

Kamu Sağlık Kurumlarında Tıbbi İşlem Verileri Üzerinden Yapay Zekâ Destekli ve Risk Tabanlı Bir Denetim Modeli

An AI-Driven and Risk-Based Audit Model Based on Medical Procedure Data in Public Health Institutions

Harun BULUT¹ 

Muhammed Hanefi CALP² 

DOI:10.33461/uybisbbd.1929597

Öz

Makale Bilgileri

Makale Türü:

Araştırma Makalesi

Geliş Tarihi:

13.04.2026

Kabul Tarihi:

17.06.2026

©2026 UYBISBBDD
Tüm hakları saklıdır.



Bu çalışmada, kamu sağlık kurumlarında gerçekleştirilen tıbbi işlemlerin denetimine yönelik yapay zekâ destekli ve risk tabanlı bütünlük bir analitik model geliştirilmiştir. Araştırmanın temel amacı, yüksek hacimli sağlık verileri üzerinde olağandışı işlem örüntülerini belirleyerek denetim süreçlerinde erken uyarı, risk önceliklendirme ve karar destek mekanizmalarının oluşturulmasına katkı sağlamaktır. Bu doğrultuda, bir Ağız ve Diş Sağlığı Merkezine ait, 2024–2025 dönemini kapsayan ve kimlik bilgileri anonimleştirilmiş toplam 42174 tıbbi işlem kaydı analiz edilmiştir. Modelleme süreci; veri ön işleme, gözetimsiz anomali tespiti, kümeleme analizi ve sınıflandırma aşamalarını içeren çok katmanlı bir yapıda kurgulanmıştır. Anomali tespiti aşamasında İzolasyon Ormanı (Isolation Forest) ve Yerel Uzaklık Faktörü (Local Outlier Factor) algoritmaları kullanılmış; sırasıyla 2109 ve 1745 işlem potansiyel anomali olarak belirlenmiştir. Kümeleme analizlerinde k-Ortalamalar (k-Means) yöntemi veri setini üç anlamlı işlem segmentine ayırırken, Yoğunluk Tabanlı Uzamsal Kümeleme (DBSCAN) yöntemi veri yapısı içerisindeki lokal yoğunluk farklılıklarını ve aykırı kümelene örüntülerini ortaya koymuştur. Sınıflandırma aşamasında, gözetimsiz öğrenme yöntemleriyle elde edilen risk etiketleri hedef değişken olarak kullanılarak Rastgele Orman (Random Forest) ve Gradyan Artırma (Gradient Boosting) modelleri geliştirilmiştir. Performans değerlendirme sonuçlarına göre Rastgele Orman modeli, $AUC=0.999$ ve doğruluk oranı= 0.992 değerleriyle en yüksek sınıflandırma başarısını göstermiştir. Elde edilen yüksek AUC değeri, modelin normal işlem örüntüleri ile potansiyel riskli işlemleri ayırt etme kapasitesinin oldukça güçlü olduğunu ortaya koymaktadır. Gerçekleştirilen değişken önem analizleri, olağandışı işlem davranışlarının özellikle güvence türü, kronik hastalık bilgisi, işlem kodları ve mali göstergeler ile güçlü ilişkiler taşıdığını göstermiştir. Sonuç olarak geliştirilen yapay zekâ destekli risk tabanlı denetim modeli; kamu sağlık kurumlarında denetim kaynaklarının etkin kullanımı, riskli işlemlerin önceliklendirilmesi ve veri temelli erken uyarı mekanizmalarının oluşturulması açısından uygulanabilir bir karar destek çerçevesi sunmaktadır. Çalışmanın, sağlık denetimi alanında proaktif, analitik ve veri odaklı denetim yaklaşımlarının geliştirilmesine önemli katkılar sağlayacağı değerlendirilmektedir.

Anahtar Kelimeler: kamu sağlık kurumu, tıbbi veri, risk tabanlı denetim, yapay zekâ, anomali tespiti.

Article Info

Paper Type:

Research Paper

Received:

13.04.2026

Accepted:

17.06.2026

©2026 UYBISBBDD
All rights reserved.



Abstract

This study develops an AI-powered, risk-based integrated analytical model for monitoring medical procedures performed in public healthcare institutions. The main objective of the research is to contribute to the creation of early warning, risk prioritization, and decision support mechanisms in monitoring processes by identifying unusual procedure patterns on high-volume healthcare data. To this end, a total of 42174 medical procedure records from a Dental and Oral Health Center, covering the period 2024–2025 and with anonymized identifying information, were analyzed. The modeling process was designed in a multi-layered structure including data preprocessing, unsupervised anomaly detection, clustering analysis, and classification stages. In the anomaly detection stage, the Isolation Forest and Local Outlier Factor algorithms were used; 2109 and 1745 procedures were identified as potential anomalies, respectively. In clustering analyses, the k-Means method divided the dataset into three meaningful transaction segments, while the Density-Based Spatial Clustering (DBSCAN) method revealed local density differences and outlier clustering patterns within the data structure. In the classification phase, Random Forest and Gradient Boosting models were developed using risk labels obtained through unsupervised learning methods as target variables. According to the performance evaluation results, the Random Forest model showed the highest classification success with $AUC=0.999$ and accuracy rate= 0.992 . The high AUC value obtained reveals that the model has a very strong capacity to distinguish between normal transaction patterns and potentially risky transactions. Variable significance analyses showed that unusual transaction behaviors have strong relationships, especially with assurance type, chronic disease information, transaction codes, and financial indicators. Consequently, the developed AI-supported risk-based audit model offers a viable decision support framework for the effective use of audit resources in public healthcare institutions, prioritization of risky processes, and the creation of data-driven early warning mechanisms. It is assessed that this study will make significant contributions to the development of proactive, analytical, and data-driven audit approaches in the field of healthcare auditing.

