verilen geri dönüş doğrultusunda, en yüksek benzerlik oranına sahip ilk 50 aramanın 43'ünün dolandırıcı olduğu doğru olarak tespit edildiği onaylanmıştır. Böylelikle yöntem yaklaşık olarak %86 başarı elde etmiştir.

Doğrulama sürecinde uzmanın deneyimlerine ve sezgilerine ek olarak, dolandırıcılık tespitini etkileyen ve kural olarak tanımlanabilen çeşitli bariz durumlar bulunmaktadır. Yapılan incelemelerde, analiz ve testlerde elde edilen bulgular doğrultusunda bu durumlar dokuz maddede şu şekilde ele alınabilir:

1. Uzun süreli az sayıda çağrı,
2. Kısa süreli çok sayıda çağrı,
3. Eş zamanlı aktif çağrı,
4. Belirli bir zaman diliminde uzun süreli çağrı,
5. Belirli bir zaman diliminde belirli bir kategoriye yönelik çok sayıda çağrı,
6. Belirli bir zaman diliminde belirli bir kategoriye yönelik uzun süreli çağrı,
7. Belirli bir zaman diliminde belirli bir kategoriye yönelik çok sayıda uzun süreli çağrı,
8. Aynı kullanıcıdan aynı hedefe aynı anda uzun süreli birden fazla çağrı ve
9. Aynı kullanıcıdan aynı hedefe aynı anda çok sayıda kısa süreli çağrıdır.

Bu bağlamda yapılan değerlendirmeler doğrultusunda, servis kullanıcıları arasında daha önceden meşru oldukları zannedilen kullanıcıların davranışlarındaki anormal durumlar ortaya çıkarılmış ve aslında bu kullanıcıların dolandırıcı davranışlarına benzeyen eylemlerde bulundukları tespit edilmiştir.

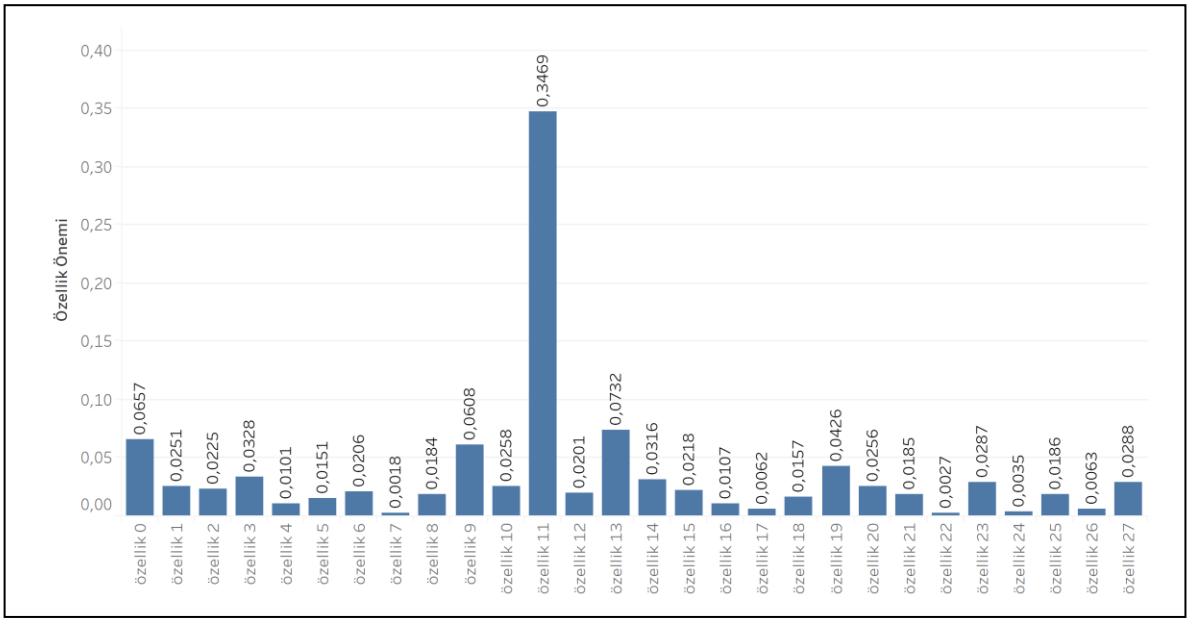
Dağıtık yeniden örnekleme ile kredi kartı dolandırıcılığı tespiti yönteminde, veri kümesi değiştirildikten sonra GBT sınıflandırıcısı kullanılmıştır. GBT, zayıf tahmin modellerinden oluşan bir dizi karar ağacını aşamalı, toplamalı ve sıralı bir şekilde eğiterek, tahminleri iyileştirir ve sınıflandırma performansını artırır. Ağaçları derinlik ve ağaç sayısı bakımından güçlendirmek modelin doğruluğunu arttırırken, analiz hızını ve yorumlanabilirliği de azaltmaktadır.

Çizelge 6.1. DFDDS sonuçları için uzman görüşü raporu

Şüpheli Kullanıcıya Ait Anonimleştirilmiş Telefon Numarası	Tahmin		Şüpheli Kullanıcıya Ait Anonimleştirilmiş Telefon Numarası	Tahmin
eeVrTVYuLTrTe@ruVnVenVnTL	DOĞRU		eeVrVVLTLTee@ruVnVenVnTL	YANLIŞ
8j+VV8+rYu@ruVnVenVnTL	DOĞRU		8+YLjTeL++@ruVnVenVnTL	YANLIŞ
8+8LYuTL8r@ruVnVenVnTL	DOĞRU		eeYe8euj+VjrYV@ruVnVenVnTL	YANLIŞ
8+YujrVLVj@ruVnVenVnTL	DOĞRU		8++jLeVuYY@ruVnVenVnTL	YANLIŞ
8+VruuY8uY@ruVnVenVnTL	DOĞRU		eejYruuu8TL8eL@ruVnVenVnTL	YANLIŞ
8+VjurL+re@ruVnVenVnTL	DOĞRU		eeYe8eT+LerVTT@ruVnVenVnTL	YANLIŞ
eeYur8ej+VY8L+@ruVnVenVnTL	DOĞRU		ee+LYueVjVeL8@ruVnVenVnTL	YANLIŞ
8+jLYLrjje@ruVnVenVnTL	DOĞRU			
88++e+LuL+@ruVnVenVnTL	DOĞRU			
8jYV8LeeVj@ruVnVenVnTL	DOĞRU			
eeYe88+++VV+++@ruVnVenVnTL	DOĞRU			
eeYe8e8eu8ueLj@ruVnVenVnTL	DOĞRU			
8+LejLr8TT@ruVnVenVnTL	DOĞRU			
8+8jrjV8jT@ruVnVenVnTL	DOĞRU			
eeYe8j+uVYe+Vr@ruVnVenVnTL	DOĞRU			
eejeuTVLVY8Tj@ruVnVenVnTL	DOĞRU			
8+LVr+V+jT@ruVnVenVnTL	DOĞRU			
eeVjreV+uLjVL@ruVnVenVnTL	DOĞRU			
8+r8eLLuLe@ruVnVenVnTL	DOĞRU			
*8+uLjreYVu@ruVnVenVnTL	DOĞRU			
8+VV+rYjY8@ruVnVenVnTL	DOĞRU			
eeVV+Yjj8urTV@ruVnVenVnTL	DOĞRU			
eeYe88+T8T8Yrr@ruVnVenVnTL	DOĞRU			
eeVrT+jY88rY@ruVnVenVnTL	DOĞRU			
eeYe8j+VreVeVV@ruVnVenVnTL	DOĞRU			
8+YTTLTrVV@ruVnVenVnTL	DOĞRU			
VrV++jVrer@ruVnVenVnTL	DOĞRU			
8+88Y8ruLT@ruVnVenVnTL	DOĞRU			
eejYruejL88LLY@ruVnVenVnTL	DOĞRU			
8+Vue8L8uu@ruVnVenVnTL	DOĞRU			
8+uT8eVLYr@ruVnVenVnTL	DOĞRU			
8j8ujeeT8j@ruVnVenVnTL	DOĞRU			
8+rYee8j8j@ruVnVenVnTL	DOĞRU			
8+uV8Y8VjV@ruVnVenVnTL	DOĞRU			
VrV8rueYee@ruVnVenVnTL	DOĞRU			
8jT8+YjTr8@ruVnVenVnTL	DOĞRU			
8+LYL+88ru@ruVnVenVnTL	DOĞRU			
VrV+erYeee@ruVnVenVnTL	DOĞRU			
88VTjeueue@ruVnVenVnTL	DOĞRU			
8+jer+ruuV@ruVnVenVnTL	DOĞRU			
jYr8VruLurje8@ruVnVenVnTL	DOĞRU			
8+Vr+erYVL@ruVnVenVnTL	DOĞRU			
8+eY8j8+YV@ruVnVenVnTL	DOĞRU			

GBT için açıklanabilirlik, Şekil 6.2'de verildiği üzere özelliklerin önem düzeylerinin çıkarılmasıyla ve Şekil 6.3'de verildiği üzere ağacın oluşturduğu kuralların çıkarılmasıyla sağlanmıştır.

Özellik önemi, hedef değişkeni tahmin etmede ne kadar yararlı olduklarına bağlı olarak giriş özniteliklerine puan atayan teknikleri ifade etmektedir. Öznitelik önemini bilmek; modeli sadece doğrulamakla kalmaz, veriye ve modele ilişkin anlayış sağlar, modelin verimliliğini ve etkinliğini artırabilecek boyutluluk azaltma ve özellik seçimi için temel oluşturur. Ek olarak, yorumlanabilirliğin daha önemli olduğu kimi durumlar için doğruluktan feragat edilmesini sağlar. GBT için her özelliğin önemi, topluluktaki tüm ağaçlardaki öneminin ortalaması alınarak elde edilir ve son olarak önem vektörü, toplamı bir olacak şekilde normalleştirilir [340]. Bu doğrultuda European veri kümesinden elde edilen ve Şekil 6.2'de görselleştirilen, dağıtık yeniden örnekleme ile kredi kartı dolandırıcılığı tespiti modeline ait önem düzeyi vektöründe görüldüğü üzere, dolandırıcı veya meşru kararının verilmesindeki en baskın etkenin on birinci özellik olduğu görülmektedir.



Şekil 6.2. European veri kümesindeki özelliklerin önem düzeyleri

Karar kuralları, basit ve kolay yorumlanabilir bir bilgi temsili şeklidir. Karar kurallarının oldukça kesin olması, modelin mevcut örneklere çok fazla ayarlanmış ve gürültüye bağımlı olmasına yol açabilir. Kuralların yaklaşık olması ise, daha az sayıda öznitelik içeren ve nispeten iyi doğruluğa sahip kurallarla çalışmayı mümkün kılar. Dolayısıyla, kuralların

görselleştirilmesi, çıktı parametreleri doğrultusunda modelin güncellenmesini ve geliştirilmesini sağlamaktadır. Karar kuralı, bir şart ve bir tahminden oluşan basit bir IF-THEN ifadesidir ve kararın bulunduğu yaprak düğümüne giden yoldaki koşullar birleşiminden oluşur. Şekil 6.3'de GBT sınıflandırıcısının oluşturduğu ağaçlardan birine ait karar kurallarının bir kısmı verilmiştir. Bu ağacın ilk iki kuralına bakıldığında, ayırt edici unsurların özellik önem vektöründeki en etkili iki özelliğin olduğunu görmek mümkündür. Aynı mantıkla model için yeterince bilgi vermeyen veya gereksiz öngörücüler, kurallar dizisinin sonlarında yer almaktadırlar.

```
Tree 13 (weight 0.1):
    If (feature 13 <= -1.966415133174354)
     If (feature 11 <= -2.5072664768204422)
      If (feature 9 <= -1.4708993418688674)
       If (feature 15 <= 0.6731595716749437)
        If (feature 19 <= 0.8969884428558537)
         If (feature 10 <= -0.01860510438759986)
          Predict: -0.5384314669911769
         Else (feature 10 > -0.01860510438759986)
          If (feature 3 <= 0.8445681894320942)
           Predict: 0.1014273101299857
          Else (feature 3 > 0.8445681894320942)
           Predict: 0.23657162995788086
        Else (feature 19 > 0.8969884428558537)
         If (feature 14 <= 1.0029848778583776)
          If (feature 14 <= -0.19974984050045463)
           Predict: 0.02338451295369103
          Else (feature 14 > -0.19974984050045463)
           Predict: 0.36349705807782734
         Else (feature 14 > 1.0029848778583776)
          If (feature 23 <= -0.2190876831173119)
           Predict: 0.2492146595812085
          Else (feature 23 > -0.2190876831173119)
           Predict: -0.6318316803948868
       Else (feature 15 > 0.6731595716749437)
        If (feature 25 <= -1.175921157739397)
         If (feature 0 <= 0.07807622884878547)
          Predict: 0.22990528743378888
         Else (feature 0 > 0.07807622884878547)
          Predict: 0.22979240681067462
        Else (feature 25 > -1.175921157739397)
         If (feature 21 <= -0.47040091975833437)
          If (feature 2 <= -1.1130723200602832)
           Predict: -0.25971883762361486
          Else (feature 2 > -1.1130723200602832)
           Predict: -0.30674222343549884
         Else (feature 21 > -0.47040091975833437)
          Predict: 0.22784313435551606
      Else (feature 9 > -1.4708993418688674)
       If (feature 20 <= 1.2138103864917307)
        If (feature 4 <= -0.9601478083160407)
        ...
```

Şekil 6.3. Ağaç tabanlı sınıflandırıcıya ait kurallardan örnek bir bölüm

6.2. Derin Sinir Ağlarında Açıklanabilirlik İçin Yeni Bir Yaklaşım

Pek çok uygulama alanında yapay zekâ modellerinin kullanımının artmasıyla, bu modellerin sonucunu açıklama ihtiyacı; etik, adli ve güvenlik nedenleriyle önem kazanmaya başlamıştır. Dolayısıyla geliştirilen açıklamaların, yapay zekâ algoritmalarının güvenilir, şeffaf ve önyargısız olduğunu vurgulaması gerekmektedir. Derin sinir ağları, karmaşık girdi birleşimleri yoluyla verilerdeki kalıpları öğrenen, oldukça iç içe geçmiş doğrusal olmayan modeller olduğu için, geleneksel makine öğrenmesi yöntemlerine nazaran tahminlerinin kesin nedenlerini deşifre etmek daha zordur.

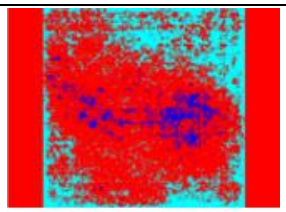
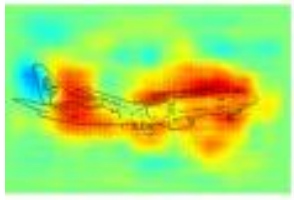
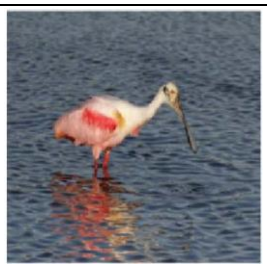
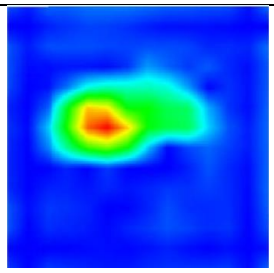
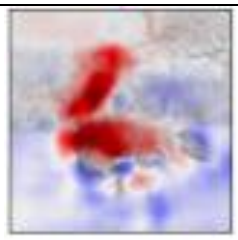
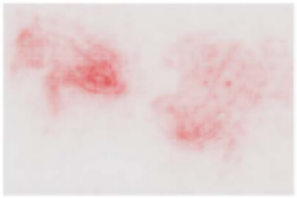
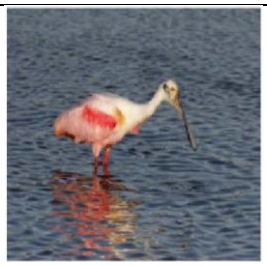
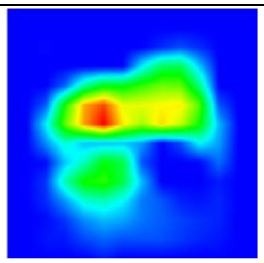
Çizelge 6.2'de görüldüğü üzere, derin öğrenmeye yönelik mevcut açıklanabilirlik yaklaşımlarının çoğu doğrudan daha yorumlanabilir modeller yapmaya çalışmak yerine, ya yorumlanabilir bazı yaklaşık modeller geliştirmeye odaklanırlar ya da model tarafından belirli bir kapasitenin üzerinde kullanılan bazı göze çarpan özellikleri ortaya çıkarmaya çalışır ve daha sonra kullanıcının modelin altında yatan bilişsel süreçleri anlamasını teşvik eder. Fakat bu yöntemler birkaç yönden sınırlıdır. İlk olarak, açıklama yöntemleriyle hesaplanan ısı haritaları birinci dereceden bilgileri görselleştirir, yani hangi giriş özelliklerinin tahminle alakalı olarak tanımlandığını gösterirken, bu özellikler arasındaki ilişkinin kendi başına mı yoksa birlikte mi oluştukları belirsizdir [341]. Diğer bir sınırlama, açıklamaların düşük soyutlama seviyesidir, yani ısı haritaları görüntüdeki nesneler veya sahne gibi daha soyut kavramlarla ilişkilendirmeden sadece belirli piksellerin önemli olduğunu vurgular [342].

Bu sınırlamalar göz önünde bulundurularak, GAF için Grad-CAM adı verilen yeni bir yaklaşım önerilmiştir. Görüntüye dönüştürme ile kredi kartı dolandırıcılığı tespiti modeli üzerinde örneklendirilen bu yaklaşımda ilk olarak, kredi kartı işlemleri zaman serisi verisi olarak değerlendirilerek GAF görüntülerine dönüştürülmüştür. GAF görüntüleri, bir zaman anındaki değerin tüm zaman anlarındaki değerlere olan kutupsal uzaklığı ile elde edilmektedir. Bu sayede elde edilen ilişkiler, öznitelik önem düzeyindekinin aksine özniteliğin etikete etkisini değil, iki özniteliğin birbirleri ile etkileşimini kapsamaktadır. Daha sonra, CNN ile sınıflandırılan GAF görüntülerinin açıklanabilirliğini sağlamak için Gradyan Ağırlıklı Sınıf Aktivasyon Haritalama (Gradient-weighted Class Activation Mapping, Grad-CAM) kullanılmıştır. Son olarak, literatürdeki yerel yaklaşımların aksine tüm test verisini özetleyebilen yeni bir global açıklanabilirlik yaklaşımı geliştirilmiştir.

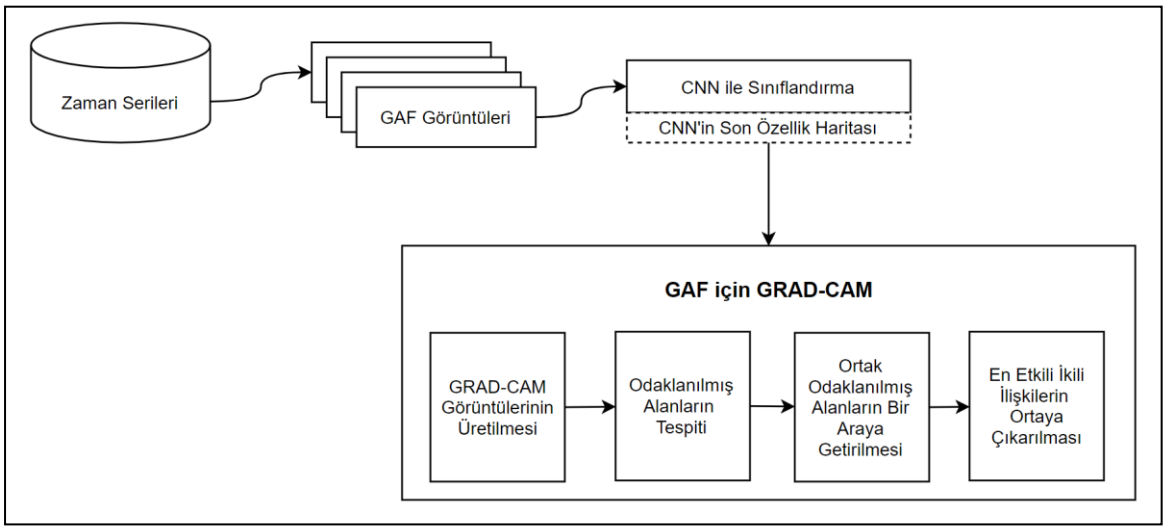
Çizelge 6.2. Derin sinir ağlarında açıklanabilirlik yaklaşımları

Kaynak	Yaklaşım	Amaç	Kapsam		Yöntem		Kullanım	
			Global	Yerel	Yayılım	Pertürbasyon	Modele Özgü	Modelden Bağımsız
[343]	Activation Maximization	Birim düzeyinde yorumlama	-	+	+	-	-	+
[344]	Saliency Maps	Nesne bölütleme	-	+	+	-	-	+
[345]	DeConvNet	Ara özellik katmanlarının işlevini ortaya çıkarma	-	+	+	-	-	+
[346]	Bayesian Case Model (BCM)	Prototip çıkarma	+	-	-	-	+	-
[347]	Layer-wise Relevance Propagation (LRP)	Piksel bazında ayrıştırma	+	+	+	-	-	+
[348]	Bayesian Rule Lists (BRL)	Karar listeleri oluşturma	+	-	-	-	+	-
[349]	Class Activation Maps (CAM)	Ayırt edici bölge yerelleştirme	-	+	+	-	-	+
[350]	Local Interpretable Model-agnostic Explanations (LIME)	Yerel olarak yorumlanabilir bir modele yaklaştırma	+	+	-	+	-	+
[351]	SHapley Additive exPlanations (SHAP)	Önem değeri atama	+	+	-	+	-	+
[352]	Gradient-weighted Class Activation Mapping (Grad-CAM)	Önemli bölgeleri vurgulayan yerelleştirme haritası	-	+	+	-	-	+
[353]	Prediction Difference Analysis (PDA)	Piksellerinin karar için önemli ya da karşı kanıt olduğunu gösterme	-	+	-	+	-	+
[354]	Deep Taylor Decomposition	Sınıflandırma kararını girdi öğelerinin katkılarına bölerek yorumlama	-	+	-	-	-	+
[355]	Randomized Input Sampling for Explanation (RISE)	Piksel seviyesinde önem haritası oluşturma	-	+	-	+	-	+
[355]	Grad-CAM++	Grad-CAM'den daha iyi açıklama sağlama ve görüntüdeki birden çok nesneyi açıklama	-	+	+	-	-	+
[357]	Salient Relevance (SR) Map	Görüntüleri tanıma ve alanlardan özellikleri öğrenme sürecini açıklama	-	+	+	-	-	+
Önerilen Yaklaşım	GAF için Grad-CAM	GAF görüntülerinin önemli ilişkilerini çıkarma ve önem haritası oluşturma	+	+	+	-	-	+

Çizelge 6.3. Çeşitli açıklanabilirlik yaklaşımlarına ait örnek görüntüler

Kaynak	Yaklaşım	Orijinal Görüntü	Açıklanabilirlik Yaklaşımının Sonucu
[344]	Saliency Maps		
[347]	Layer-wise Relevance Propagation (LRP)		
[352]	Gradient-weighted Class Activation Mapping (Grad-CAM)		
[353]	Prediction Difference Analysis (PDA)		
[354]	Deep Taylor Decomposition		
[355]	Grad-CAM++		
[357]	Salient Relevance (SR) Map		

Çizelge 6.3'de çeşitli görüntülerin sınıflandırılmasından sonra, bazı açıklanabilirlik yaklaşımları doğrultusunda ürettikleri görüntüler verilmiştir. Her örnekteki referans görüntüler karşılaştırıldığında, sınıflandırma için en önemli ve en önemsiz alanlar açıkça görülebilmektedir. Fakat GAF görüntüleri üzerinde açıklanabilirliği sağlamada, mevcut görselleştirme yöntemleri yetersiz kalmaktadır. Çünkü görüntüde yer alan bir nesne veya arka plan gibi bölümlerin sınıflandırıcının kararında etkili olduğunu gösteren vurgulamalar, zaman serilerinden elde edilen ve belirli bir ölçekteki renk haritalarından oluşan GAF görüntüleri için ilk bakışta yeterli anlam ifade etmemektedir. Başka bir deyişle, açıklanabilirlik sağlayan bu yaklaşımlar, GAF görüntüleri için fazladan bir açıklamayı daha gerektirmektedir.



Şekil 6.4. GAF için Grad-CAM yaklaşımının metodolojisi

Algoritma 1: GAF için Grad-CAM Yaklaşımı

Function *GAF için Grad-CAM(GAF_görüntüleri)*:
    eğitim, test = EğitimTestAyır(GAF_görüntüleri)
    model = Sınıflandırma.EvrişimliSinirAğı(eğitim)
    NöronÖnemAğırlıkları⟵ $OrtalamaGradyanHesaplama(model.sonÖzellikHaritası)$
    for *each görüntü in test* **do**
        GradCAM⟵ $RELU(görüntü, NöronÖnemAğırlıkları)$
        eşikDeğer = parlakAlanlar(GradCAM)
        odaklanılmışAlanlar⟵ $ikiliMaskeOlustur(GradCAM, esikDeğer)$
        son.Birleştir(odaklanılmışAlanlar)
    end
    Yayınla(son)
end

Şekil 6.5. GAF için Grad-CAM yönteminin sözde kodu

Bu bağlamda, kararın verilmesini sağlayan ilişkilerin daha net ortaya çıkarılmasını amaçlayan ve GAF için Grad-CAM olarak adlandırılan yeni bir yaklaşım geliştirilmiş ve test edilmiştir. GAF için Grad-CAM yaklaşımının Şekil 6.4'de metodolojisi ve Şekil 6.5'de sözde kodu verilmiştir.

Zaman serisini oluşturan ikili ilişkiler arasında, benzerlikten ya da farklılıktan kaynaklanan bir örüntü var ise bu örüntünün GAF görüntülerine de aktarıldığı bilinmektedir. Bu görüntüler, CNN kullanılarak dolandırıcı ya da meşru olarak sınıflandırılmıştır. Daha sonra, kara kutu olarak tanımlanan CNN modelini daha şeffaf hale getirmek için görsel açıklamalar üreten Grad-CAM tekniği kullanılmıştır. Grad-CAM tekniği GAF görüntülerinin anlamlandırılması için yetersiz kaldığından, bu aşamada elde edilen çıktılar ile süreç tekrar özellikler arası ilişki boyutuna taşınmıştır.

Grad-CAM, CNN'nin son katmanına akan gradyan bilgisini kullanarak, girdi görüntüsünün sınıflandırıcıyı en çok etkileyen bölgelerini vurgulayan bir yerelleştirme haritası üretir [353]. *L* olarak gösterilen bu haritayı elde etmek için, Eş. 6.1'de verildiği üzere, *c* sınıfının gradyanı yani sınıf puanı olan y_c, evrişim katmanının son özelik haritası A^k ile hesaplanır.

$$y_c = \sum_k w_k^c \frac{1}{Z} \sum_i \sum_j A_{ij}^k \tag{6.1}$$

Bu gradyanlar, α_k^c ağırlıklarını elde etmek için Eş. 6.2'de verildiği üzere havuzlanır. Son olarak, Eş. 6.3'de verilen Grad-CAM, düzeltilmiş doğrusal birim fonksiyonu ile özellik haritalarının ağırlıklı bir kombinasyonu olarak üretilir.

$$\alpha_k^c = \frac{1}{Z} \sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k} \tag{6.2}$$

$$L_{Grad-CAM}^c = \text{RELU}(\sum_k \alpha_k^c A^k) \tag{6.3}$$

Sınıflandırma üzerinde daha fazla etkisi olan değişken ilişkilerini ayırt etmek için görüntülere eşikleme uygulanmıştır. Görüntü yoğunluğunun, sabit bir değerden daha düşük olduğu her pikselin değeri siyah olarak değiştirilmiştir. Böylece, ısı haritasında belirli bir eşik değerinin üzerinde olan ilişkilere filtre uygulanarak odaklanılmış Grad-CAM sonucu

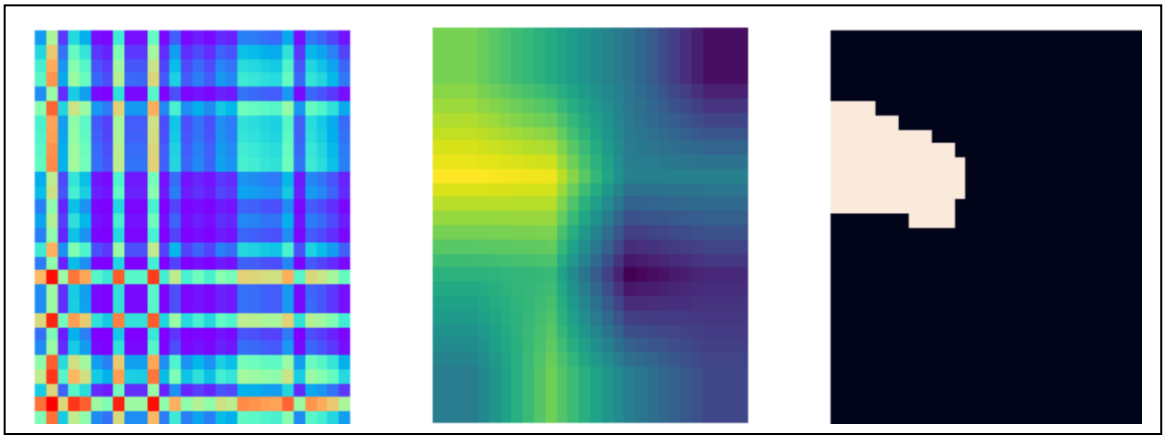
olan ikili maskeler elde edilmiştir. θ, görüntüdeki parlak alanların piksel değeri olan eşik değeri olmak üzere F^c odaklanılmış Grad-CAM görüntüleri Eş. 6.4'de verildiği şekilde hesaplanmaktadır.

$$F^c(i,j) = \begin{cases} 1, & L^c(i,j) \geq \theta \\ 0, & diğer \end{cases} \tag{6.4}$$

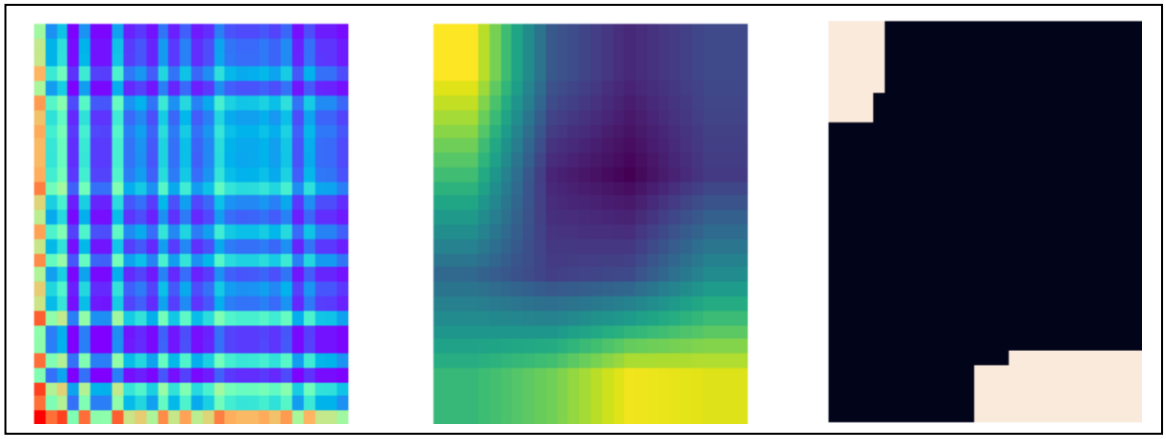
Son olarak, en çok önem addedilen ilişkileri tespit etmek ve görselleştirmek için test kümesindeki her sınıfın odaklanılmış Grad-CAM görüntüleri birleştirilmiştir. *mxm* boyutundaki *n* adet odaklanılmış Grad-CAM görüntüsünün bulunduğu bir test kümesinde, önemli piksellerin vurgulanma sıklığı olan T^c Eş. 6.5'de verildiği şekilde hesaplanmaktadır.

$$T^c(i,j) = \sum_{k=1}^{n} \sum_{i=1}^{m} \sum_{j=1}^{m} k_{i,j}^c \tag{6.5}$$

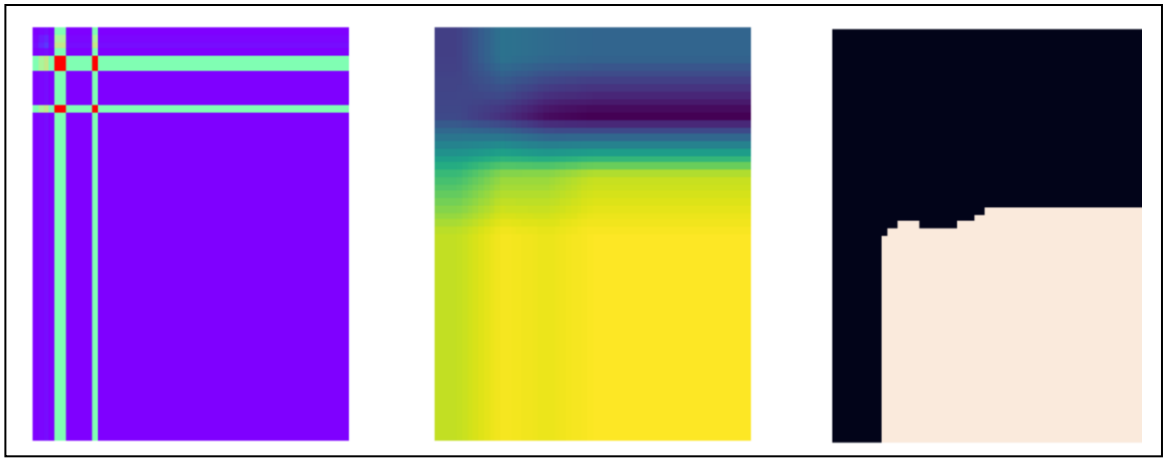
FDIC yönteminin GAF için Grad-CAM yaklaşımı ile açıklanabilirlik aşaması; European veri kümesine ait Şekil 6.6'da bir dolandırıcı örnek, Şekil 6.7'de bir meşru örnek ve ccFraud veri kümesine ait Şekil 6.8'de bir dolandırıcı örnek ve Şekil 6.9'da bir meşru örnek özelinde incelendiğinde, gelinen noktada zaman serisi GAF ile görüntüye dönüştürülmüş ve CNN ile sınıflandırılmıştır. GAF'daki kırmızı ya da mavi alanlar, kesişen iki değişken arasındaki ilişkinin en güçlü olduğu alanları gösterirken, yeşil ve turuncu alanlar orta şiddetli ilişkileri göstermektedir. GAF'daki kırmızı ya da mavi alanlar, kesişen iki değişken arasındaki ilişkinin en güçlü olduğu alanları gösterirken, yeşil ve turuncu alanlar orta şiddetli ilişkileri göstermektedir.