Keywords: public health institution, medical data, risk-based audit, artificial intelligence, anomaly detection.

Atıf/ to Cite (APA): Bulut H., Calp M. H. (2026). Kamu Sağlık Kurumlarında Tıbbi İşlem Verileri Üzerinden Yapay Zekâ Destekli ve Risk Tabanlı Bir Denetim Modeli. Uluslararası Yönetim Bilişim Sistemleri ve Bilgisayar Bilimleri Dergisi, 10, 68-89. DOI: 10.33461/uybisbbd.1929597

¹ Department of Computer Technologies, Vocational High School, Biruni University, bulut.3693@hotmail.com, İstanbul, Türkiye.

² Department of Management Information Systems, Ankara Hacı Bayram Veli University, hanefi.calp@hbv.edu.tr. Ankara, Türkiye.

1. GİRİŞ

Kamu sağlık kurumları, yüksek bütçe büyüklükleri, yoğun tıbbi işlem hacmi ve çok sayıda mali sürecin eş zamanlı yürütülmesi nedeniyle kamu mali yönetimi açısından denetimi en güç alanlardan biri olarak değerlendirilmektedir. Sağlık hizmetlerinin finansmanı; genel bütçe ve döner sermaye kaynaklarının yanı sıra Sosyal Güvenlik Kurumu (SGK) geri ödeme mekanizmaları aracılığıyla yürütülmekte, bu durum tıbbi işlem kayıtları ile mali verilerin bütüncül biçimde izlenmesini zorunlu kılmaktadır (Güley & Bozdemir, 2024; Kılınç & Boztosun, 2024). Sağlık Bakanlığı bütçesinin 2025 yılı itibarıyla 1 trilyon TL'yi aşması ise sağlık alanında mali risklerin büyüklüğünü ve etkin denetim ihtiyacını gösterir (Cumhurbaşkanlığı Strateji ve Bütçe Başkanlığı, 2025).

Türkiye'de kamu mali yönetimi ve denetim sistemi, 5018 sayılı Kamu Malî Yönetimi ve Kontrol Kanunu ile düzenlenmiş olup, kamu kaynaklarının etkili, ekonomik ve verimli kullanımı temel amaç olarak benimsenmiştir (T.C. Resmî Gazete, 2003). Ancak uygulamada denetim faaliyetlerinin önemli bir bölümünün şikâyet ve ihbar temelli yürütülmesi, programlı denetimlerde ise çok sayıda birimin kapsanmaya çalışılması nedeniyle yüksek işlem hacmine sahip sağlık kurumlarında inceleme derinliğinin sınırlı kaldığı görülmektedir. Bu durum, mali usulsüzlüklerin ve sistematik hataların erken aşamada tespit edilmesini zorlaştırmakta ve kamu zararının artması riskini beraberinde getirmektedir (Küçük, 2023).

Sağlık kurumlarında mali işlemler yalnızca klasik muhasebe kayıtlarından ibaret olmayıp, medikal muhasebe kapsamında değerlendirilen tıbbi hizmet üretimiyle doğrudan ilişkilidir. Faturalandırılmayan işlemler, hatalı veya uyumsuz işlem kodları, mükerrer işlem kayıtları ve olağandışı işlem yoğunlukları, sağlık kurumlarında mali risklerin temel kaynakları arasında yer almaktadır (Kılınç & Boztosun, 2024). Sayıştay denetim raporları ve akademik çalışmalar, özellikle döner sermaye işletmelerinde muhasebe kayıtlarının tutarlılığı ve gelir-gider dengesi açısından tekrar eden yapısal sorunlara işaret etmektedir (Güley & Bozdemir, 2024). Bu bulgular, denetimin yalnızca mevzuata uygunluk kontrolüyle sınırlı kalmaması, tıbbi işlem verileri ile mali kayıtların birlikte değerlendirilmesi gerektiğini göstermektedir.

Mevcut denetim yapısında yaygın olarak kullanılan örnekleme temelli inceleme yöntemleri ve manuel kontroller, büyük hacimli ve çok boyutlu veri üreten sağlık kurumlarında tüm riskli işlemlerin zamanında tespit edilmesini güçleştirmektedir. Denetim kapasitesinin sınırlı olması ve denetim süreçlerinin çoğunlukla geriye dönük yürütülmesi, bazı usulsüzlüklerin uzun süre fark edilmeden devam etmesine neden olabilmektedir (Küçük, 2023). Bu durum, denetim kaynaklarının daha etkin kullanılabilmesi için risk temelli ve önleyici yaklaşımların benimsenmesini zorunlu kılmaktadır.