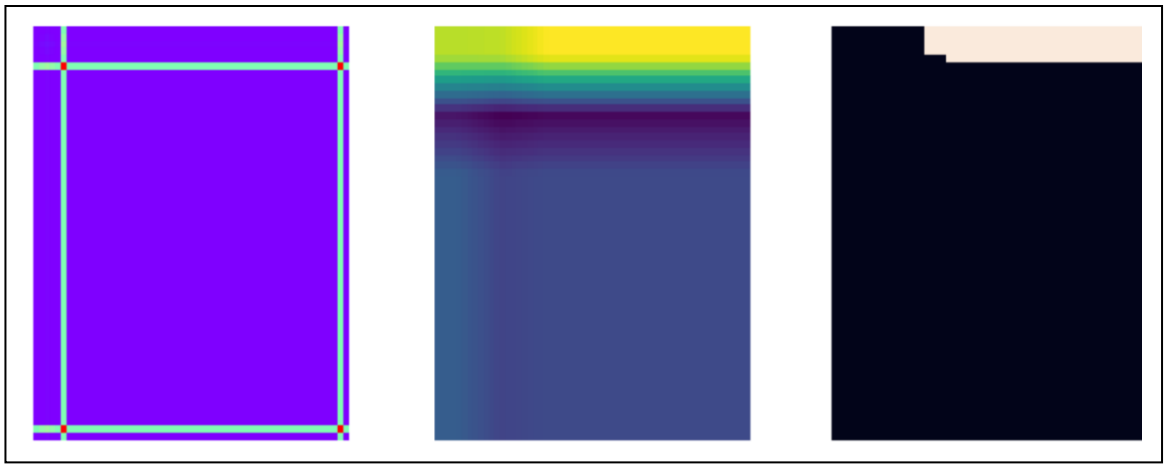
Şekil 6.6. European veri kümesindeki bir dolandırıcılık örneği için sırasıyla GASF, Grad-CAM ve odaklanılmış Grad-CAM



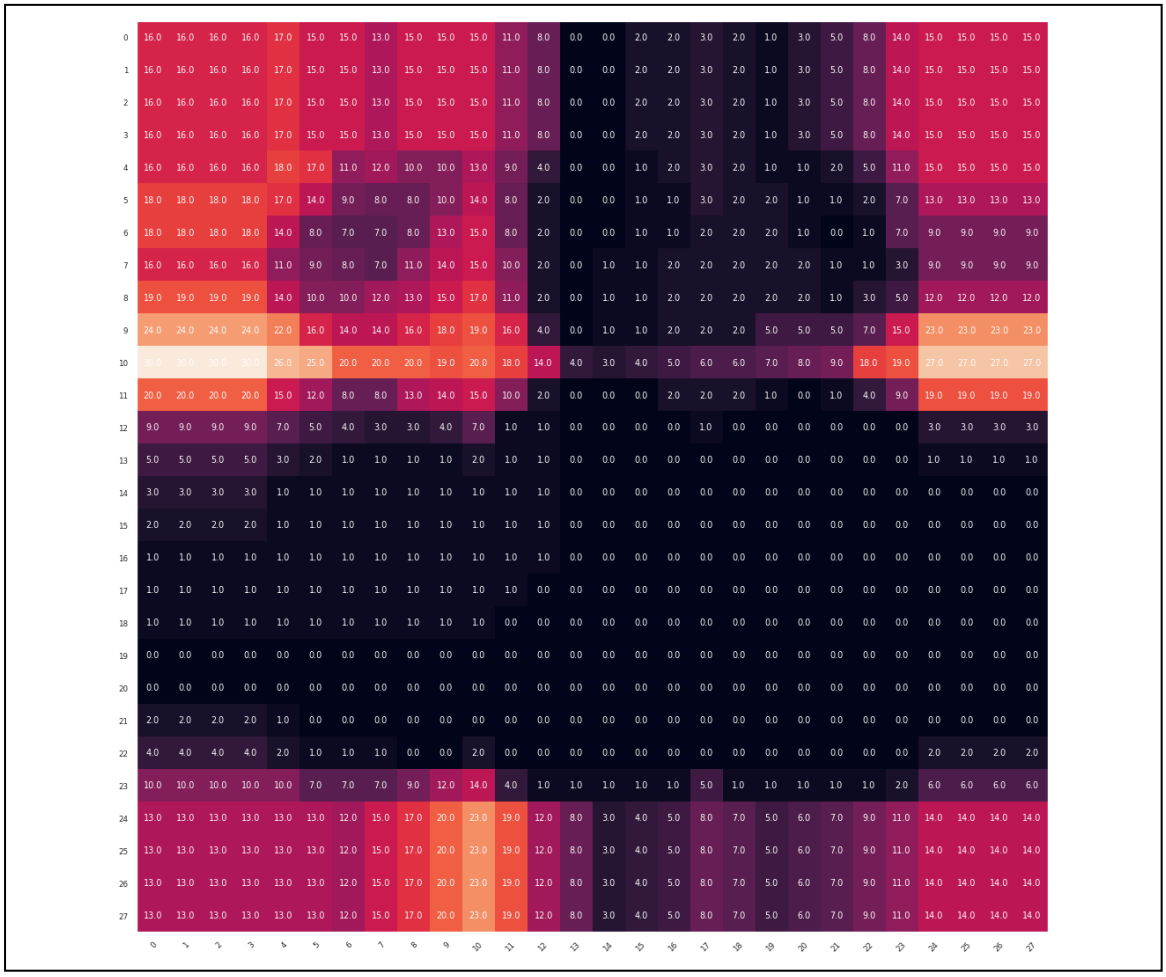
Şekil 6.7. European veri kümesindeki bir meşru örneği için sırasıyla GASF, Grad-CAM ve odaklanılmış Grad-CAM



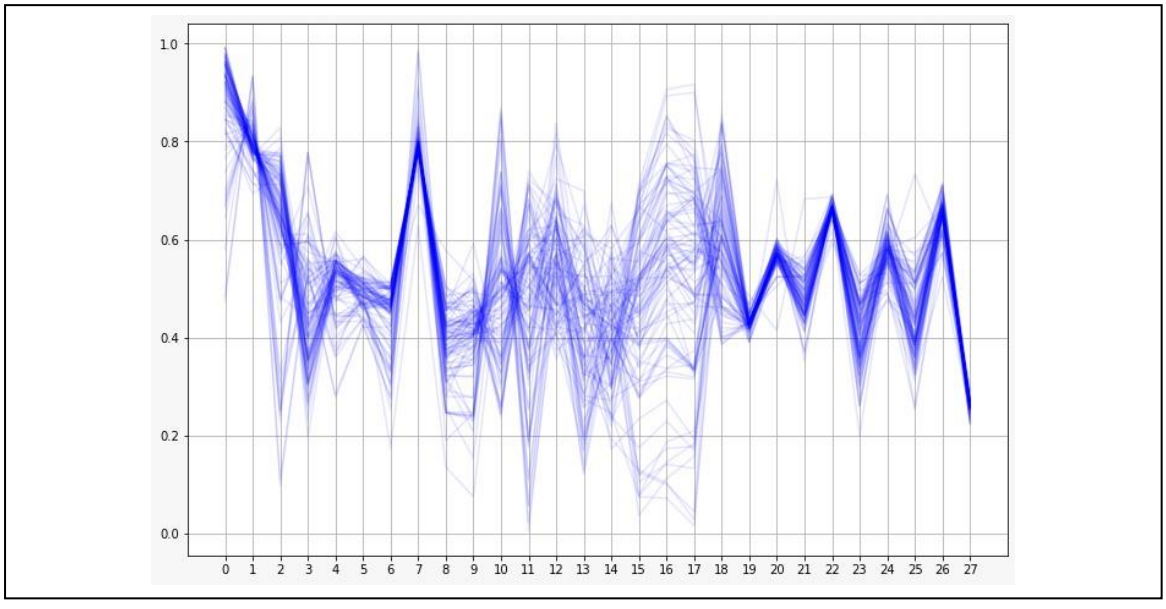
Şekil 6.8. ccFraud veri kümesindeki bir dolandırıcılık örneği için sırasıyla GASF, Grad-CAM ve odaklanılmış Grad-CAM



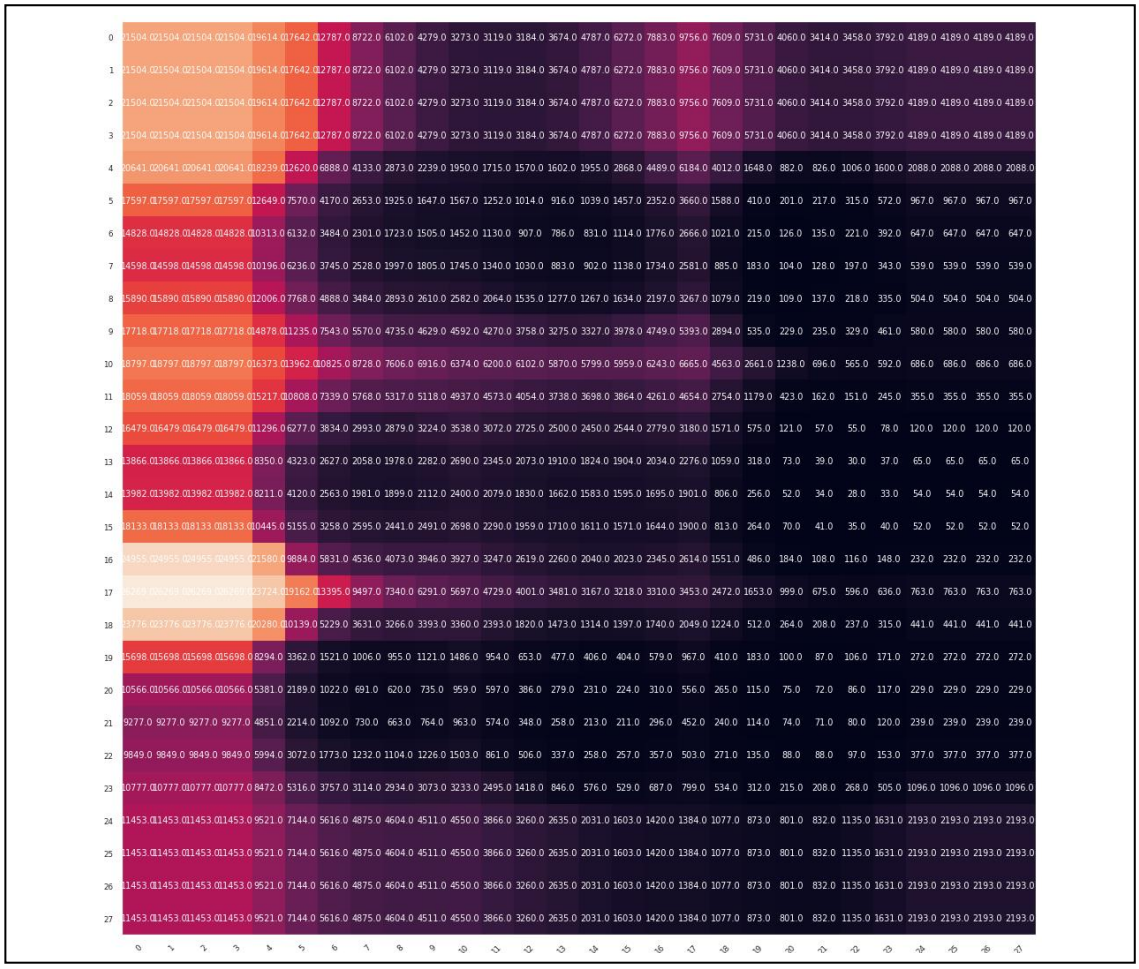
Şekil 6.9. ccFraud veri kümesindeki bir meşru örneği için sırasıyla GASF, Grad-CAM ve odaklanılmış Grad-CAM



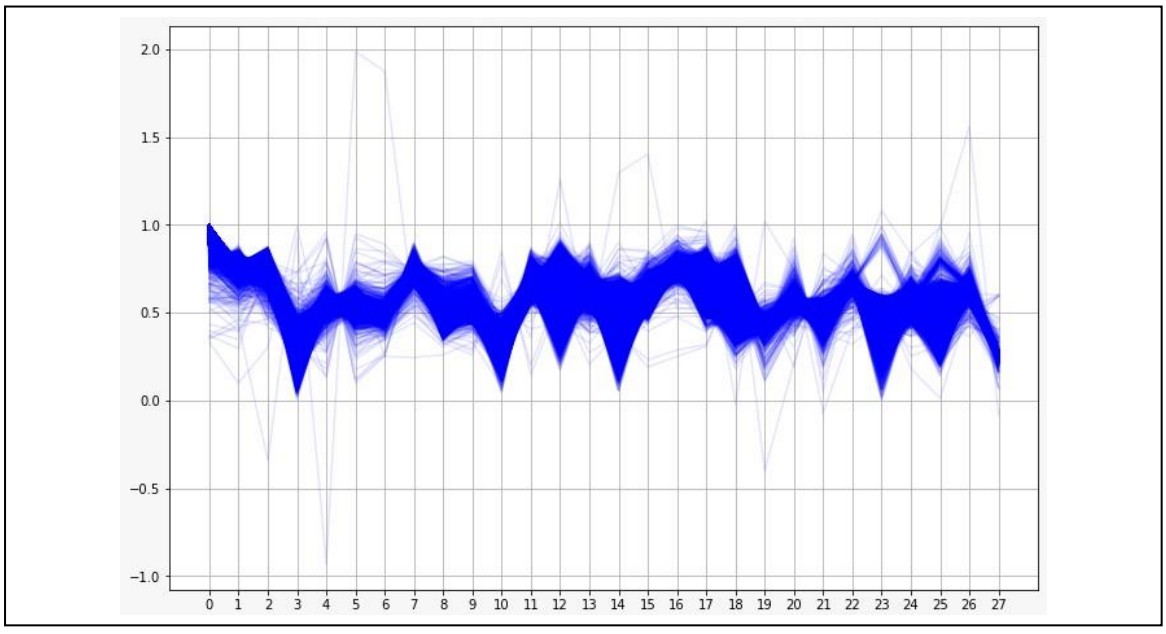
Şekil 6.10. European veri kümesindeki dolandırıcılık örnekleri için sınıflandırıcının Odaklandığı noktalar



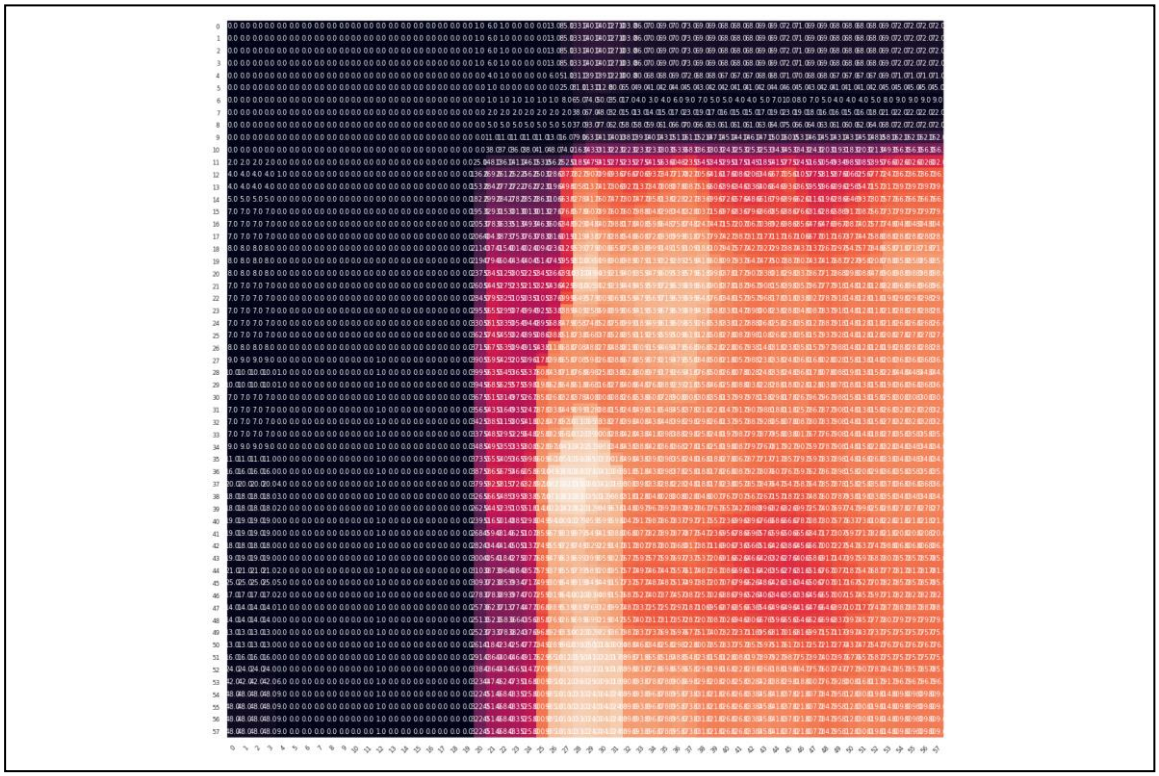
Şekil 6.11. European veri kümesindeki dolandırıcılık örneklerin gerçek değerleri



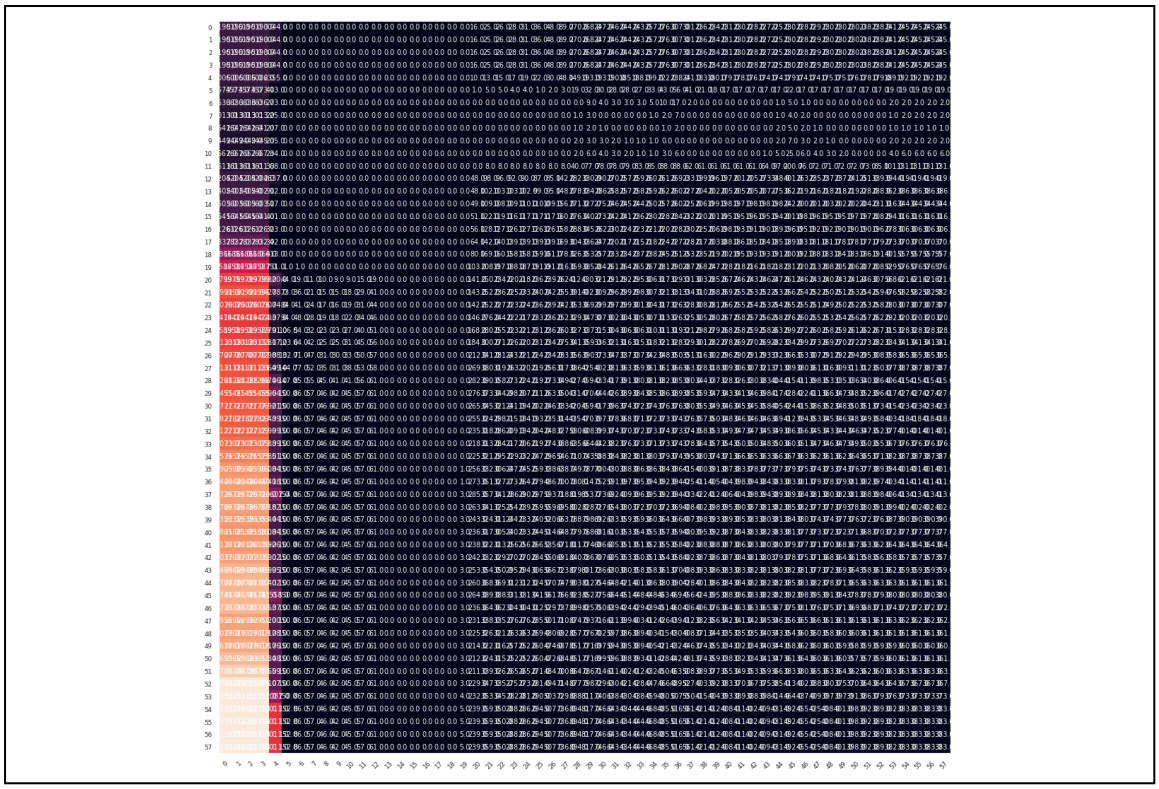
Şekil 6.12. European veri kümesindeki meşru örnekler için sınıflandırıcının odaklandığı noktalar



Şekil 6.13. European veri kümesindeki meşru örneklerin gerçek değerleri



Şekil 6.14. ccFraud veri kümesindeki dolandırıcılık örnekleri için sınıflandırıcının odaklandığı noktalar



Şekil 6.15. ccFraud veri kümesindeki meşru örnekler için sınıflandırıcının odaklandığı noktalar

Sınıflandırma sonrasında test verisine uygulanarak elde edilen Grad-CAM çıktısındaki parlak alanlar, sınıflandırma sırasında modelin odaklandığı bölgeleri yani en ayırt edici ilişkileri işaret etmektedir. Mavi renkten sarı renge doğru olan geçiş, en az etkili piksellerden en çok etkili piksellere doğru geçişi görselleştirmektedir. Odaklanılmış Grad-CAM ise, ısı haritası olarak ifade edilen Grad-CAM görüntüsünün sadece en baskın alanlarını özetleyen ikili maske olarak sunulmaktadır.

Son olarak, European veri kümesine ait Şekil 6.10'da dolandırıcılık örnekleri ve Şekil 6.11'de meşru örnekler için çıkarılan odaklanılmış Grad-CAM görüntüleri verilmiştir. Şekil 6.12'de ve Şekil 6.13'de ise modelin test edildiği görüntülerin gerçek sinyalleri verilmiştir. Model dolandırıcı sınıf için karar sürecinde özellikle ilk altı değişkenin onuncu değişken ile ve son dört değişkenin yine onuncu değişken ile ilişkisi üzerine odaklanmaktadır. Geliştirilen modelin; meşru sınıf için ise özellikle ilk beş değişken ile on altıncı, on yedinci ve on sekizinci özellikler arasındaki ilişkilere göre karar verdiği görülmektedir.

ccFraud veri kümesine ait Şekil 6.14'de dolandırıcılık örnekleri ve Şekil 6.15'de meşru örnekler için çıkarılan odaklanılmış Grad-CAM görüntüleri değerlendirildiğinde ise, dolandırıcılık tespiti için işlemin yapıldığı yeri temsil eden kategorik özelliğin değerlerini temsili olan piksellerin modelin karar verme sürecinde etkili olduğu görülmektedir. Geliştirilen model meşru sınıf için karar verirken ise; bakiye, limit ve işlem sayısını belirten ilk özelliklerin işlemin yapıldığı yerler ile olan ilişkisinin ayırt edici olduğu tespit edilmiştir.

Dolandırıcılık ve meşru faaliyetlerin günbegün arttığı büyük veri çağında, daha fazla artırılmış analitiğe doğru ilerledikçe, geliştirilen modellerin ve iç görülerin açıklanabilirliği; güven, yasal uyumluluk ve marka itibar yönetimi açısından kritik hale gelmeye başlamıştır. Bu sebeple, geliştirilen modellerin güçlü ve zayıf yönlerini vurgulayan, olası davranışını tahmin eden ve olası önyargıları belirleyen açıklanabilir yapay zekâ, kuruluşlar ile müşterileri ve paydaşları arasında daha fazla güven sağlanmasına yardımcı olacaktır.

7. SONUÇ VE ÖNERİLER

Bu tez kapsamında, literatürdeki büyük veride sınıf dengesizliği probleminin sebep olduğu sınırlamalar ve çıkmazlar değerlendirilerek, bu probleme çözüm üreten üç yeni tespit yöntemi ile bir açıklanabilir yapay zekâ yaklaşımı önerilmiştir.

İlk yöntem, *"kullanıcı davranışlarındaki sapmalar ve bilinen dolandırıcılık işlemleri beraber değerlendirilerek dolandırıcılık tespit yöntemi geliştirilebilir mi?"* sorusuna cevaben geliştirilen ve *"Dinamik İmzalar İle Dağıtık Dolandırıcılık Tespiti"* olarak isimlendirilen yöntem, 3151243 numaralı TÜBİTAK TEYDEB 1501 projesi kapsamında gerçekleştirilmiş olup gerçek dünya verileri üzerinde telekom dolandırıcılığını tespit edilmiştir. Dolandırıcılık etiketinin edinilmesindeki zorluk göz önünde bulundurularak, kullanıcıların davranış değişiklikleri üzerinden tespitin dinamik olarak ve MapReduce paradigması ile büyük kümeler üzerinde dağıtık analizi gerçekleştirebilecek şekilde tasarlanmıştır. Önerilen yöntemde; her arayanın belirli bir zaman diliminde yaptığı aktiviteleri özetleyen imzası oluşturulur, dolandırıcı olduğu bilinmeyen arayanın yeni ve bir sonraki zaman diliminde oluşan aktivitesi eskisinden oldukça farklı ise bu durum şüpheli imza olarak değerlendirilir ve eğer şüpheli imza, dolandırıcı olduğu bilinen imzalardan herhangi birine yakın benzerlik gösterirse de uyarı olarak değerlendirilir ve raporlanır. Yöntemin doğrulanması için bilinen gerçek 50 dolandırıcılık verisi ile test edilmiş ve bunlardan 43'ü doğru olarak tespit edilmiştir. Yöntem aynı zamanda açıklanabilirliğin sağlanması amacıyla uzman görüşüne sunulmuştur. Yöntemin çıktısı olan şüpheli listesinde dolandırıcı olma ihtimali en yüksek olan ilk 50 kullanıcıya ait aramalar sorumlu uzmana raporlanmıştır. Uzman tarafından, bu 50 aramanın 43'nün doğru yapıldığı belirtilmiş olup, yöntem, %86 başarı elde edilmiştir. Daha önceden farkında olunmayan dolandırıcılık faaliyetlerini ortaya çıkarması sebebiyle bu oran, proje hedefleri doğrultusunda yeterli görülmüş ve önerilen yöntem, NETAŞ'a ait dolandırıcılık yönetim sistemi olan NOVA FMS ürününe entegre edilmiştir. Geliştirilen yöntem firmanın lansmanlarında tanıtılmıştır.

İkinci yöntem, *"veri kümesindeki dolandırıcı ve meşru örnek dağılımı değiştirilerek daha iyi tespit performansı elde edilebilir mi?"* sorusuna cevaben tasarlanmış ve *"Dağıtık Yeniden Örnekleme İle Dolandırıcılık Tespiti"* olarak isimlendirilen bir yöntem olup, dağıtılmış küme tabanlı yeniden örneklemeye dayanmaktadır ve kredi kartı dolandırıcılığını tespit eder.

Hem çoğunluk ve azınlık sınıf arasındaki dengesizliği hem de her sınıf içerisindeki alt kavramların varlığından kaynaklanan dengesizliği gidermek amacıyla, veri manipülasyonu yapılarak MapReduce paradigması ile büyük kümeler üzerinde dağıtık analizi gerçekleştirebilecek şekilde tasarlanmıştır. Önerilen ikinci yöntem; hem sınıflar arası hem de sınıf içi dengesizliklere çözüm sunabilmek için her sınıfı kendi yapısına uygun olarak kümeleme ile başlar, belirlenen on beş senaryo dâhilinde bu kümeler üzerinde çoğunluk sınıf için veri azaltma uygulanırken azınlık sınıf için veri artırma ya da sentetik veri üretme yöntemlerini uygular ve elde edilen yeni veri kümesi üzerinde ağaç tabanlı bir dağıtık sınıflandırma algoritması koşturur. İki farklı veri kümesi üzerinde gerçekleştirilen analizler ve testler doğrultusunda F_1-score üzerinde %3 ila %7 puanlık performans artışı elde edilmiştir. Artışa sebep olan senaryo, az kayıt sayısına sahip veri kümesi için sadece azınlık sınıfın artırılması iken çok kayıtlı veri kümesi için hem çoğunluk sınıfın azaltılması hem de azınlık sınıfın artırılması olarak tespit edilmiştir. Dolandırıcılık tespit sistemlerinde gerçekleştirilen iyileştirmelerin oranı küçük dahi olsa, sağlayacağı maddi kazancın boyutu da dikkate alınmalıdır. Bu yöntemin, veri kümesinin boyutu arttıkça daha fazla performans artışı sağladığı ve sınıfların yakınlaştığı sınır noktalarında daha iyi örnekleme gerçekleştirdiği görülmüştür. Bir senaryonun diğerinden daha iyi performans göstermesinin tek yolunun, çözümün problemin yapısına göre özelleştirilmiş olması gerektiği sonucu çıkarılmıştır. Geliştirilen yöntem, ağaç tabanlı olduğu için, açıklanabilirliğini sağlamak amacıyla özniteliklerin önem düzeyleri hesaplanmış ve ağacın ürettiği kurallar çıkarılmıştır.

Üçüncü yöntem ise, *"tekil özelliklerin dolandırıcılık tespitine etkisini modellemek yerine, özelliklerin ikili ilişkisinin etkisini ortaya koyan bir yöntem geliştirilebilir mi?"* sorusuna cevaben geliştirilmiş ve *"Görüntüye Dönüştürme İle Dolandırıcılık Tespiti"* olarak isimlendirilmiş, farklı bir perspektif sunularak dolandırıcılık tespiti problemi görüntü sınıflandırma problemine dönüştürülerek çözülmüştür. Kredi kartı işlemleri, ayrık olaylar dizisinden oluşan bir zaman serisi olarak kabul edilmiş ve GAF tekniği kullanılarak bu zaman serisi kutupsal koordinat sisteminde temsil edilmiştir. Kutupsal koordinat sisteminde yer aldıkları açıların simetri matrisine dönüştürülmesi ile de belirli bir renk ölçeğindeki piksellerden oluşan görüntüler elde edilmiştir. Daha sonra, bu görüntüler bir evrişimli sinir ağı mimarisi tasarlanarak dolandırıcı veya meşru olarak sınıflandırılmıştır. İki farklı veri kümesi üzerinde gerçekleştirilen analizler, literatürdeki aynı veri kümelerinin kullanıldığı ilgili çalışmalardan F_1-score bağlamında daha üstün sonuç üreterek %59,47 ve %85,49 başarı elde edilmiştir. Bu durum önerilen yöntemin, gerçek pozitifi yani yüksek maliyetle

ilişkili olan dolandırıcı örnekleri tespit etmede daha iyi olduğunu göstermektedir. Ayrıca, geliştirilebilecek yeni çözümler için de altyapı oluşturmaktadır.

Son olarak, *"ilk bakışta anlamlandırılması zor olan GAF görüntüleri ile oluşturulmuş kara kutu modelleri için yapay zekâ açıklanabilirliği sağlanabilir mi?*" araştırma sorusu doğrultusunda yeni bir açıklanabilirlik yaklaşımı geliştirilmiştir. *"GAF için Grad-CAM"* olarak isimlendirilen yaklaşımda, belirli bir ölçekteki renk haritalarından oluşan GAF görüntüleri üzerindeki ilişkiler daha net ortaya çıkarılmış ve sınıflandırma üzerinde daha fazla etkisi olan ilişkileri ayırt etmek için görüntülere eşikleme uygulanarak odaklanılmış görüntüler elde edilmiştir. En çok önem addedilen ilişkileri tespit etmek ve görselleştirmek için de, test kümesindeki her sınıfın odaklanılmış görüntüleri birleştirilmiştir. Bu yaklaşım ile, hem tüm test verisi özetlenerek global bir açıklanabilirlik sağlanmış hem de elde edilen odaklanılmış ısı haritası sayesinde modelin dolandırıcı ve meşru sınıf ayrımını sağlarken kullandığı en etkili ikili ilişkiler ortaya çıkarılmıştır. Bu yaklaşımın doğruluğu ise, üçüncü yönteme ait sonuçlar üzerinde test edilerek değerlendirilmiştir.

Karşılaştırmalı olarak geliştirilen üç yöntem ele alındığında, her yöntem farklı gereksinimlere çözüm üretmektedir. Dinamik İmzalar İle Dağıtık Dolandırıcılık Tespiti, etiket ediniminin zor olduğu veya etiket kirliliğinin olabileceği veri kümelerinde tespit sağlamak için ideal bir yöntem olacaktır. Dağıtık Yeniden Örnekleme İle Dolandırıcılık Tespiti, hem sınıf dengesizliğinin hem de veri hacminin oldukça yüksek olduğu veri kümelerinde daha iyi başarım artışı sağlaması beklenen bir yöntemdir. Görüntüye Dönüştürme İle Dolandırıcılık Tespiti ise, özellik sayısının az ve kayıt sayısının fazla olduğu durumlarda, daha iyi bir sınıflandırma için bütün ikili özellik ilişkilerinin devreye sokulduğu bir yöntemdir. Özellik sayısının fazla olduğu durumlarda görüntü boyutu da yükseleceği için, yöntem analiz zamanının artmasına yol açacaktır. Önerilen yöntemler karmaşıklık açısından ele alındığında; bir MapReduce görevinin karmaşıklığının; anahtar-değer çiftlerinin azami boyutuna, çalışma süresine, bellek kullanımına ve sayısına [358], bir CNN modelinin karmaşıklığının ise; katmanların sayısına, katmanlardaki filtrelerin sayısına, giriş kanallarının sayısına ve çıktı özellik haritasının boyutuna [359] bağlı olduğu bilinmektedir. Dolayısıyla modellerin karmaşıklığı hakkında net bir karşılaştırma yapmak pek mümkün değildir. Son olarak, yöntemler gerekli uyarlamaların yapılması dâhilinde gerçek zamanlı analiz açısından değerlendirildiği zaman; Dinamik İmzalar İle Dağıtık Dolandırıcılık Tespiti mini toplu işlemler barındırdığından yarı-gerçek zamanlı, Dağıtık Yeniden Örnekleme İle

Dolandırıcılık Tespiti verinin tamamına ihtiyaç duyduğundan toplu ve Görüntüye Dönüştürme İle Dolandırıcılık Tespiti ise gerçek zamanlı analizler için çözüm sunan yöntemler olabilecektir.

Dolandırıcılık tespiti amacıyla yöntem geliştirme sürecinde karşılaşılan zorluklar ve bunların aşılmasına yönelik öneriler aşağıda verilmiştir.

1. Veri Kümesi Sınırlamaları: Gizlilik endişeleri nedeniyle, dolandırıcılık tespiti üzerinde çalışılabilecek açık veri kümeleri oldukça sınırlıdır ve bu veri kümelerindeki örnek sayısı da oldukça azdır.
2. Yüksek Veri Hacmi: Dolandırıcılık tespit sistemleri tarafından işlenecek işlem sayısı genellikle çok fazladır. Örüntü bulmayı ve arama yapmayı zorlaştıran bu durum aynı zamanda tespit süresinin uzamasına da yol açmaktadır.
3. Etiket Edinimi: İşlemlerin dolandırıcı olarak etiketlenmesi genellikle gözetim sistemleri ya da müşteri şikâyetleri sayesinde gerçekleştirilir. Kimi dolandırıcılık faaliyetleri ise, tespit edilemediğinden veri kümelerinde meşruymuş gibi yer almaktadır.
4. Sınıf Dengesizliği: Dolandırıcılık faaliyeti olduğu bilinen işlemlerin oranı, doğası gereği bütün işlemlere nazaran azdır. Üstelik bu oran, işlem sayılarının gün geçtikçe arttığı büyük veri çağında daha da azalmaktadır.
5. Kavram Sapması: Hem meşru kullanıcıların hem de dolandırıcılık faaliyetleri gerçekleştiren kullanıcıların davranışları zamanla değişmektedir ve hatta bazen çakışmaktadır.
6. Model Gürbüzlüğü: Dolandırıcılık tespit modeli geliştirilirken, genellikle hem meşru hem de hileli faaliyetlerdeki değişimleri kaldırabilecek esnekliğin sağlanması oldukça zordur.
7. Değerlendirme Ölçütü: Doğruluk gibi kimi ölçütler, sınıflar arası ve sınıf içi dengesizlikler barındıran dolandırıcılık tespiti modelleri için uygun değillerdir ve genellikle çoğunluk sınıfa doğru yanlılık gösterirler.
8. Yanlış Sınıflandırma Maliyeti: Gerçekleşen her işlemin farklı bir maliyeti vardır. Kimi zaman gözden kaçan bir dolandırıcılık faaliyeti, tespit edilen birkaç faaliyetten daha maliyetli olabilmektedir.
9. Açıklanabilirlik İhtiyacı: Dolandırıcılık tespit modellerinin, aşırı uydurma ya da yetersiz uydurma sorunları barındırma yatkınlığı fazla olduğu için, açıklanabilirlik ile model işleyişinin net veya anlaşılması kolay hale getirilmesi gerekmektedir.

Tüm bu zorluklar göz önünde bulundurularak çözümlerin üretildiği bu tez kapsamında, büyük veri ve derin öğrenme yaklaşımları bağlamında güncel teknik ve teknolojiler kullanılarak dolandırıcılık tespiti yöntemleri geliştirilmiş ve bu yöntemlerin çıktıları, performans değerlendirmelerine ek olarak açıklanabilir yapay zekâ açısından incelenmiştir. Literatüre farklı bakış açıları ile yaklaşarak katkı sunan bu yöntemler her ne kadar telekom ve kredi kartı dolandırıcılığı özelinde uygulanmış olsa da, diğer dolandırıcılık tespiti türlerinde ya da farklı uygulama alanlarında kolayca uyarlanabilecek özelliklere sahiptirler.

Gelecek çalışmalarda, farklı veri kaynaklarından beslenen, karmaşıklığı ve hacmi daha yüksek olan veri kümeleri kullanılarak ölçeklenebilir yöntemler tasarlanabilir. Bu sayede, daha önceden keşfedilmemiş örüntüleri ve değerleri ortaya çıkaran gürbüz modeller geliştirilebilir. Öğrenmeyi öğrenmek olarak da bilinen meta-öğrenme gibi, birkaç eğitim örneği ile yeni beceriler öğrenebilen veya yeni ortamlara hızla uyum sağlayabilen [360] modeller de problemin çözümünde kullanılabilir. Bu modellerin değerlendirilmesi için ise, açıklanabilir yapay zekânın da bir faktör olarak ele alındığı yeni değerlendirme ölçütleri geliştirilebilir.

KAYNAKLAR

1. Krawczyk, B. (2016). Learning from imbalanced data: open challenges and future directions. *Progress in Artificial Intelligence*, 5(4), 221-232.

2. Haixiang, G., Yijing, L., Shang, J., Mingyun, G., Yuanyue, H., and Bing, G. (2017). Learning from class-imbalanced data: Review of methods and applications. *Expert Systems with Applications*, 73, 220-239.

3. Thabtah, F., Hammoud, S., Kamalov, F., and Gonsalves, A. (2020). Data imbalance in classification: Experimental evaluation. *Information Sciences*, 513, 429-441.

4. Fernández, A., del Río, S., Chawla, N. V., and Herrera, F. (2017). An insight into imbalanced big data classification: outcomes and challenges. *Complex & Intelligent Systems*, 3(2), 105-120.

5. Laleh, N., and Azgomi, M. A. (2009). A taxonomy of frauds and fraud detection techniques. *International Conference on Information Systems, Technology and Management*, Ghaziabad, India, 256-267.

6. Flegel, U., Vayssiere, J., and Bitz, G. (2010). A state of the art survey of fraud detection technology. *Insider Threats in Cyber Security*, 49, 73-84.

7. Abdallah, A., Maarof, M. A., and Zainal, A. (2016). Fraud detection system: A survey. *Journal of Network and Computer Applications*, 68, 90-113.

8. Ackoff, R. L. (1989). From data to wisdom. *Journal of Applied Systems Analysis*, 16(1), 3-9.

9. Rowley, J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. *Journal of Information Science*, 33(2), 163-180.

10. Murtagh, F., and Devlin, K. (2018). The Development of Data Science: Implications for Education, Employment, Research, and the Data Revolution for Sustainable Development. *Big Data and Cognitive Computing*, 2(2), 1-16.

11. Kotu, V., and Deshpande, B. (2018). *Data Science: Concepts and Practice* (Second edition). Cambridge: Morgan Kaufmann, 4.