Bu bağlamda risk tabanlı denetim yaklaşımları, denetim faaliyetlerinin geçmiş işlemleri inceleyen reaktif bir yapıdan çıkarılarak, riskli alanların önceliklendirildiği proaktif bir çerçeveye dönüştürülmesini hedeflemektedir. Türkiye'de Sağlık Bakanlığı Teftiş Kurulu Başkanlığı bünyesinde oluşturulan Risk Esaslı Denetim Sistemi (REDES), denetim kaynaklarının riskli alanlara yönlendirilmesi açısından önemli bir kurumsal adım olarak değerlendirilmektedir (T.C. Sağlık Bakanlığı Teftiş Kurulu Başkanlığı, 2024). Ancak mevcut uygulamaların büyük ölçüde manuel analizlere ve sınırlı istatistiksel araçlara dayanması, yüksek hacimli tıbbi işlem verilerinden bütüncül ve öngörücü sonuçlar üretilmesini kısıtlamaktadır.

Son yıllarda yapay zekâ ve veri analitiği temelli yöntemler, denetim süreçlerinin dönüştürülmesinde önemli bir potansiyel sunmaktadır. Anomali tespiti, sınıflandırma ve kümeleme gibi makine öğrenmesi yaklaşımları, büyük veri setleri içerisinde olağan işlem davranışlarından sapma gösteren örüntülerin otomatik olarak belirlenmesine imkân tanımaktadır. Literatürde, yapay zekâ destekli denetim modellerinin hile ve usulsüzlüklerin tespitinde geleneksel yöntemlere kıyasla daha hızlı ve etkin sonuçlar üretebildiği vurgulanmaktadır (Efe & Tunçbilek, 2023; Albayrak, 2022).

Bu çalışmanın amacı, kamu sağlık kurumlarında tıbbi işlem verileri üzerinden yapay zekâ destekli risk tabanlı bir denetim modeli geliştirmektir. Önerilen model; tıbbi işlem kayıtları ile mali

göstergeleri birlikte analiz ederek olağandışı işlem örüntülerini belirlemeyi ve riskli işlemleri önceliklendirmeyi hedeflemektedir.

Çalışmada özellikle şu noktaların vurgulanması gerekir:

- Mevcut çalışmaların büyük ölçüde özel sağlık sigortası sahtekarlığına odaklandığı,
- Kamu sağlık kurumlarında işlem bazlı risk denetimi üzerine bütünleşik yapay zekâ modeli çalışmalarının sınırlı olduğu,
- Çalışmanın anomali tespiti, kümeleme ve sınıflandırma süreçlerini entegre eden bütüncül bir denetim modeli önerdiği,
- Klinik ve mali göstergelerin birlikte ele alınmasının literatürdeki önemli katkılardan biri olduğu.

Bu bağlamda çalışmanın özgün katkısını, kamu sağlık kurumlarında tıbbi işlem verileri ile mali göstergeleri birlikte analiz eden, yapay zekâ destekli ve risk tabanlı bütünleşik bir denetim modeli olarak özetlemek mümkündür.

2. KAVRAMSAL ÇERÇEVE

Literatür tarandığında, mevcut çalışmaların büyük bölümü bireysel sağlık sigortası sahtekarlığına odaklanmakta olup, kamu sağlık kurumlarında işlem verisi üzerinden kurumsal risk tabanlı denetim modeli geliştiren bütünleşik bir çerçeve sunmamaktadır. Kamu sağlık kurumlarında mali işlemler, yalnızca bütçe ve muhasebe kayıtlarından ibaret olmayıp tıbbi hizmet üretimi, döner sermaye uygulamaları ve faturalama süreçleriyle birlikte ilerleyen çok bileşenli bir yapı göstermektedir. Bu nedenle denetim, yalnızca finansal uygunluk kontrolü değil, aynı zamanda süreç bütünlüğünün izlenmesini ve işlem bazlı uyumsuzlukların görünür kılınmasını gerektiren daha geniş bir çerçevede ele alınmalıdır. Sağlık Bakanlığı döner sermaye işletmelerine yönelik denetim raporlarının muhasebesel incelemeleri, gelir-gider dengesi ve kayıt tutarlılığı açısından tekrar eden yapısal sorunların varlığına işaret ederek bu gerekliliği güçlendirmektedir (Güley & Bozdemir, 2024).