12. Cao, L. (2017). Data science: a comprehensive overview. *ACM Computing Surveys*, 50(3), 1-42.

13. Gehl, R. W. (2015). Sharing, knowledge management and big data: A partial genealogy of the data scientist. *European Journal of Cultural Studies*, 18(4-5), 413-428.

14. Provost, F., and Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data*, 1(1), 51-59.

15. Kotu, V., and Deshpande, B. (2018). *Data Science: Concepts and Practice* (Second edition). Cambridge: Morgan Kaufmann, 19.

16. Vicario, G., and Coleman, S. (2020). A review of data science in business and industry and a future view. *Applied Stochastic Models in Business and Industry*, 36(1), 6-18.

17. Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI magazine*, 17(3), 37-54.

18. Wirth, R., and Hipp, J. (2000). CRISP-DM: Towards a standard process model for data mining. *International conference on the practical applications of knowledge discovery and data mining*, London, UK, 29-39.

19. De Mast, J., and Lokkerbol, J. (2012). An analysis of the Six Sigma DMAIC method from the perspective of problem solving. *International Journal of Production Economics*, 139(2), 604-614.

20. Shafique, U., and Qaiser, H. (2014). A comparative study of data mining process models (KDD, CRISP-DM and SEMMA). *International Journal of Innovation and Scientific Research*, 12(1), 217-222.

21. Larose, C.D., and Larose, D.T. (2019). *Data Science Using Python and R*. Hoboken:John Wiley & Sons, 2.

22. Kilkenny, M. F., and Robinson, K. M. (2018). Data quality:"Garbage in–garbage out". *Health Information Management Journal*, 47(3), 103-105.

23. Cai, L., and Zhu, Y. (2015). The challenges of data quality and data quality assessment in the big data era. *Data Science Journal*, 14(2), 1-10.

24. Das, S., Datta, S., and Chaudhuri, B. B. (2018). Handling data irregularities in classification: Foundations, trends, and future challenges. *Pattern Recognition*, 81, 674-693.

25. García, S., Ramírez-Gallego, S., Luengo, J., Benítez, J. M., and Herrera, F. (2016). Big data preprocessing: methods and prospects. *Big Data Analytics*, 1(9), 1-22.

26. Ali, A., Shamsuddin, S. M., and Ralescu, A. L. (2015). Classification with class imbalance problem: a review. *International Journal of Advances in Soft Computing and its Applications*, 7(3), 176-204.

27. López, V., Fernández, A., García, S., Palade, V., and Herrera, F. (2013). An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics. *Information Sciences*, 250, 113-141.

28. Gao, J., Ding, B., Fan, W., Han, J., and Philip, S. Y. (2008). Classifying data streams with skewed class distributions and concept drifts. *IEEE Internet Computing*, 12(6), 37-49.

29. Zhu, X., and Wu, X. (2004). Class noise vs. attribute noise: A quantitative study. *Artificial Intelligence Review*, 22(3), 177-210.

30. Dong, Y., and Peng, C. Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus*, 2, 1-17.

31. Hsu, A. S., Horng, A., Griffiths, T. L., and Chater, N. (2017). When absence of evidence is evidence of absence: Rational inferences from absent data. *Cognitive Science*, 41, 1155-1167.

32. Loukides, M. (2011). *What is data science?*. Boston: O'Reilly Media, 4.

33. Lv, Z., Song, H., Basanta-Val, P., Steed, A., and Jo, M. (2017). Next-generation big data analytics: State of the art, challenges, and future research topics. *IEEE Transactions on Industrial Informatics*, 13(4), 1891-1899.

34. İnternet: Laney, D., 3D data management: Controlling data volume, velocity and variety. *META Group*, URL: https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf , Son Erişim Tarihi: 09.09.2020.

35. Buyya, R., Calheiros, R. N., and Dastjerdi, A. V. (2016). *Big data: principles and paradigms*. Cambridge: Morgan Kaufmann, 3-38.

36. Alsaig, A., Alagar, V., and Ormandjieva, O. (2018). A critical analysis of the V-model of big data. *International Conference On Trust, Security And Privacy In Computing And Communications/ International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*, New York, USA, 1809-1813.

37. Hu, H., Wen, Y., Chua, T. S., and Li, X. (2014). Toward scalable systems for big data analytics: A technology tutorial. *IEEE Access*, 2, 652-687.

38. Mikalef, P., Pappas, I. O., Krogstie, J., and Giannakos, M. (2018). Big data analytics capabilities: a systematic literature review and research agenda. *Information Systems and e-Business Management*, 16(3), 547-578.

39. Rodríguez-Mazahua, L., Rodríguez-Enríquez, C. A., Sánchez-Cervantes, J. L., Cervantes, J., García-Alcaraz, J. L., and Alor-Hernández, G. (2016). A general perspective of Big Data: applications, tools, challenges and trends. *The Journal of Supercomputing*, 72(8), 3073-3113.

40. Akerkar, R. (2013). *Big Data Computing*. Boca Raton: CRC Press, 107-111.

41. Chen, C. P., and Zhang, C. Y. (2014). Data-intensive applications, challenges, techniques and technologies: A survey on Big Data. *Information Sciences*, 275, 314-347.

42. L'heureux, A., Grolinger, K., Elyamany, H. F., and Capretz, M. A. (2017). Machine learning with big data: Challenges and approaches. *IEEE Access*, 5, 7776-7797.

43. Nasser, T., and Tariq, R. S. (2015). Big data challenges. *Journal of Computer Engineering & Information Technology*, 4(3), 1-10.

44. Grover, V., Chiang, R. H., Liang, T. P., and Zhang, D. (2018). Creating strategic business value from big data analytics: A research framework. *Journal of Management Information Systems*, 35(2), 388-423.

45. Jackson, J. C., Vijayakumar, V., Quadir, M. A., and Bharathi, C. (2015). Survey on programming models and environments for cluster, cloud, and grid computing that defends big data. *Procedia Computer Science*, 50, 517-523.

46. Dean, J., and Ghemawat, S. (2008). MapReduce: simplified data processing on large clusters. *Communications of the ACM*, 51(1), 107-113.

47. Li, R., Hu, H., Li, H., Wu, Y., and Yang, J. (2016). MapReduce parallel programming model: a state-of-the-art survey. *International Journal of Parallel Programming*, 44(4), 832-866.

48. Lee, K. H., Lee, Y. J., Choi, H., Chung, Y. D., and Moon, B. (2012). Parallel data processing with MapReduce: a survey. *ACM SIGMOD Record*, 40(4), 11-20.

49. Singh, D., and Reddy, C. K. (2015). A survey on platforms for big data analytics. *Journal of Big Data*, 1(8), 1-20.

50. Zaharia, M., Chowdhury, M., Franklin, M. J., Shenker, S., and Stoica, I. (2010). Spark: Cluster computing with working sets. *USENIX conference on Hot topics in cloud computing*, Boston, United States, 1-7.

51. Chambers, B., and Zaharia, M. (2018). *Spark: The definitive guide: Big data processing made simple*. Boston: O'Reilly Media, 14.

52. Salloum, S., Dautov, R., Chen, X., Peng, P. X., and Huang, J. Z. (2016). Big data analytics on Apache Spark. *International Journal of Data Science and Analytics*, 1(3-4), 145-164.

53. Meng, X., Bradley, J., Yavuz, B., Sparks, E., Venkataraman, S., Liu, D., Freeman, J.,Tsai, D. B., Made, M.,Owen, S., Xin, D., Xin, R., Franklin, M.J., Zadeh, R., Zaharia, M., and Talwalkar, A. (2016). MLlib: Machine learning in apache spark. *Journal of Machine Learning Research*, 34, 1-7.

54. Oussous, A., Benjelloun, F. Z., Lahcen, A. A., and Belfkih, S. (2018). Big Data technologies: A survey. *Journal of King Saud University-Computer and Information Sciences*, 30(4), 431-448.

55. Mohamed, A., Najafabadi, M. K., Wah, Y. B., Zaman, E. A. K., and Maskat, R. (2020). The state of the art and taxonomy of big data analytics: view from new big data framework. *Artificial Intelligence Review*, 53(2), 989-1037.

56. Zaharia, M., Xin, R. S., Wendell, P., Das, T., Armbrust, M., Dave, A., Meng, X., Rosen, J., Venkataraman, S., Franklin, M. J., Ghodsi A., Gonzalez, J., Shenker, S., and

Stoica I. (2016). Apache spark: a unified engine for big data processing. *Communications of the ACM*, 59(11), 56-65.

57. Fernández, A., del Río, S., López, V., Bawakid, A., del Jesus, M. J., Benítez, J. M., and Herrera, F. (2014). Big Data with Cloud Computing: an insight on the computing environment, MapReduce, and programming frameworks. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 4(5), 380-409.

58. Duan, Y., Edwards, J. S., and Dwivedi, Y. K. (2019). Artificial intelligence for decision making in the era of Big Data–evolution, challenges and research agenda. *International Journal of Information Management*, 48, 63-71.

59. Joshi, A. V. (2020). *Machine Learning and Artificial Intelligence*. Switzerland: Springer, 3-7.

60. Lu, Y. (2019). Artificial intelligence: a survey on evolution, models, applications and future trends. *Journal of Management Analytics*, 6(1), 1-29.

61. Sivarajah, U., Kamal, M. M., Irani, Z., and Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 70, 263-286.

62. Landset, S., Khoshgoftaar, T. M., Richter, A. N., and Hasanin, T. (2015). A survey of open source tools for machine learning with big data in the Hadoop ecosystem. *Journal of Big Data*, 2(24), 1-36.

63. Johnston, R. (2019). *Foundations of AI: A Framework For AI in Mining*. Canada: Global Mining Guidelines Group, 5-6.

64. Hilbert, M. (2016). Big data for development: A review of promises and challenges. *Development Policy Review*, 34(1), 135-174.

65. Qiu, J., Wu, Q., Ding, G., Xu, Y., and Feng, S. (2016). A survey of machine learning for big data processing. *EURASIP Journal on Advances in Signal Processing*, 2016(67), 1-16.

66. Lee, I. (2017). Big data: Dimensions, evolution, impacts, and challenges. *Business Horizons*, 60(3), 293-303.

67. Zhou, L., Pan, S., Wang, J., and Vasilakos, A. V. (2017). Machine learning on big data: Opportunities and challenges. *Neurocomputing*, 237, 350-361.

68. Gupta, P., Sharma, A., and Jindal, R. (2016). Scalable machine- learning algorithms for big data analytics: a comprehensive review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 6(6), 194-214.

69. Samek, W., Montavon, G., Vedaldi, A., Hansen, L. K., and Müller, K. R. (2019). *Explainable AI: interpreting, explaining and visualizing deep learning*. Switzerland: Springer, 5-22.

70. Sheh, R., and Monteath, I. (2018). Defining explainable ai for requirements analysis. *Künstliche Intelligenz*, 32(4), 261-266.

71. Singh, A., Sengupta, S., and Lakshminarayanan, V. (2020). Explainable deep learning models in medical image analysis. *Journal of Imaging*, 6(52), 1-19.

72. Preece, A. (2018). Asking 'Why'in AI: Explainability of intelligent systems–perspectives and challenges. *Intelligent Systems in Accounting, Finance and Management*, 25(2), 63-72.

73. Lipton, Z. C. (2018). The mythos of model interpretability. Queue, 16(3), 31-57.

74. Rai, A. (2020). Explainable AI: From black box to glass box. *Journal of the Academy of Marketing Science*, 48(1), 137-141.

75. Adadi, A., and Berrada, M. (2018). Peeking inside the black-box: A survey on Explainable Artificial Intelligence (XAI). *IEEE Access*, 6, 52138-52160.

76. Gunning, D., and Aha, D. W. (2019). DARPA's explainable artificial intelligence program. *AI Magazine*, 40(2), 44-58.

77. Arrieta, A. B., Díaz-Rodríguez, N., Del Ser, J., Bennetot, A., Tabik, S., Barbado, A., Garcia, S., Gil-Lopez, S., Molina,D., Benjamins, R., Chatila, R., and Herrera, F. (2020). Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Information Fusion*, 58, 82-115.

78. Hagras, H. (2018). Toward human-understandable, explainable AI. *Computer*, 51(9), 28-36.

79. Mohseni, S., Zarei, N., and Ragan, E. D. (2018). A Multidisciplinary Survey and Framework for Design and Evaluation of Explainable AI Systems. *arXiv-1811*.

80. Das, A., and Rad, P. (2020). Opportunities and Challenges in Explainable Artificial Intelligence (XAI): A Survey. *arXiv:2006.11371*.

81. Sokol, K., and Flach, P. (2020). Explainability fact sheets: a framework for systematic assessment of explainable approaches. *ACM Conference on Fairness, Accountability, and Transparency*, New York, United States, 56-67.

82. Samek, W., Wiegand, T., and Müller, K. R. (2017). Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. *arXiv:1708.08296*.

83. He, H., and Garcia, E. A. (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering*, 21(9), 1263-1284.

84. Hasanin, T., Khoshgoftaar, T. M., Leevy, J. L., and Bauder, R. A. (2019). Severely imbalanced Big Data challenges: investigating data sampling approaches. *Journal of Big Data*, 6(107), 1-25.

85. Sun, Y., Wong, A. K., and Kamel, M. S. (2009). Classification of imbalanced data: A review. *International Journal of Pattern Recognition and Artificial Intelligence*, 23(04), 687-719.

86. Luque, A., Carrasco, A., Martín, A., and Heras, A. (2019). The impact of class imbalance in classification performance metrics based on the binary confusion matrix. *Pattern Recognition*, 91, 216-231.

87. García, V., Sánchez, J. S., and Mollineda, R. A. (2012). On the effectiveness of preprocessing methods when dealing with different levels of class imbalance. *Knowledge-Based Systems*, 25(1), 13-21.

88. Japkowicz, N., and Stephen, S. (2002). The class imbalance problem: A systematic study. *Intelligent Data Analysis*, 6(5), 429-449.

89. Hart, P. (1968). The condensed nearest neighbor rule. *IEEE Transactions on Information Theory*, 14(3), 515-516.

90. Tomek, I. (1976). Two modifications of CNN. *IEEE Transactions on Systems, Man, and Cybernetics*, 6(11), 769-772.

91. Yen, S. J., and Lee, Y. S. (2009). Cluster-based under-sampling approaches for imbalanced data distributions. *Expert Systems with Applications*, 36(3), 5718-5727.

92. García, S., and Herrera, F. (2009). Evolutionary undersampling for classification with imbalanced datasets: Proposals and taxonomy. *Evolutionary Computation*, 17(3), 275-306.

93. Wong, G. Y., Leung, F. H., and Ling, S. H. (2014). An under-sampling method based on fuzzy logic for large imbalanced dataset. *IEEE International Conference on Fuzzy Systems*, Beijing, China, 1248-1252.

94. Ng, W. W., Hu, J., Yeung, D. S., Yin, S., and Roli, F. (2014). Diversified sensitivity-based undersampling for imbalance classification problems. *IEEE Transactions on Cybernetics*, 45(11), 2402-2412.

95. Fu, Y., Zhang, H., Bai, Y., and Sun, W. (2016). An under-sampling method: Based on principal component analysis and comprehensive evaluation model. *IEEE International Conference on Software Quality, Reliability and Security Companion*, Vienna, Austria, 414-415.

96. Ha, J., and Lee, J. S. (2016). A new under-sampling method using genetic algorithm for imbalanced data classification. *ACM International Conference on Ubiquitous Information Management and Communication*, New York, United States, 1-6.

97. Devi, D., and Purkayastha, B. (2017). Redundancy-driven modified Tomek-link based undersampling: a solution to class imbalance. *Pattern Recognition Letters*, 93, 3-12.

98. Bunkhumpornpat, C., and Sinapiromsaran, K. (2017). DBMUTE: density-based majority under-sampling technique. *Knowledge and Information Systems*, 50(3), 827-850.

99. Yap, B. W., Abd Rani, K., Abd Rahman, H. A., Fong, S., Khairudin, Z., and Abdullah, N. N. (2013). An application of oversampling, undersampling, bagging and boosting in handling imbalanced datasets. *International Conference on Advanced Data and Information Engineering*, Kuala Lumpur, Malaysia, 13-22.

100. Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P. (2002). SMOTE: synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, 16, 321-357.

101. Han, H., Wang, W. Y., and Mao, B. H. (2005). Borderline-SMOTE: a new over-sampling method in imbalanced data sets learning. *IEEE International Conference on Intelligent Computing*, Hefei, China, 878-887.

102. He, H., Bai, Y., Garcia, E. A., and Li, S. (2008). ADASYN: Adaptive synthetic sampling approach for imbalanced learning. *IEEE International Joint Conference on Neural Networks*, Hong Kong, China, 1322-1328.

103. Maciejewski, T., and Stefanowski, J. (2011). Local neighbourhood extension of SMOTE for mining imbalanced data. *IEEE Symposium on Computational Intelligence and Data Mining*, Paris, France, 104-111.

104. Bunkhumpornpat, C., Sinapiromsaran, K., and Lursinsap, C. (2009). Safe-level-smote: Safe-level-synthetic minority over-sampling technique for handling the class imbalanced problem. *Pacific-Asia conference on Knowledge Discovery and Data Mining*, Bangkok, Thailand, 475-482.

105. Bunkhumpornpat, C., Sinapiromsaran, K., and Lursinsap, C. (2012). DBSMOTE: density-based synthetic minority over-sampling technique. *Applied Intelligence*, 36(3), 664-684.

106. Abdi, L., and Hashemi, S. (2015). To combat multi-class imbalanced problems by means of over-sampling techniques. *IEEE Transactions on Knowledge and Data Engineering*, 28(1), 238-251.

107. Ramentol, E., Caballero, Y., Bello, R., and Herrera, F. (2012). SMOTE-RSB*: a hybrid preprocessing approach based on oversampling and undersampling for high imbalanced data-sets using SMOTE and rough sets theory. *Knowledge and Information Systems*, 33(2), 245-265.

108. Wang, Q. (2014). A hybrid sampling SVM approach to imbalanced data classification. *Abstract and Applied Analysis*, 972786, 1-7.

109. Prachuabsupakij, W. (2015). CLUS: A new hybrid sampling classification for imbalanced data. *IEEE International Joint Conference on Computer Science and Software Engineering*, Hatyai, Thailand, 281-286.

110. Jian, C., Gao, J., and Ao, Y. (2016). A new sampling method for classifying imbalanced data based on support vector machine ensemble. *Neurocomputing*, 193, 115-122.

111. Sáez, J. A., Luengo, J., Stefanowski, J., and Herrera, F. (2015). SMOTE–IPF: Addressing the noisy and borderline examples problem in imbalanced classification by a re-sampling method with filtering. *Information Sciences*, 291, 184-203.

112. Kang, Q., Chen, X., Li, S., and Zhou, M. (2016). A noise-filtered under-sampling scheme for imbalanced classification. *IEEE Transactions on Cybernetics*, 47(12), 4263-4274.

113. Junsomboon, N., and Phienthrakul, T. (2017). Combining over-sampling and under-sampling techniques for imbalance dataset. *ACM International Conference on Machine Learning and Computing*, Ho Chi Minh City, Vietnam, 243-247.

114. Gónzalez, S., García, S., Lázaro, M., Figueiras-Vidal, A. R., and Herrera, F. (2017). Class switching according to nearest enemy distance for learning from highly imbalanced data-sets. *Pattern Recognition*, 70, 12-24.

115. Ling, C. X., and Sheng, V. S. (2008). Cost-sensitive learning and the class imbalance problem. *Encyclopedia of machine learning*, 2011, 231-235.

116. Cheng, F., Zhang, J., Wen, C., Liu, Z., and Li, Z. (2017). Large cost-sensitive margin distribution machine for imbalanced data classification. *Neurocomputing*, 224, 45-57.

117. Xiao, W., Zhang, J., Li, Y., Zhang, S., and Yang, W. (2017). Class-specific cost regulation extreme learning machine for imbalanced classification. *Neurocomputing*, 261, 70-82.

118. Ohsaki, M., Wang, P., Matsuda, K., Katagiri, S., Watanabe, H., and Ralescu, A. (2017). Confusion-matrix-based kernel logistic regression for imbalanced data classification. *IEEE Transactions on Knowledge and Data Engineering*, 29(9), 1806-1819.

119. Chen, Y., and Mani, S. (2011). Active learning for unbalanced data in the challenge with multiple models and biasing. *Active Learning and Experimental Design Workshop*, Sardinia, Italy, 113-126.

120. He, H., and Ma, Y. (2013). *Imbalanced Learning: Foundations, Algorithms, and Applications.* Hoboken: John Wiley & Sons, 101-149.

121. Ferdowsi, Z., Ghani, R., and Settimi, R. (2013). Online active learning with imbalanced classes. *IEEE International Conference on Data Mining*, Dallas, United States, 1043-1048.

122. You, X., Wang, R., and Tao, D. (2014). Diverse expected gradient active learning for relative attributes. *IEEE Transactions on Image Processing*, 23(7), 3203-3217.

123. Zhang, J., Wu, X., and Shengs, V. S. (2014). Active learning with imbalanced multiple noisy labeling. *IEEE Transactions on Cybernetics*, 45(5), 1095-1107.

124. Guo, H., and Wang, W. (2015). An active learning-based SVM multi-class classification model. *Pattern Recognition*, 48(5), 1577-1597.

125. Zhang, X., Yang, T., and Srinivasan, P. (2016). Online asymmetric active learning with imbalanced data. *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, United States, 2055-2064.

126. Khanchi, S., Heywood, M. I., and Zincir-Heywood, A. N. (2017). Properties of a GP active learning framework for streaming data with class imbalance. *ACM Genetic and Evolutionary Computation Conference*, Berlin, Germany 945-952.

127. Wu, S., and Amari, S. I. (2002). Conformal transformation of kernel functions: A data-dependent way to improve support vector machine classifiers. *Neural Processing Letters*, 15(1), 59-67.

128. Zhang, Y., Fu, P., Liu, W., and Chen, G. (2014). Imbalanced data classification based on scaling kernel-based support vector machine. *Neural Computing and Applications*, 25(3-4), 927-935.

129. Li, X., Wang, L., and Sung, E. (2008). AdaBoost with SVM-based component classifiers. *Engineering Applications of Artificial Intelligence*, 21(5), 785-795.

130. Wu, G., and Chang, E. Y. (2005). KBA: Kernel boundary alignment considering imbalanced data distribution. *IEEE Transactions on Knowledge and Data Engineering*, 17(6), 786-795.

131. Maratea, A., and Petrosino, A. (2011). Asymmetric kernel scaling for imbalanced data classification. *International Workshop on Fuzzy Logic and Applications*, Trani, Italy, 196-203.

132. Maratea, A., Petrosino, A., and Manzo, M. (2014). Adjusted F-measure and kernel scaling for imbalanced data learning. *Information Sciences*, 257, 331-341.

133. Galar, M., Fernandez, A., Barrenechea, E., Bustince, H., and Herrera, F. (2011). A review on ensembles for the class imbalance problem: bagging-, boosting-, and hybrid-based approaches. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 42(4), 463-484.

134. Chawla, N. V., Lazarevic, A., Hall, L. O., and Bowyer, K. W. (2003). SMOTEBoost: Improving prediction of the minority class in boosting. *European Conference on Principles of Data Mining and Knowledge Discovery*, Prague, Czech Republic, 107-119.

135. Guo, H., and Viktor, H. L. (2004). Learning from imbalanced data sets with boosting and data generation: the databoost-im approach. *ACM SIGKDD Explorations Newsletter*, 6(1), 30-39.

136. Mease, D., Wyner, A., and Buja, A. (2007). Cost-weighted boosting with jittering and over/under-sampling: Jous-boost. *Journal of Machine Learning Research*, 8, 409-439.

137. Seiffert, C., Khoshgoftaar, T. M., Van Hulse, J., and Napolitano, A. (2009). RUSBoost: A hybrid approach to alleviating class imbalance. *IEEE Transactions on Systems, Man, and Cybernetics, Part A: Systems and Humans*, 40(1), 185-197.

138. Wang, S., and Yao, X. (2009). Diversity analysis on imbalanced data sets by using ensemble models. *IEEE Symposium on Computational Intelligence and Data Mining*, Nashville, United States, 324-331.

139. Liu, X. Y., Wu, J., and Zhou, Z. H. (2008). Exploratory undersampling for class-imbalance learning. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics,* 39(2), 539-550.

140. Chen, S., He, H., and Garcia, E. A. (2010). RAMOBoost: ranked minority oversampling in boosting. *IEEE Transactions on Neural Networks*, 21(10), 1624-1642.

141. Akbani, R., Kwek, S., and Japkowicz, N. (2004). Applying support vector machines to imbalanced datasets. *European Conference on Machine Learning*, Pisa, Italy, 39-50.

142. López, V., Fernández, A., Moreno-Torres, J. G., and Herrera, F. (2012). Analysis of preprocessing vs. cost-sensitive learning for imbalanced classification. Open problems on intrinsic data characteristics. *Expert Systems with Applications*, 39(7), 6585-6608.

143. Hsu, J. L., Hung, P. C., Lin, H. Y., and Hsieh, C. H. (2015). Applying under-sampling techniques and cost-sensitive learning methods on risk assessment of breast cancer. *Journal of Medical Systems*, 39(40), 1-13.

144. Del Río, S., López, V., Benítez, J. M., and Herrera, F. (2014). On the use of MapReduce for imbalanced big data using Random Forest. *Information Sciences*, 285, 112-137.

145. Patil, S. S., and Sonavane, S. P. (2017). Enriched over_sampling techniques for improving classification of imbalanced big data. *IEEE International Conference on Big Data Computing Service and Applications*, Boston, United States, 1-10.

146. Ghanavati, M., Wong, R. K., Chen, F., Wang, Y., and Perng, C. S. (2014). An effective integrated method for learning big imbalanced data. *IEEE International Congress on Big Data*, Washington DC, United States, 691-698.

147. Galpert, D., Del Río, S., Herrera, F., Ancede-Gallardo, E., Antunes, A., and Agüero-Chapin, G. (2015). An effective big data supervised imbalanced classification approach for ortholog detection in related yeast species. *BioMed Research International*, 748681, 1-12.

148. Triguero, I., del Río, S., López, V., Bacardit, J., Benítez, J. M., and Herrera, F. (2015). ROSEFW-RF: the winner algorithm for the ECBDL'14 big data competition: an

extremely imbalanced big data bioinformatics problem. *Knowledge-Based Systems*, 87, 69-79.

149. Triguero, I., Galar, M., Vluymans, S., Cornelis, C., Bustince, H., Herrera, F., and Saeys, Y. (2015). Evolutionary undersampling for imbalanced big data classification. *IEEE Congress on Evolutionary Computation*, Sendai, Japan, 715-722.

150. Kamal, S., Ripon, S. H., Dey, N., Ashour, A. S., and Santhi, V. (2016). A MapReduce approach to diminish imbalance parameters for big deoxyribonucleic acid dataset. *Computer Methods and Programs in Biomedicine*, 131, 191-206.

151. Bhagat, R. C., and Patil, S. S. (2015). Enhanced SMOTE algorithm for classification of imbalanced big-data using random forest. *IEEE International Advance Computing Conference*, Bangalore, India 403-408.

152. Maurya, C. K., Toshniwal, D., and Venkoparao, G. V. (2016). Online sparse class imbalance learning on big data. *Neurocomputing*, 216, 250-260.

153. Tang, M., Yang, C., Zhang, K., and Xie, Q. (2014). Cost-sensitive support vector machine using randomized dual coordinate descent method for big class-imbalanced data classification. *Abstract and Applied Analysis*, 2014, 1-9.

154. López, V., Del Río, S., Benítez, J. M., and Herrera, F. (2015). Cost-sensitive linguistic fuzzy rule based classification systems under the MapReduce framework for imbalanced big data. *Fuzzy Sets and Systems*, 258, 5-38.

155. Wang, Z., Xin, J., Yang, H., Tian, S., Yu, G., Xu, C., and Yao, Y. (2017). Distributed and weighted extreme learning machine for imbalanced big data learning. *Tsinghua Science and Technology*, 22(2), 160-173.

156. Abdel-Hamid, N. B., ElGhamrawy, S., El Desouky, A., and Arafat, H. (2018). A dynamic spark-based classification framework for imbalanced big data. *Journal of Grid Computing*, 16(4), 607-626.

157. Deng, L. (2014). A tutorial survey of architectures, algorithms, and applications for deep learning. *APSIPA Transactions on Signal and Information Processing*, 3(2), 1-29.

158. Johnson, J. M., and Khoshgoftaar, T. M. (2019). Survey on deep learning with class imbalance. *Journal of Big Data*, 6(27), 1-54.

159. Hensman, P., and Masko, D. (2015). The impact of imbalanced training data for convolutional neural networks. *KTH Royal Institute of Technology*, 1-28.

160. Lee, H., Park, M., and Kim, J. (2016). Plankton classification on imbalanced large scale database via convolutional neural networks with transfer learning. *IEEE International Conference on Image Processing*, Phoenix, United States, 3713-3717.

161. Wang, S., Liu, W., Wu, J., Cao, L., Meng, Q., and Kennedy, P. J. (2016). Training deep neural networks on imbalanced data sets. *IEEE International Joint Conference on Neural Networks*, Vancouver, Canada, 4368-4374.

162. Huang, C., Li, Y., Loy, C. C., and Tang, X. (2016). Learning deep representation for imbalanced classification. *IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, USA, 5375-5384.

163. Khan, S. H., Hayat, M., Bennamoun, M., Sohel, F. A., and Togneri, R. (2017). Cost-sensitive learning of deep feature representations from imbalanced data. *IEEE Transactions on Neural Networks and Learning Systems*, 29(8), 3573-3587.

164. Ando, S., and Huang, C. Y. (2017). Deep over-sampling framework for classifying imbalanced. *Joint European Conference on Machine Learning and Knowledge Discovery in Databases*, Skopje, Macedonia, 770-785.

165. Ding, W., Huang, D. Y., Chen, Z., Yu, X., and Lin, W. (2017). Facial action recognition using very deep networks for highly imbalanced class distribution. *IEEE Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, Kuala Lumpur, Malaysia, 1368-1372.

166. Pouyanfar, S., Tao, Y., Mohan, A., Tian, H., Kaseb, A. S., Gauen, K., Dailey, R., Aghajanzadeh, S., Lu, Y. H., Chen, S. C., and Shyu, M.L. (2018). Dynamic sampling in convolutional neural networks for imbalanced data classification. *IEEE Conference on Multimedia Information Processing and Retrieval*, Miami, United States, 112-117.

167. Buda, M., Maki, A., and Mazurowski, M. A. (2018). A systematic study of the class imbalance problem in convolutional neural networks. *Neural Networks*, 106, 249-259.

168. Zhang, Y., Shuai, L., Ren, Y., and Chen, H. (2018). Image classification with category centers in class imbalance situation. *IEEE Youth Academic Annual Conference of Chinese Association of Automation*, Nanjing, China, 359-363.

169. Free, C. (2015). Looking through the fraud triangle: A review and call for new directions. *Meditari Accountancy Research*, 23(2), 175-196.

170. Pejic-Bach, M. (2010). Profiling intelligent systems applications in fraud detection and prevention: survey of research articles. *IEEE International Conference on Intelligent Systems, Modelling and Simulation*, Liverpool, United Kingdom, 80-85.

171. Rivera, K., Rohn, C., Donker, J., and Butter, C. (2020). Fighting fraud: A never-ending battle, *PwC's Global Economic Crime and Fraud Survey*, 1-14.

172. Youngblood, J.R., (2015). *A comprehensive look at fraud identification and prevention*. Boca Raton: CRC Press, 3-8.

173. Türk Ceza Kanunu. (2004). 5237 Sayılı Kanun. URL: https://www.mevzuat.gov.tr/MevzuatMetin/1.5.5237.pdf. Son Erişim Tarihi: 09.09.2020.

174. Huang, S. Y., Lin, C. C., Chiu, A. A., and Yen, D. C. (2017). Fraud detection using fraud triangle risk factors. *Information Systems Frontiers*, 19(6), 1343-1356.

175. Dorminey, J. W., Fleming, A. S., Kranacher, M. J., and Riley Jr, R. A. (2010). Beyond the fraud triangle. *The CPA Journal*, 80(7), 17-23.

176. Prabowo, H. Y. (2011). Building our defence against credit card fraud: a strategic view. *Journal of Money Laundering Control*, 14(4), 1-16.

177. Behdad, M., Barone, L., Bennamoun, M., and French, T. (2012). Nature-inspired techniques in the context of fraud detection. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, 42(6), 1273-1290.

178. Adewumi, A. O., and Akinyelu, A. A. (2017). A survey of machine-learning and nature-inspired based credit card fraud detection techniques. *International Journal of System Assurance Engineering and Management*, 8(2), 937-953.

179. Phua, C., Lee, V., Smith, K., and Gayler, R. (2010). A comprehensive survey of data mining-based fraud detection research. *arXiv:1009.6119*.

180. Baesens, B., Van Vlasselaer, V., and Verbeke, W. (2015). *Fraud analytics using descriptive, predictive, and social network techniques: a guide to data science for fraud detection*. Hoboken: John Wiley & Sons, 15-16.

181. Patil, D. (2011). *Building Data Science Teams*. Boston: O'Reilly Media, 8-9.

182. Arif, M., and Dar, A. R. (2015). Survey on fraud detection techniques using data mining. *International Journal of u-and e-Service, Science and Technology*, 8(3), 165-170.

183. Gama, J., Žliobaitė, I., Bifet, A., Pechenizkiy, M., and Bouchachia, A. (2014). A survey on concept drift adaptation. *ACM Computing Surveys* , 46(4), 1-37.

184. Malekian, D., and Hashemi, M. R. (2013). An adaptive profile based fraud detection framework for handling concept drift. *IEEE International Iranian Society of Cryptology Conference on Information Security and Cryptology*, Yazd, Iran, 1-6.

185. Dal Pozzolo, A., Boracchi, G., Caelen, O., Alippi, C., and Bontempi, G. (2015). Credit card fraud detection and concept-drift adaptation with delayed supervised information. *IEEE International Joint Conference on Neural Networks*, Killarney, Ireland, 1-8.

186. Kulkarni, P., and Ade, R. (2015). Logistic regression learning model for handling concept drift with unbalanced data in credit card fraud detection system. *International Conference on Computer and Communication Technologies*, Hyderabad, India, 681-689.

187. Veeramachaneni, K., Arnaldo, I., Korrapati, V., Bassias, C., and Li, K. (2016). AI^ 2: training a big data machine to defend. *IEEE International Conference on Big Data Security on Cloud*, New York, United States, 49-54.

188. Somasundaram, A., and Reddy, S. (2019). Parallel and incremental credit card fraud detection model to handle concept drift and data imbalance. *Neural Computing and Applications*, 31(1), 3-14.

189. Sisodia, D. S., Reddy, N. K., and Bhandari, S. (2017). Performance evaluation of class balancing techniques for credit card fraud detection. *IEEE International Conference on Power, Control, Signals and Instrumentation Engineering*, Chennai, India, 2747-2752.

190. Pérez, J. M., Muguerza, J., Arbelaitz, O., Gurrutxaga, I., and Martín, J. I. (2005). Consolidated tree classifier learning in a car insurance fraud detection domain with class imbalance. *International Conference on Pattern Recognition and Image Analysis*, Bath, United Kingdom, 381-389.

191. Hordri, N. F., Yuhaniz, S. S., Azmi, N. F. M., and Shamsuddin, S. M. (2018). Handling class imbalance in credit card fraud using resampling methods. *International Journal of Advanced Computer Science and Applications*, 9(11), 390-396.

192. Bauder, R. A., Khoshgoftaar, T. M., and Hasanin, T. (2018). Data sampling approaches with severely imbalanced big data for medicare fraud detection. *International Conference on Tools with Artificial Intelligence*, Volos, Greece, 137-142.

193. Moepya, S. O., Akhoury, S. S., and Nelwamondo, F. V. (2014). Applying cost-sensitive classification for financial fraud detection under high class-imbalance. *IEEE International Conference on Data Mining Workshop*, Shenzhen, China, 183-192.

194. Bahnsen, A. C., Aouada, D., Stojanovic, A., and Ottersten, B. (2016). Feature engineering strategies for credit card fraud detection. *Expert Systems with Applications*, 51, 134-142.

195. Herland, M., Khoshgoftaar, T. M., and Bauder, R. A. (2018). Big data fraud detection using multiple medicare data sources. *Journal of Big Data*, 5(29), 1-21.

196. Li, Y., Yan, C., Liu, W., and Li, M. (2018). A principle component analysis-based random forest with the potential nearest neighbor method for automobile insurance fraud identification. *Applied Soft Computing*, 70, 1000-1009.

197. Shiguihara-Juarez, P., and Murrugarra-Llerena, N. (2019). Reducing Dimensionality of Variables for a Classification Problem: Fraud Detection. *IEEE Sciences and Humanities International Research Conference*, Lima, Peru, 1-4.

198. Khan, A. U. S., Akhtar, N., and Qureshi, M. N. (2014). Real-time credit-card fraud detection using artificial neural network tuned by simulated annealing algorithm. *International Conference on Recent Trends in Information, Telecommunication and Computing*, Chandigarh , India, 113-121.

199. Niu, K., Jiao, H., Deng, N., and Gao, Z. (2016). A Real-Time Fraud Detection Algorithm Based on Intelligent Scoring for the Telecom Industry. *IEEE International Conference on Networking and Network Applications*, Hakodate, Japan, 303-306.