Medikal muhasebe perspektifi, tıbbi işlemlerin mali kayıtlara yansıma biçiminin denetim açısından kritik olduğunu ortaya koymaktadır. Özellikle faturalandırılmayan işlemlere yönelik tahakkukların eksik veya hatalı izlenmesi, kurum içi mali raporlamada kayıplara ve tutarsızlıklara yol açabilmekte; bu da denetimde tıbbi işlem verileri ile mali göstergelerin birlikte değerlendirilmesini zorunlu hale getirmektedir (Kılınç & Boztosun, 2024; Nazlıoğlu, 2020). Denetim ve kontrol kavramları, kamu kurumlarının hedeflerine uygun biçimde işleyişini sağlamak amacıyla faaliyetlerin izlenmesi ve sapmaların düzeltilmesi üzerine kuruludur. Sağlık hizmetlerinde kontrol işlevi, çok katmanlı harcama ve hizmet süreçleri nedeniyle yalnızca finansal uygunlukla sınırlı kalmamakta; süreçlerin bütünlüğünü ve tutarlılığını da kapsayan bir nitelik taşımaktadır. Bu çerçevede iç denetim, risk yönetimi ve kontrol sistemlerinin etkinliğine ilişkin yönetime güvence sağlayan bağımsız bir değerlendirme süreci olarak sağlık kurumlarında hem mali hem de operasyonel alanlara temas eden risklerin izlenmesini gerekli kılmaktadır (Arslan Kurtuluş, 2021; Karagöz & Yaman, 2019; T.C. Resmî Gazete, 2003).

Sağlık kurumlarında mali risk alanları, usulsüzlük türlerinin çeşitliliği nedeniyle daha da belirginleşmektedir. Sahte fatura düzenleme, mükerrer işlem kayıtları, kodlama hileleri (upcoding) ve uygulama dışı hizmet beyanı gibi örnekler; hem kamu bütçesi açısından doğrudan kayıp yaratmakta hem de işlemsel bütünlük ve kurumsal güven üzerinde olumsuz etkiler doğurmaktadır. Türkiye bağlamında sağlık harcamaları ile yolsuzluk arasındaki ilişkinin ampirik olarak gösterilmesi, mali risklerin yalnızca kurumsal değil aynı zamanda toplumsal sonuçları olduğunu da gösterir (Küçük, 2023; Albayrak, 2022; Demir & Kurt, 2017). Bu noktada risk tabanlı denetim yaklaşımı, denetim kaynaklarının sınırlı olduğu koşullarda riskli alanların önceliklendirilmesini ve denetimin daha proaktif biçimde yürütülmesini amaçlamaktadır. Türkiye’de Sağlık Bakanlığı bünyesinde risk

esaslı denetim yapılanmasını düzenleyen yönerge, risk odaklı denetim anlayışının kurumsal bir zemine taşındığını göstermektedir. Bununla birlikte risk odaklı iç denetime ilişkin çalışmalar, yaklaşımın etkin sonuç üretmesi için denetim planlamasında risk kriterlerinin sistematik biçimde uygulanması ve analitik araçların kullanımının güçlendirilmesi gerektiğine işaret etmektedir (Mandalas & Can, 2020; T.C. Sağlık Bakanlığı Teftiş Kurulu Başkanlığı, 2024; Türedi et al., 2015).