200. Manunza, L., Marseglia, S., and Romano, S. P. (2017). Kerberos: A real-time fraud detection system for IMS-enabled VoIP networks. *Journal of Network and Computer Applications*, 80, 22-34.

201. Carcillo, F., Dal Pozzolo, A., Le Borgne, Y. A., Caelen, O., Mazzer, Y., and Bontempi, G. (2018). Scarff: a scalable framework for streaming credit card fraud detection with spark. *Information Fusion*, 41, 182-194.

202. Kamaruddin, S., and Ravi, V. (2016). Credit card fraud detection using big data analytics: use of PSOAANN based one-class classification. *ACM International Conference on Informatics and Analytics*, Pondicherry, India, 1-8.

203. Bauder, R. A., and Khoshgoftaar, T. M. (2018). The effects of varying class distribution on learner behavior for medicare fraud detection with imbalanced big data. Health Information Science and Systems, 6(9), 1-14.

204. Bauder, R., and Khoshgoftaar, T. (2018). Medicare fraud detection using random forest with class imbalanced big data. *IEEE International Conference on Information Reuse and Integration*, Salt Lake City, United States, 80-87.

205. Makki, S., Haque, R., Taher, Y., Assaghir, Z., Ditzler, G., Hacid, M. S., and Zeineddine, H. (2017). Fraud Analysis Approaches in the Age of Big Data-A Review of State of the Art. *IEEE International Workshops on Foundations and Applications of Self Systems*, Tucson, United States, 243-250.

206. Leevy, J. L., Khoshgoftaar, T. M., Bauder, R. A., and Seliya, N. (2018). A survey on addressing high-class imbalance in big data. *Journal of Big Data*, 5(42), 1-30.

207. Burnett, R., Brunstorm, A., and Nilsson A.G. (2004). Perspectives on Multimedia Communication, Media and Information Technology, West Sussex: John Wiley & Sons, 125-144.

208. Li, Z. (2018). *Telecommunication 4.0 Reinvention of the Communication Network*, Singapore: Springer, 17-19.

209. Koi-Akrofi, G. Y., Koi-Akrofi, J., Odai, D. A., and Twum, E. O. (2019). Global telecommunications fraud trend analysis. *International Journal of Innovation and Applied Studies*, 25(3), 940-947.

210. Sahin, M., Francillon, A., Gupta, P., and Ahamad, M. (2017). SoK: Fraud in telephony networks. *IEEE European Symposium on Security and Privacy*, Paris, France, 235-250.

211. Wood, L. (2020). *Intelligence Report - Key Global Telecom Industry Statistics*. Australia: Paul Budde Communication, 1-39.

212. Howell, J. (2019*). Global Telecom Fraud Survey*. New Jersey: Communications Fraud Control Association, 1-2.

213. Huang, Z. (2017). Causes and Prevention of Telecommunication Network Fraud. *International Conference on Humanities Science and Society Development,* Xiamen, China, 164-173.

214. Wallis, D., Saad, D., Zhou, H., and Lee, Y. (2015). *Data synthesis in fraud detection and prevention for telecommunications service providers*. London: Deloitte & Touche LLP, 1-6.

215. Kurgan, D., and Makela, J. (2018). *Taking Action Against Fraud Demonstrating the International Wholesale Industry's Leadership Against Telecoms Fraud*, London: Global Leaders' Forum and Delta Partners, 1-44.

216. Wilson, S., and Gibson, C. (2019). *Cyber-Telecom Crime Report 2019*, Tokyo: Trend Micro Research, 1-57.

217. Sahin, M., Relieu, M., and Francillon, A. (2017). Using chatbots against voice spam: Analyzing Lenny's effectiveness. *Symposium on Usable Privacy and Security*, Santa Clara, United States, 319-337.

218. Gibson, C. (2018). *Toll fraud, international revenue share fraud and more: How criminals monetise hacked cell phones and IoT devices for telecom fraud*. Tokyo: Trend Micro Research, 1-8.

219. Tseng, V. S., Ying, J. C., Huang, C. W., Kao, Y., and Chen, K. T. (2015). Fraudetector: A graph-mining-based framework for fraudulent phone call detection. *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, New York, United States, 2157-2166.

220. Wiens, A., Kübler, S., Wiens, T., and Massoth, M. (2015). Improvement of User Profiling, Call Destination Profiling and Behavior Pattern Recognition Approaches for Telephony Toll Fraud Detection. *International Journal on Advances in Security*, 8(1-2), 1-15.

221. Sallehuddin, R., Ibrahim, S., Zain, A. M., and Elmi, A. H. (2015). Detecting SIM box fraud by using support vector machine and artificial neural network. *Jurnal Teknologi*, 74(1), 131-143.

222. Kübler, S., Massoth, M., Wiens, A., and Wiens, T. (2015). Toll Fraud Detection in Voice over IP Networks Using Communication Behavior Patterns on Unlabeled Data. *International Conference on Networks*, 191-197.

223. Subudhi, S., and Panigrahi, S. (2016). Use of fuzzy clustering and support vector machine for detecting fraud in mobile telecommunication networks. *International Journal of Security and Networks*, 11(1-2), 3-11.

224. Subudhi, S., and Panigrahi, S. (2017). Use of Possibilistic fuzzy C-means clustering for telecom fraud detection. International Conference on Computational Intelligence in Data Mining, Bhubaneswar, India, 633-641.

225. Yao, W., Ruan, N., Yu, F., Jia, W., and Zhu, H. (2017). Privacy-preserving fraud detection via cooperative mobile carriers with improved accuracy. *IEEE International Conference on Sensing, Communication, and Networking*, San Diego, United States, 1-9.

226. Li, R., Zhang, Y., Tuo, Y., and Chang, P. (2018). A Novel Method for Detecting Telecom Fraud User. *IEEE International Conference on Information Systems Engineering*, Shanghai, China, 46-50.

227. Chouiekh, A., and Haj, E. H. I. E. (2018). Convnets for fraud detection analysis. *International Conference On Intelligent Computing in Data S*ciences, Fez, Morocco, 133-138.

228. Yan, H., Jiang, Y., and Liu, G. (2018). Telecomm Fraud Detection via Attributed Bipartite Network. *IEEE International Conference on Service Systems and Service Management*, Hangzhou, China, 1-6.

229. Zheng, Y. J., Zhou, X. H., Sheng, W. G., Xue, Y., and Chen, S. Y. (2018). Generative adversarial network based telecom fraud detection at the receiving bank. *Neural Networks*, 102, 78-86.

230. Guo, J., Liu, G., Zuo, Y., and Wu, J. (2018). Learning sequential behavior representations for fraud detection. *IEEE International Conference on Data Mining*, Singapore, Singapore, 127-136.

231. Subudhi, S., and Panigrahi, S. (2018). A hybrid mobile call fraud detection model using optimized fuzzy C-means clustering and group method of data handling-based network. *Vietnam Journal of Computer Science*, 5(3-4), 205-217.

232. Arafat, M., Qusef, A., and Sammour, G. (2019). Detection of Wangiri Telecommunication Fraud Using Ensemble Learning. *IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology*, Amman, Jordan, 330-335.

233. Zhong, R., Dong, X., Lin, R., and Zou, H. (2019). An Incremental Identification Method for Fraud Phone Calls Based on Broad Learning System. *IEEE International Conference on Communication Technology*, Xi'an, China, 1306-1310.

234. Lin, H., Liu, G., Wu, J., Zuo, Y., Wan, X., and Li, H. (2019). Fraud Detection in Dynamic Interaction Network. *IEEE Transactions on Knowledge and Data Engineering*, 1-14.

235. Chang, Q., Lin, S., and Liu, X. (2019). Stacked-SVM: A Dynamic SVM Framework for Telephone Fraud Identification from Imbalanced CDRs. *ACM International Conference on Algorithms, Computing and Artificial Intelligence*, Sanya, China, 112-120.

236. Cortes, C., and Pregibon, D. (2001). Signature-based methods for data streams. *Data Mining and Knowledge Discovery*, 5(3), 167-182.

237. Globa, L. S., Novogrudska, R. L., and Koval, A. V. (2018). Ontology Model of Telecom Operator Big Data. *IEEE International Black Sea Conference on Communications and Networking*, Batumi, Georgia, 1-5.

238. Yayah, F. C., Ghauth, K. I., and Ting, C. Y. (2017). Adopting Big Data Analytics Strategy in Telecommunication Industry. *Journal of Computer Science & Computational Mathematics*, 7(3), 57-67.

239. Rebahi, Y., Nassar, M., Magedanz, T., and Festor, O. (2011). A survey on fraud and service misuse in voice over IP (VoIP) networks. *Information Security Technical Report*, 16(1), 12-19.

240. Sengar, H. (2014). VoIP Fraud: Identifying a wolf in sheep's clothing. *ACM SIGSAC Conference on Computer and Communications Security*, Scottsdale, United States, 334-345.

241. İnternet: Nova FMS, Next generation fraud management system. *NETAŞ Telekomunikasyon A.Ş*, URL: http://www.netas.com.tr/media/102281/fms-brosur-web.pdf , Son Erişim Tarihi: 09.09.2020.

242. Abello, J., Pardalos, P. M., and Resende, M. G. (2013). *Handbook of massive data sets*, Dallas: Springer, 911-929.

243. Rebahi, Y., Thanh, T. Q., Busse, R., and Lorenz, P. (2014). On the Use of Unsupervised Techniques for Fraud Detection in VoIP Networks. *Emerging Trends in ICT Security*, 359-373.

244. Cahill, M. H., Lambert, D., Pinheiro, J. C., and Sun, D. X. (2002). Detecting fraud in the real world. *Handbook of massive data sets*. Boston: Springer, 911-929.

245. Ferreira, P., Alves, R., Belo, O., and Cortesão, L. (2006). Establishing fraud detection patterns based on signatures. *Industrial Conference on Data Mining*, Leipzig, Germany, 526-538.

246. Tang, J., Cheng, Y., and Zhou, C. (2009). Sketch-based SIP flooding detection using Hellinger distance. *IEEE Global Telecommunications Conference*, Honolulu, United States, 1-6.

247. Spathoulas, G. P., and Katsikas, S. K. (2010). Reducing false positives in intrusion detection systems. *Computers & Security*, 29(1), 35-44.

248. Juszczak, P., Adams, N. M., Hand, D. J., Whitrow, C., and Weston, D. J. (2008). Off-the-peg and bespoke classifiers for fraud detection. *Computational Statistics & Data Analysis*, 52(9), 4521-4532.

249. Gold, S. (2014). The evolution of payment card fraud. *Computer Fraud & Security*, 2014(3), 12-17.

250. İnternet: Nilson Report, The Nilson Report 1164. URL: https://nilsonreport.com/publication_newsletter_archive_issue.php?issue=1164 , Son Erişim Tarihi: 09.09.2020.

251. İnternet: Visa, The Evolution of Payment Security. URL: https://tn.visamiddleeast.com/dam/VCOM/regional/na/us/visa-everywhere/documents/visa-security-timeline.pdf , Son Erişim Tarihi: 09.09.2020.

252. İnternet: Gabelgård, S., European Fraud Report – Payments Industry Challenges. Nets, URL: https://www.nets.eu/solutions/fraud-and-dispute-services/Documents/Nets-Fraud-Report-2019.pdf , Son Erişim Tarihi: 09.09.2020.

253. Sivakumar, N., and Balasubramanian, R. (2015). Fraud detection in credit card transactions: classification, risks and prevention techniques. *International Journal of Computer Science and Information Technologies*, 6(2), 1379-1386.

254. Barker, K. J., D'amato, J., and Sheridon, P. (2008). Credit card fraud: awareness and prevention. *Journal of Financial Crime*, 15(4), 398-410.

255. Delamaire, L., Abdou, H., and Pointon, J. (2009). Credit card fraud and detection techniques: a review. *Banks and Bank systems*, 4(2), 57-68.

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

257. Sethi, N., and Gera, A. (2014). A revived survey of various credit card fraud detection techniques. *International Journal of Computer Science and Mobile Computing*, 3(4), 780-791.

258. Awoyemi, J. O., Adetunmbi, A. O., and Oluwadare, S. A. (2017). Credit card fraud detection using machine learning techniques: A comparative analysis. *IEEE International Conference on Computing Networking and Informatics*, Lagos, Nigeria, 1-9.

259. Van Vlasselaer, V., Bravo, C., Caelen, O., Eliassi-Rad, T., Akoglu, L., Snoeck, M., and Baesens, B. (2015). APATE: A novel approach for automated credit card transaction fraud detection using network-based extensions. *Decision Support Systems*, 75, 38-48.

260. Charleonnan, A. (2016). Credit card fraud detection using RUS and MRN algorithms. *IEEE Management and Innovation Technology International Conference*, Bang-San, Thailand, 73-76.

261. Mahmud, M. S., Meesad, P., and Sodsee, S. (2016). An evaluation of computational intelligence in credit card fraud detection. *IEEE International Computer Science and Engineering Conference*, Chiang Mai, Thailand,1-6.

262. Carneiro, N., Figueira, G., and Costa, M. (2017). A data mining based system for credit-card fraud detection in e-tail. *Decision Support Systems*, 95, 91-101.

263. Mishra, C., Gupta, D. L., and Singh, R. (2017). Credit Card Fraud Identification Using Artificial Neural Networks. *International Journal of Computer Systems*, 4(7), 151-159.

264. Kültür, Y., and Çağlayan, M. U. (2017). Hybrid approaches for detecting credit card fraud. *Expert Systems*, 34(2), 1-13.

265. Kazemi, Z., and Zarrabi, H. (2017). Using deep networks for fraud detection in the credit card transactions. *IEEE International Conference on Knowledge-Based Engineering and Innovation*, Tehran, Iran, 630-633.

266. Braun, F., Caelen, O., Smirnov, E. N., Kelk, S., and Lebichot, B. (2017). Improving card fraud detection through suspicious pattern discovery. *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, Arras, France, 181-190.

267. Kho, J. R. D., and Vea, L. A. (2017). Credit card fraud detection based on transaction behavior. *IEEE Region 10 Conference*, Penang, Malaysia, 1880-1884.

268. Artikis, A., Katzouris, N., Correia, I., Baber, C., Morar, N., Skarbovsky, I., Fournier, F., and Paliouras, G. (2017). Industry Paper: A Prototype for Credit Card Fraud Management. *ACM International Conference on Distributed and Event-based Systems*, Barcelona, Spain, 1-12.

269. Kavitha, M., and Suriakala, M. (2017). Real time credit card fraud detection on huge imbalanced data using meta-classifiers. *IEEE International Conference on Inventive Computing and Informatics*, Coimbatore, India, 881-887.

270. Porwal, U., and Mukund, S. (2018). Credit card fraud detection in e-commerce: An outlier detection approach. *arXiv:1811.02196*.

271. Xuan, S., Liu, G., Li, Z., Zheng, L., Wang, S., and Jiang, C. (2018). Random forest for credit card fraud detection. *IEEE International Conference on Networking, Sensing and Control*, Zhuhai, China, 1-6.

272. Fiore, U., De Santis, A., Perla, F., Zanetti, P., and Palmieri, F. (2019). Using generative adversarial networks for improving classification effectiveness in credit card fraud detection. *Information Sciences*, 479, 448-455.

273. Wang, C., and Han, D. (2019). Credit card fraud forecasting model based on clustering analysis and integrated support vector machine. *Cluster Computing*, 22(6), 13861-13866.

274. Gaumer, Q., Mortier, S., and Moutaib, A. (2016). Financial institutions and cyber crime Between vulnerability and security. *Financial Stability Review*, 20, 45-52.

275. Rosenberg, H. (2019). Banking & Financial Services Cyber Threat Landscape Report. *IntSights*, 1-14.

276. Gupta, A. (2018). The evolution of fraud: Ethical implications in the age of large-scale data breaches and widespread artificial intelligence solutions deployment. *International Telecommunication Union Journal*, 1, 1-7.

277. West, J., and Bhattacharya, M. (2016). Intelligent financial fraud detection: a comprehensive review. *Computers & Security*, 57, 47-66.

278. Estabrooks, A., Jo, T., and Japkowicz, N. (2004). A multiple resampling method for learning from imbalanced data sets. *Computational Intelligence*, 20(1), 18-36.

279. Huang, J. W., Chiang, C. W., and Chang, J. W. (2018). Email security level classification of imbalanced data using artificial neural network: The real case in a world-leading enterprise. *Engineering Applications of Artificial Intelligence*, 75, 11-21.

280. Jo, T., and Japkowicz, N. (2004). Class imbalances versus small disjuncts. *ACM SIGKDD Explorations Newsletter*, 6(1), 40-49.

281. Agrawal, A., Viktor, H. L., and Paquet, E. (2015). SCUT: Multi-class imbalanced data classification using SMOTE and cluster-based undersampling. *IEEE International Joint Conference on Knowledge Discovery, Knowledge Engineering and Knowledge Managemen*t, Lisbon, Portugal, 226-234.

282. Lin, W. C., Tsai, C. F., Hu, Y. H., and Jhang, J. S. (2017). Clustering-based undersampling in class-imbalanced data. *Information Sciences*, 409, 17-26.

283. Nekooeimehr, I., and Lai-Yuen, S. K. (2016). Adaptive semi-unsupervised weighted oversampling (A-SUWO) for imbalanced datasets. *Expert Systems with Applications*, 46, 405-416.

284. Yen, S. J., and Lee, Y. S. (2006). Under-sampling approaches for improving prediction of the minority class in an imbalanced dataset. *International Conference on Intelligent Computing*, Kunming, China, 731-740.

285. Guo, H., Zhou, J., and Wu, C. A. (2018). Imbalanced learning based on data-partition and SMOTE. *Information*, 9, 1-22.

286. İnternet: Université Libre de Bruxelles Machine Learning Group, URL: https://www.kaggle.com/mlg-ulb/creditcardfraud , Son Erişim Tarihi: 09.09.2020.

287. İnternet: Revolution Analytics, URL: https://packages.revolutionanalytics.com/datasets/ccFraud.csv , Son Erişim Tarihi: 09.09.2020.

288. Fernández-Delgado, M., Cernadas, E., Barro, S., and Amorim, D. (2014). Do we need hundreds of classifiers to solve real world classification problems?. *Journal of Machine Learning Research*, 15(1), 3133-3181.

289. Zecevic, P., and Bonaci, M. (2016). *Spark in Action*. Shelter Island: Manning Publications, 18-20.

290. Patil, D. V., and Bichkar, R. S. (2012). Issues in optimization of decision tree learning: A survey. *International Journal of Applied Information Systems*, 3(5), 13-30.

291. Shaik, A. B., and Srinivasan, S. (2019). A brief survey on random forest ensembles in classification model. International Conference on Innovative Computing and Communications, Ostrava, Czech Republic, 253-260.

292. Krauss, C., Do, X. A., and Huck, N. (2017). Deep neural networks, gradient-boosted trees, random forests: Statistical arbitrage on the S&P 500. *European Journal of Operational Research*, 259(2), 689-702.