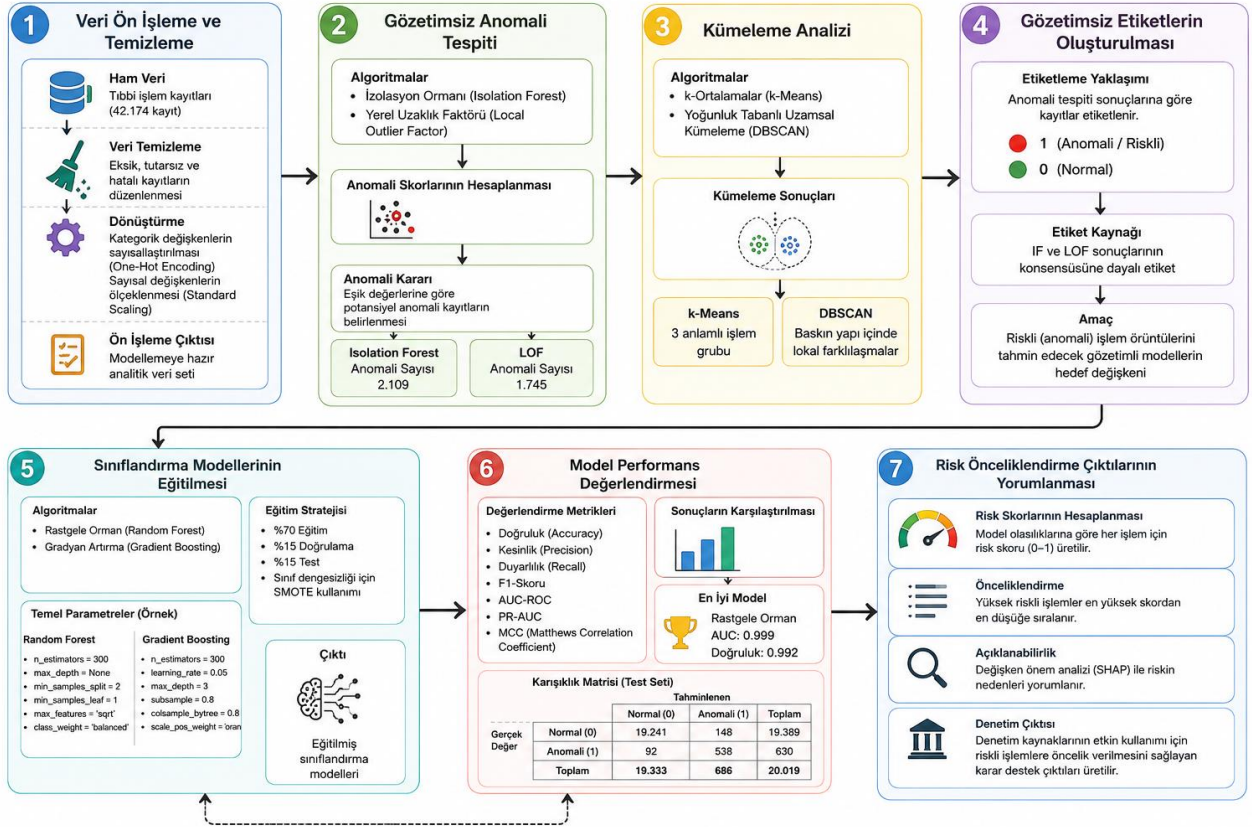
Teknoloji temelli denetim araçları ve veri analitiği yaklaşımları, geleneksel denetimin sınırlılıklarını aşma potansiyeli taşımaktadır. Özellikle sağlık sektöründe bilgi sistemleri üzerinden üretilen büyük hacimli veriler, anomali tespiti ve örüntü analizi gibi yöntemlerle riskli işlemlerin daha erken aşamada belirlenmesine imkân sağlayabilir. Pandemi döneminde uzaktan iç denetim uygulamalarına ilişkin bulgular da denetimin veri altyapısına ve dijital kapasiteye bağımlılığını görünür kılmakta; güçlü veri sistemleri ve iletişim altyapısının önemini vurgulamaktadır. Denetimde yapay zekâ kullanımına ilişkin değerlendirmeler ise bu araçların denetçinin yerini almak yerine, denetim kararlarını destekleyen ve önceliklendirmeyi güçlendiren bir karar destek rolüyle ele alınması gerektiğini belirtmektedir (Efe & Tunçbilek, 2023; Koçak & Ergün, 2023; Korkmaz & Coşkun, 2022; Koyuncuğil & Özgülbaş, 2010).

Sonuç olarak literatür, kamu sağlık kurumlarında denetimin süreç bütünlüğü, veri kalitesi ve risk önceliklendirmesi ekseninde güçlendirilmesi gerektiğini göstermektedir. Bu çalışma, söz konusu kavramsal çerçeveye dayanarak tıbbi işlem verileri ile mali göstergeleri birlikte ele alan, yapay zekâ destekli yöntemlere dayalı uygulamalı bir risk tabanlı denetim modeli geliştirmiş ve bu model aracılığıyla riskli işlem örüntülerinin belirlenerek denetim kaynaklarının daha etkin kullanılmasına yönelik bulguları tartışmıştır.

3. MATERYAL VE YÖNTEM

Bu çalışmada, kamu sağlık kurumlarında yürütülen tıbbi işlemlere ait büyük hacimli veriler üzerinden mali usulsüzlük risklerinin erken aşamada tespit edilebilmesine yönelik yapay zekâ destekli ve risk tabanlı bir denetim modeli geliştirilmiştir. Modelleme süreci aşağıdaki mantıksal akış doğrultusunda yapılandırılmıştır (Şekil 1):

1. Veri ön işleme ve temizleme
2. Gözetimsiz anomali tespiti
3. Kümeleme analizi
4. Gözetimsiz etiketlerin oluşturulması
5. Sınıflandırma modellerinin eğitilmesi
6. Model performans değerlendirmesi
7. Risk önceliklendirme çıktılarının yorumlanması



Şekil 1: Modelleme süreci

Araştırma, uygulamaya dayalı nicel bir yöntem çerçevesinde tasarlanmış; denetim süreçlerinde üretilen yüksek hacimli ve çok boyutlu verilerin yapay zekâ algoritmaları kullanılarak analiz edilmesi esas alınmıştır. Çalışma kapsamında, bir Ağız ve Diş Sağlığı Merkezine (ADSM) ait anonimleştirilmiş işlem verileri üzerinde uygulama gerçekleştirilmiş, geliştirilen yaklaşımın kamu sağlık kurumlarında mali usulsüzlüklerin erken tespitine katkı sunabilecek veri temelli ve uygulanabilir bir denetim aracı oluşturma potansiyeli ortaya konulmuştur.

Bu doğrultuda;

- Gözetimsiz öğrenme yöntemleri ile elde edilen etiketlerin gerçek suç/usulsüzlük etiketi olmadığı,
- Modelin denetçilere karar destek sağlayan önceliklendirme aracı niteliğinde olduğu,
- Çalışmanın erken uyarı ve risk odaklı inceleme yaklaşımını desteklediği,
- Model çıktılarının doğrudan yaptırım değil, analitik değerlendirme amacı taşıdığı vurgulanmaktadır.

Yani; çalışmada kullanılan etiketler gözetimsiz yöntemlerle üretilmiş olup, model gerçek usulsüzlük tespitinden ziyade potansiyel risk örüntülerinin belirlenmesine yöneliktir.

Çalışmanın yöntemsel yaklaşımı, geleneksel örnekleme temelli ve geriye dönük denetim yöntemlerinin sınırlılıklarını aşmak amacıyla, tüm işlem evrenini kapsayan veri temelli ve öngörücü bir denetim perspektifi çerçevesinde yapılandırılmıştır.

3.1. Çalışmanın Kapsamı ve Metodolojik Yaklaşım

Bu çalışmanın metodolojik yaklaşımı, kamu sağlık kurumlarında denetim süreçlerinde ortaya çıkan yüksek işlem hacmi ve veri karmaşıklığı dikkate alınarak oluşturulmuştur. Sağlık hizmetlerinin finansal yapısı, tıbbi işlem kayıtları ile doğrudan ilişkili çok sayıda işlem kaleminin eş zamanlı olarak yürütülmesini gerektirmektedir. Bu durum denetim faaliyetlerinde sistematik ve veri temelli analiz ihtiyacını artırmaktadır. Bu nedenle çalışma, denetim süreçlerinde riskli işlem davranışlarının analitik yöntemlerle belirlenmesine odaklanan bir metodolojik çerçeveye içerisinde ele alınmıştır.

Çalışmanın kapsamı, kamu sağlık kurumlarında gerçekleştirilen tıbbi işlemlere ait kayıtların işlem bazlı olarak analiz edilmesi ile sınırlandırılmıştır. Analizlerde işlem kodları, işlem türleri, işlem tekrar sıklıkları ve belirli işlem gruplarında gözlenen yoğunlaşmalar temel alınmıştır. Bu yaklaşım, olağan işlem davranışlarından sapma gösteren örüntülerin tespit edilmesine ve denetim açısından potansiyel risk alanlarının görünür kılınmasına yönelik olarak yapılandırılmıştır.

Bu çerçevede süreç, tıbbi işlem verilerinin analize uygun hâle getirilmesi, denetim açısından anlamlı değişkenlerin oluşturulması ve bu veriler üzerinden anomali ve risk değerlendirmesi yapılmasına imkân tanıyacak şekilde kurgulanmıştır. Bu çerçevede analiz süreci, riskli işlem örüntülerinin belirlenmesi ve denetim kaynaklarının önceliklendirilebilmesine katkı sağlayacak bir karar destek mantığı doğrultusunda yürütülmüştür.

3.2. Veri Seti ve Veri Kaynağı

Bu çalışmada kullanılan veri seti, bir Ağız ve Diş Sağlığı Merkezinde (ADSM) 2024-2025 yılları arasında sunulan ağız ve diş sağlığı hizmetlerine ilişkin tıbbi işlem kayıtlarından elde edilmiştir. Veriler, kurum bünyesinde kullanılan Hastane Bilgi Yönetim Sistemi (HBYS) üzerinden temin edilmiştir. Çalışmada, ilgili dönemde gerçekleştirilen tüm tıbbi işlemler analiz kapsamına alınmış; veri seti herhangi bir örnekleme veya filtreleme yapılmaksızın tam evren yaklaşımıyla kullanılmıştır. Bu sayede, ADSM’de yürütülen hizmet sunumuna ilişkin işlem hacmi ve dağılımı denetim perspektifiyle bütüncül olarak ele alınmıştır.

Veri setinde yer alan hasta ve hekim bilgileri, kişisel verilerin korunması ilkeleri doğrultusunda anonimleştirilmiş olup kişisel veri içermemektedir. Analiz sürecinde kimlik tanımlayıcı nitelikteki alanlar kullanılmamıştır. Ayrıca, ilgili sağlık kurumundan da etik kurul onayı alınmıştır.

Çalışmada yalnızca tıbbi işlemlere ilişkin kayıtlar esas alınmış olup, veriler denetim süreçlerinde kullanılacak analitik değerlendirmelere imkân tanıyacak şekilde ele alınmıştır. Verisetinde anomali oranı yaklaşık %5 düzeyinde olduğundan sınıflar arası dengesizlik söz konusudur. Bu nedenle model performansı yalnızca doğruluk ölçütü ile değerlendirilmemiş; Kesinlik (Precision), Duyarlılık (Recall), F1 skoru ve Matthews Korelasyon Katsayısı (MCC) gibi dengesiz veri setlerine duyarlı metrikler de kullanılmıştır. Ayrıca anomali sınıfına ilişkin yanlış negatif oranı özellikle izlenmiş ve denetim perspektifi açısından kritik performans göstergesi olarak ele alınmıştır.