293. Murtagh, F. (1991). Multilayer perceptrons for classification and regression. *Neurocomputing*, 2(5-6), 183-197.

294. Whitaker, T., Beranger, B., and Sisson, S. A. (2019). Logistic regression models for aggregated data. *arXiv:1912.03805*.

295. Saputra, D. M., Saputra, D., and Oswari, L. D. (2020). Effect of Distance Metrics in Determining K-Value in K-Means Clustering Using Elbow and Silhouette Method. *Advances in Intelligent Systems Research*, 172, 341-346.

296. Priscilla, C. V., and Prabha, D. P. (2019). Credit Card Fraud Detection: A Systematic Review. *International Conference on Information, Communication and Computing Technology*, Karachi, Pakistan, 290-303.

297. Zhang, C., Liu, C., Zhang, X., and Almpanidis, G. (2017). An up-to-date comparison of state-of-the-art classification algorithms. *Expert Systems with Applications*, 82, 128-150.

298. Adam, S. P., Alexandropoulos, S. A. N., Pardalos, P. M., and Vrahatis, M. N. (2019). No free lunch theorem: a review. *Approximation and Optimization*, 145, 57-82.

299. Seyedhossein, L., and Hashemi, M. R. (2010). A Timelier Credit Card Fraud Detection by Mining Transaction Time Series. *International Journal of Information and Communication Technology*, 2(3), 21-28.

300. Esling, P., and Agon, C. (2012). Time-series data mining. *ACM Computing Surveys*, 45(1), 1-34.

301. Durlauf, S., and Blume, L. (2016). *Macroeconometrics and time series analysis*. London: Palgrave Macmillan, 60-67.

302. Kirchgässner, G., and Wolters, J. (2007). *Introduction to modern time series analysis*. Berlin: Springer, 12.

303. Chatfield, C. (2000). Time-series forecasting. Boca Raton: CRC Press, 23-25.

304. Maimon, O., and Rokach, L. (2005). *Data mining and knowledge discovery handbook*, Boston: Springer, 1069-1103.

305. Granger, C. W. (2004). Time series analysis, cointegration, and applications. *American Economic Review*, 94(3), 421-425.

306. Fu, T. C. (2011). A review on time series data mining. *Engineering Applications of Artificial Intelligence*, 24(1), 164-181.

307. Faloutsos, C., Ranganathan, M., and Manolopoulos, Y. (1994). Fast subsequence matching in time-series databases. *ACM SIGMOD Record*, 23(2), 419-429.

308. Korn, F., Jagadish, H. V., and Faloutsos, C. (1997). Efficiently supporting ad hoc queries in large datasets of time sequences. *ACM SIGMOD Record*, 26(2), 289-300.

309. Popivanov, I., and Miller, R. J. (2002). Similarity search over time-series data using wavelets. *IEEE International Conference on Image Analysis and Signal*, Zhejiang, China, 212-221.

310. Aghabozorgi, S., Shirkhorshidi, A. S., and Wah, T. Y. (2015). Time-series clustering–a decade review. *Information Systems*, 53, 16-38.

311. Zolhavarieh, S., Aghabozorgi, S., and Teh, Y. W. (2014). A review of subsequence time series clustering. *The Scientific World Journal*, 2014, 1-20.

312. Wilson, S. J. (2017). Data representation for time series data mining: time domain approaches. *Wiley Interdisciplinary Reviews: Computational Statistics*, 9, 1-6.

313. Schäfer, P. (2016). Scalable time series classification. *Data Mining and Knowledge Discovery*, 30(5), 1273-1298.

314. Zhang, R., Zheng, F., and Min, W. (2018). Sequential Behavioral Data Processing Using Deep Learning and the Markov Transition Field in Online Fraud Detection. *ACM SIGKDD Conference On Knowledge Discovery And Data Mining Data Science in Fintech Workshop*, 1-5.

315. Karimi-Bidhendi, S., Munshi, F., and Munshi, A. (2018). Scalable classification of univariate and multivariate time series. *IEEE International Conference on Big Data*, Seattle, United States, 1598-1605.

316. Sánchez, F. R., and Cervera, J. G. (2019). ECG Classification Using Artificial Neural Networks. *Journal of Physics: Conference Series*, 1221(2019), 1-6.

317. Hsueh, Y., Ittangihala, V. R., Wu, W. B., Chang, H. C., and Kuo, C. C. (2019). Condition monitor system for rotation machine by CNN with recurrence plot. *Energies*, 12(3221), 1-13.

318. Said, A. B., and Erradi, A. (2019). Deep-Gap: A deep learning framework for forecasting crowdsourcing supply-demand gap based on imaging time series and residual learning. *IEEE International Conference on Cloud Computing Technology and Science*, Sydney, Australia, 279-2896.

319. Yang, C. L., Yang, C. Y., Chen, Z. X., and Lo, N. W. (2019). Multivariate time series data transformation for convolutional neural network. *IEEE/SICE International Symposium on System Integration*, Paris, France, 188-192.

320. Qin, Z., Zhang, Y., Meng, S., Qin, Z., and Choo, K. K. R. (2020). Imaging and fusing time series for wearable sensor-based human activity recognition. *Information Fusion*, 53, 80-87.

321. Hong, Y. Y., Martinez, J. J. F., and Fajardo, A. C. (2020). Day-Ahead Solar Irradiation Forecasting Utilizing Gramian Angular Field and Convolutional Long Short-Term Memory. *IEEE Access*, 8, 18741-18753.

322. Liu, X., Cai, H., Zhong, R., Sun, W., and Chen, J. (2020). Learning Traffic as Images for Incident Detection Using Convolutional Neural Networks. *IEEE Access*, 8, 7916-7924.

323. Yang, C. L., Chen, Z. X., and Yang, C. Y. (2020). Sensor Classification Using Convolutional Neural Network by Encoding Multivariate Time Series as Two-Dimensional Colored Images. *Sensors*, 20(168), 1-15.

324. Estebsari, A., and Rajabi, R. (2020). Single Residential Load Forecasting Using Deep Learning and Image Encoding Techniques. *Electronics*, 9(68), 1-17.

325. Wang, Z., and Oates, T. (2015). Encoding time series as images for visual inspection and classification using tiled convolutional neural networks. *Association for the Advancement of Artificial Intelligence Conference on Artificial Intelligence*, Austin, United States, 1-7.

326. Wang, Z., and Oates, T. (2015). Imaging time-series to improve classification and imputation. *ACM International Conference on Artificial Intelligence*, Buenos Aires, Argentina, 3939-3945.

327. Vasimalla, K. (2014). A survey on time series data mining. Intern. J. of Innovative Research in Computer and Communication Engineering, 2, 170-179.

328. Gamboa, J. C. B. (2017). Deep learning for time-series analysis. *arXiv:1701.01887*.

329. Fawaz, H. I., Forestier, G., Weber, J., Idoumghar, L., and Muller, P. A. (2019). Deep learning for time series classification: a review. *Data Mining and Knowledge Discovery*, 33(4), 917-963.

330. Alom, M. Z., Taha, T. M., Yakopcic, C., Westberg, S., Sidike, P., Nasrin, M. S., Hasan M., Essen, B. C. V., Awwal, A. A. S., and Asari, V. K. (2019). A state-of-the-art survey on deep learning theory and architectures. *Electronics*, 8(3), 292, 1-67.

331. Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., Liu, T., Wang, X., Cai, J., and Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern Recognition*, 77, 354-377.

332. Liu, W., Wang, Z., Liu, X., Zeng, N., Liu, Y., and Alsaadi, F. E. (2017). A survey of deep neural network architectures and their applications. *Neurocomputing*, 234, 11-26.

333. Janocha, K., and Czarnecki, W. M. (2017). On loss functions for deep neural networks in classification. *Theoretical Foundations of Machine Learning*, Cracow, Poland, 1-10.

334. Abakarim, Y., Lahby, M., and Attioui, A. (2018). An Efficient Real Time Model For Credit Card Fraud Detection Based On Deep Learning. *ACM International Conference on Intelligent Systems: Theories and Applications*, Rabat, Morocco. 1-7.

335. Al-Shabi, M. A. (2019). Credit Card Fraud Detection Using Autoencoder Model in Unbalanced Datasets. *Journal of Advances in Mathematics and Computer Science*, 33(5), 1-16.

336. Mohammed, R. A., Wong, K. W., Shiratuddin, M. F., and Wang, X. (2018). Scalable machine learning techniques for highly imbalanced credit card fraud detection: a comparative study. *Pacific Rim International Conference on Artificial Intelligence*, Nanjing, China ,237-246.

337. Makki, S., Haque, R., Taher, Y., Assaghir, Z., Hacid, M. S., and Zeineddine, H. (2018). A Cost-Sensitive Cosine Similarity K-Nearest Neighbor for Credit Card Fraud Detection. *International Conference on Big Data and Cybersecurity Intelligence*, Beirut, Lebanon, 42-47.

338. Xie, N., Ras, G., van Gerven, M., and Doran, D. (2020). Explainable deep learning: A field guide for the uninitiated. *arXiv:2004.14545*.

339. Spinner, T., Schlegel, U., Schäfer, H., and El-Assady, M. (2019). explAIner: A visual analytics framework for interactive and explainable machine learning. *IEEE Transactions on Visualization and Computer Graphics*, 26(1), 1064-1074.

340. Friedman, J., Hastie, T., Tibshirani, R. and Friedman, J., (2001). *The Elements of Statistical Learning Data Mining, Inference, and Prediction* (Second Edition), New York: Springer, 353-358.

341. Thomson, R., and Schoenherr, J. R. (2020). Knowledge-to-Information Translation Training (KITT): An Adaptive Approach to Explainable Artificial Intelligence. *International Conference on Human-Computer Interaction*, Copenhagen, Denmark 187-204.

342. Miller, T. (2019). "But why?" Understanding explainable artificial intelligence. *The ACM Magazine for Students*, 25(3), 20-25.

343. İnternet: Erhan, D., A. Courville, Y. Bengio, Understanding representations learned in deep architectures, Department d'Informatique et Recherche Operationnelle URL: https://aaroncourville.files.wordpress.com/2011/11/invariances_techreport-3.pdf , Son Erişim Tarihi: 09.09.2020.

344. Simonyan, K., Vedaldi, A., and Zisserman, A. (2013). Deep inside convolutional networks: Visualising image classification models and saliency maps. *arXiv:1312.6034*.

345. Zeiler, M. D., and Fergus, R. (2014). Visualizing and understanding convolutional networks. *European Conference on Computer Vision*, Zurich, Switzerland, 818-833.

346. Kim, B., Rudin, C., and Shah, J. A. (2014). The bayesian case model: A generative approach for case-based reasoning and prototype classification. *Advances in neural information processing systems*, Montréal, Canada, 1-9.

347. Bach, S., Binder, A., Montavon, G., Klauschen, F., Müller, K. R., and Samek, W. (2015). On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PloS One*, 10(7), 1-46.

348. Letham, B., Rudin, C., McCormick, T. H., and Madigan, D. (2015). Interpretable classifiers using rules and bayesian analysis: Building a better stroke prediction model. *The Annals of Applied Statistics*, 9(3), 1350-1371.

349. Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., and Torralba, A. (2016). Learning deep features for discriminative localization. *IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, United States, 2921-2929.

350. Ribeiro, M. T., Singh, S., and Guestrin, C. (2016). "Why should I trust you?" Explaining the predictions of any classifier. *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, United States, 1135-1144.

351. Lundberg, S. M., and Lee, S. I. (2017). A unified approach to interpreting model predictions. Advances in Neural Information Processing Systems, Long Beach, United States, 4765-4774.

352. Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. *IEEE International Conference on Computer Vision*, Venice, Italy, 618-626.

353. Zintgraf, L. M., Cohen, T. S., Adel, T., and Welling, M. (2017). Visualizing deep neural network decisions: Prediction difference analysis. *International Conference on Learning Representations*, Toulon, France, 1-12.

354. Montavon, G., Lapuschkin, S., Binder, A., Samek, W., and Müller, K. R. (2017). Explaining nonlinear classification decisions with deep taylor decomposition. *Pattern Recognition*, 65, 211-222.

355. Petsiuk, V., Das, A., and Saenko, K. (2018). Rise: Randomized input sampling for explanation of black-box models. *arXiv:1806.07421*.

356. Chattopadhay, A., Sarkar, A., Howlader, P., and Balasubramanian, V. N. (2018). Grad-cam++: Generalized gradient-based visual explanations for deep convolutional

networks. *IEEE Winter Conference on Applications of Computer Vision*, Lake Tahoe, United States, 839-847.

357. Li, H., Tian, Y., Mueller, K., and Chen, X. (2019). Beyond saliency: understanding convolutional neural networks from saliency prediction on layer-wise relevance propagation. *Image and Vision Computing*, 83, 70-86.

358. Fish, B., Kun, J., Lelkes, Á. D., Reyzin, L., and Turán, G. (2015). On the computational complexity of mapreduce. *International Symposium on Distributed Computing*, Tokyo, Japan, 1-5.

359. He, K., and Sun, J. (2015). Convolutional neural networks at constrained time cost. *IEEE International Conference on Computer Vision and Pattern Recognition*, Boston, United States, 5353-5360.

360. Hospedales, T., Antoniou, A., Micaelli, P., and Storkey, A. (2020). Meta-learning in neural networks: A survey. *arXiv:2004.05439*.

ÖZGEÇMİŞ

Kişisel Bilgiler

Soyadı, adı : SİNANÇ TERZİ, Duygu
Uyruğu : T.C.
Doğum tarihi ve yeri : 20.09.1988, Elazığ
Medeni hali : Evli
e-mail : duygusinanc@gazi.edu.tr

Eğitim

Derece	Eğitim Birimi	Mezuniyet Tarihi
Doktora	Gazi Üniversitesi / Bilgisayar Mühendisliği	Devam ediyor
Yüksek lisans	Gazi Üniversitesi / Bilgisayar Mühendisliği	2014
Lisans	Selçuk Üniversitesi / Bilgisayar Mühendisliği	2011
Lise	Elazığ Anadolu Lisesi	2006

İş Deneyimi

Yıl	Yer	Görev
2013-Halen	Gazi Üniversitesi	Araştırma Görevlisi
2012-2013	Amasya Üniversitesi	Araştırma Görevlisi
2012-2012	Pamukkale Üniversitesi	Araştırma Görevlisi

Yabancı Dil

İngilizce

Yayınlar

1. Terzi, D. S., Sagiroglu, S. (2019). A New Big Data Model Using Distributed Cluster-Based Resampling for Class-Imbalance Problem. *Applied Computer Systems*, 24(2), 104-110.
2. Terzi, D. S., Sagiroglu, S. (2018). Siber Güvenlik için Büyük Veri Yaklaşımları. *Siber Güvenlik ve Savunma Farkındalık ve Caydırıcılık*. Ankara: Grafiker Yayınları, 375-386.

3. Terzi, D. S., Arslan, B., Sagiroglu, S. (2018). Smart grid security evaluation with a big data use case. *IEEE International Conference on Compatibility, Power Electronics and Power Engineering*, 1-6, Doha, Qatar.
4. Terzi, D. S., Sagiroglu, S. (2017). Mobil İletişim Sektöründe Büyük Veri ve Uygulama Örnekleri. *Büyük Veri ve Açık Veri Analitiği: Yöntemler ve Uygulamalar*. Ankara: Grafiker Yayınları, 301-312.
5. Sağıroğlu, Ş., Terzi, D. S., Terzi, R., Canbay Y., Gündüz M. S., Arslan B., Ayaydın A., Gökalp A. B. (2016). Büyük Veri Analitiği, Güvenliği ve Mahremiyeti, *Gazi Üniversitesi Büyük Veri ve Bilgi Güvenliği Merkezi*.
6. Terzi, D. S., Terzi, R., Sagiroglu, S. (2017). Big data analytics for network anomaly detection from netflow data. *IEEE International Conference on Computer Science and Engineering*, 592-597, Antalya, Türkiye.
7. Terzi, D. S., Demirezen, U., Sagiroglu, S. (2016). Evaluations of big data processing. *Services Transactions on Big Data*, 3(1), 44-53.
8. Terzi, D. S., Sagiroglu, S. (2016). Mobil cihaz kullanıcı davranışlarının modellenmesi için yeni bir yaklaşım. *IEEE International Symposium on Digital Forensic and Security*, 171, Little Rock, USA.
9. Ayaydin, A., Terzi, D. S., Sagiroglu, S. (2016). Security, privacy and forensics issues on big data and cloud computing, *International Conference on Computer Science and Engineering*, 1-6, Çorlu, Türkiye.
10. Terzi, D. S., Terzi, R., Sagiroglu, S. (2015). A survey on security and privacy issues in big data. *IEEE International Conference for Internet Technology and Secured Transactions*, 202-207, London, UK.
11. Sağıroğlu, Ş., Dener, M., Güneş, S., Güllü, A., Tataroğlu, A., Orman, A., Sinanç, D., Terzi, R., Şahin, H., Çetinyokuş, S., Savran, S., Akçay, H. (2015). Ulusal Veritabanı ve Atıf İndeksi Kurulumu için Stratejiler, Problemler ve Çözüm Önerileri. *Gazi Üniversitesi Fen Bilimleri Dergisi Part C: Tasarım ve Teknoloji*, 3(2), 501-512.
12. Sağıroğlu, Ş., Dener, M., Güllü, A., Tataroğlu, A., Orman, A., Güneş, S., Şahin, H., Sinanç, D., Çetinyokuş, Terzi, R., S., Savran, S., Akçay, H. (2015). Ulusal Veritabanı ve Atıf İndeksi Strateji Raporu. *Gazi Üniversitesi Fen Bilimleri Enstitüsü*.
13. Terzi, R., Yavanoglu, U., Sinanc, D., Oguz, D., Cakir, S. (2014). An Intelligent Technique for Detecting Malicious Users on Mobile Stores. *IEEE International Conference on Machine Learning and Applications*, 470-477, Detroit, USA.

14. Sinanc, D., Sahin, M., Esen, Z., Yavanoglu, U., Sagiroglu, S. (2014). An intelligent feedback control mechanism for brushless DC motors. *IEEE International Power Electronics and Motion Control Conference and Exposition*, 939-944, Antalya, Türkiye.
15. Sinanc, D., Yavanoglu, U. (2013). A new approach to detecting content anomalies in Wikipedia. *IEEE International Conference on Machine Learning and Applications*, 288-293, Miami, USA.
16. Sinanc, D., Sagiroglu, S. (2013). A review on cloud security, *ACM International Conference on Security of Information and Networks*, 321-325, Aksaray, Türkiye.
17. Sagiroglu, S., Sinanc, D. (2013). Big data: A review, *IEEE International Conference on Collaboration Technologies and Systems*, 42-47, San Diego, USA.

Hobiler

Yürüyüş, Fotoğrafçılık, Kitap Okumak

GAZİ GELECEKTİR..