Son olarak, çalışmada kullanılan etiketler gözetimsiz yöntemlerle üretilmiş olup, model gerçek usulsüzlük tespitinden ziyade potansiyel risk örüntülerinin belirlenmesine yöneliktir. Burada gözetimsiz etiketlerle ilgili olarak:

- Veri setinde önceden tanımlanmış gerçek risk/usulsüzlük etiketlerinin bulunmadığı,
- Bu nedenle anomali algoritmalarının potansiyel riskli kayıtları işaretlemek amacıyla kullanıldığı,
- Sınıflandırma modellerinin “gerçek suç” değil “potansiyel risk örüntülerini” öğrendiği,
- Bu yaklaşımın denetim süreçlerinde ön tarama ve karar destek amacı taşıdığı söylenebilir.

Bununla birlikte, uzman doğrulamasıyla bu etiketlerin anlamlılığına ilişkin ek değerlendirme de çalışmada sunulmuştur. Buna ek olarak, uzman incelemesi model tarafından riskli olarak belirlenen

kayıtların uygulama açısından anlamlı olup olmadığını değerlendirmek amacıyla gerçekleştirilmiştir. Ancak bu değerlendirme tüm veri setini kapsayan kapsamlı bir denetim çalışması niteliğinde değildir ve yalnızca seçilen kayıtlar üzerinden yapılmıştır. Uzman görüşleri mesleki deneyim ve bilgi birikimine dayandığından farklı uzman grupları tarafından kısmen farklı değerlendirmeler yapılması mümkündür. Bu nedenle uzman incelemesi, model sonuçlarının kesin doğrulaması olarak değil, elde edilen risk örüntülerinin uygulama açısından tutarlılığını ve anlamlılığını destekleyen ek bir değerlendirme olarak ele alınmalıdır. Yani, burada uzman doğrulaması yalnızca ön değerlendirme niteliğindedir. Nihai doğrulama çalışması olarak değerlendirilmemesi gerekmektedir. Örneklem büyüklüğü sınırlı olup daha sistematik klinik-denetim doğrulamaları da gereklidir.

Ayrıca, model çıktılarının gerçek ortamdaki doğruluğunu/geçerliliğini teyit etmek veya analiz etmek amacıyla, anomali olarak etiketlenen kayıtlar arasından rastgele 100 adet işlem kaydı seçilmiştir. Seçilen bu kayıtlar, alanında uzman denetçiler tarafından incelenmiştir. Yapılan incelemelerde, bu kayıtların önemli bir kısmının işlem süresi, işlem tekrarı ve mali göstergeler açısından olağan işlem davranışlarından sapma gösterdiği doğrulanmıştır. Bu bulgu, gözetimsiz öğrenme yöntemleriyle elde edilen anomali etiketli çıktıların anlamlı bir karşılığı olduğunu ortaya koymaktadır.

3.3. Veri Yapısı ve Değişken Grupları

Bu çalışmada kullanılan veriseti, işlem bazlı bir yapıya sahiptir ve her bir kayıt, belirli bir hastaya belirli bir tarih ve zaman aralığında uygulanan tek bir tıbbi işlemi temsil etmektedir. Veri seti; Sıra, Takip No, İşlem Tarihi, İşlem Başlangıç Saati, İşlem Bitiş Saati, Hekim ID, Hasta ID, Hasta Yaş, Hasta Engelli/Hak Sahibi Kodu, Hasta Kronik Hastalık Bilgisi, İşlem Kodu, İşlem Adı, SUT Adı, Radyolojik Görüntüleme Bilgisi, Fatura Tutarı, SUT Fiyatı ve Güvence Türü değişkenlerinden oluşmaktadır (Tablo 1). Bu değişkenler, denetim açısından anlamlı olacak şekilde kimlik ve takip bilgileri, zaman ve süreç göstergeleri, anonim hasta ve hekim bilgileri, klinik işlem alanları ile mali bileşenleri aynı gözlem düzeyinde bir araya getiren çok boyutlu bir veri yapısı sunmaktadır. Klinik işlemlere ilişkin alanlar ile mali göstergelerin birlikte yer aldığı bu yapı, işlem örüntülerinin bütüncül olarak analiz edilmesine ve denetim açısından olağandışı davranışların görünür hâle getirilmesine imkân sağlamaktadır.

Tablo 1: Ön İşlemeden Geçirilmiş Nihai Verisetinin Bir Kısmı

| Sıra | Takip No | İşlem Tarihi | İşlem Başlangıç Saati | İşlem Bitiş Saati | Hekim ID | Hasta ID | Hasta Yaş | Hasta Engelli / Hak Sahibi Kodu | Hasta Kronik Hastalık Bilgisi | İşlem Kodu | İşlem Adı | Sut Adı | Fatura Tutarı | SUT Fiyatı | Güvence Türü | İşlem Dakika | Radyolojik Görüntüleme |
|------|----------|--------------|-----------------------|-------------------|----------|----------|-----------|---------------------------------|-------------------------------|------------|--|--|---------------|------------|----------------------------------|--------------|------------------------|
| 1 | 4AQ3J58 | 11.02.2024 | 09:06:00 | 09:14:52 | 3E06 | 3855 | 9 | Yok | Yok | 40105 | Periapikal film, her biri (analog/dijital) | Periapikal film, her biri (analog/dijital) | 50.57 | 50.57 | SSK (SGK) SOSYAL GÜVENLİK KURUMU | 8.867 | Evet |
| 2 | 4AIJY59 | 5.02.2024 | 17:51:45 | 18:19:49 | 1C45 | 3832 | 37 | Yok | Yok | 40105 | Periapikal film, her biri (analog/dijital) | Periapikal film, her biri (analog/dijital) | 50.57 | 50.57 | SSK (SGK) SOSYAL GÜVENLİK KURUMU | 28.067 | Evet |
| 3 | 4AIJY59 | 5.02.2024 | 17:51:45 | 18:19:49 | 4G56 | 3832 | 37 | Yok | Yok | 40105 | Periapikal film, her biri (analog/dijital) | Periapikal film, her biri (analog/dijital) | 50.57 | 50.57 | SSK (SGK) SOSYAL GÜVENLİK KURUMU | 28.067 | Evet |
| 4 | 4AQ9IWP | 11.02.2024 | 15:26:15 | 15:31:49 | 3H74 | 1254 | 31 | Yok | Yok | 40105 | Periapikal film, her biri (analog/dijital) | Periapikal film, her biri (analog/dijital) | 50.57 | 50.57 | SSK (SGK) SOSYAL GÜVENLİK KURUMU | 5.567 | Evet |
| 5 | 4AQ9JRS | 11.02.2024 | 15:28:31 | 15:44:43 | 1F44 | 7365 | 28 | Yok | Yok | 40105 | Periapikal film, her biri (analog/dijital) | Periapikal film, her biri (analog/dijital) | 50.57 | 50.57 | SSK (SGK) SOSYAL GÜVENLİK KURUMU | 16.02 | Evet |
| 6 | 4BXWDRM | 12.03.2024 | 22:55:04 | 22:59:50 | 1C45 | 1057 | 31 | Yok | Yok | 40105 | Periapikal film, her biri (analog/dijital) | Periapikal film, her biri (analog/dijital) | 50.57 | 50.57 | SSK (SGK) SOSYAL GÜVENLİK KURUMU | 4.767 | Evet |
| 7 | 4BJ0I8L | 1.03.2024 | 18:24:06 | 18:30:45 | 3E06 | 2376 | 35 | Yok | Yok | 40105 | Periapikal film, her biri (analog/dijital) | Periapikal film, her biri (analog/dijital) | 50.57 | 50.57 | SSK (SGK) SOSYAL GÜVENLİK KURUMU | 6.65 | Evet |
| 8 | 4BJ2W2U | 1.03.2024 | 20:24:38 | 20:32:43 | 2L87 | 2458 | 12 | Yok | Yok | 40105 | Periapikal film, her biri (analog/dijital) | Periapikal film, her biri (analog/dijital) | 50.57 | 50.57 | SSK (SGK) SOSYAL GÜVENLİK KURUMU | 8.083 | Evet |
| 9 | 4BJCHHX | 2.03.2024 | 10:52:03 | 11:03:10 | 3E06 | 8351 | 40 | Yok | Yok | 40105 | Periapikal film, her biri (analog/dijital) | Periapikal film, her biri (analog/dijital) | 50.57 | 50.57 | SSK (SGK) SOSYAL GÜVENLİK KURUMU | 11.117 | Evet |
| 10 | 4BJH0L7 | 2.03.2024 | 14:07:08 | 14:32:25 | 1C45 | 1051 | 30 | Yok | Yok | 40105 | Periapikal film, her biri (analog/dijital) | Periapikal film, her biri (analog/dijital) | 50.57 | 50.57 | SSK (SGK) SOSYAL GÜVENLİK KURUMU | 25.283 | Evet |

3.4. Veri Ön İşleme, Temizleme Süreci ve Özellik Dönüştürme

HBYS üzerinden elde edilen ham işlem verileri, modelleme sürecinde güvenilir ve karşılaştırılabilir sonuçlar üretilebilmesi amacıyla ön işleme, temizleme ve özellik dönüştürme adımlarından geçirilmiştir. Altışar aylık iki ayrı veri dosyasının birleştirilmesi sonucunda elde edilen ham veri seti toplam 46409 işlem kaydından oluşmaktadır. Analitik sürece geçilmeden önce, veri setinin yapısal tutarlılığını sağlamak ve denetim açısından anlamlı olmayan kayıtları ayıklamak amacıyla kapsamlı bir ön işleme süreci yürütülmüştür.

Ön işleme aşamasında, veri setinde yer alan 3296 tamamen boş satır doğrudan analiz kapsamı dışında bırakılmıştır. Ayrıca işlem başlangıç zamanı bulunmasına rağmen işlem bitiş zamanı eksik olan ve bu nedenle işlem süresi hesaplanamayan 939 kayıt veri setinden çıkarılmıştır. Bu işlemler sonucunda analizlerde kullanılan nihai veri seti 42174 işlem kaydından oluşmuştur. Hasta ve hekim kimliklerine ilişkin alanlarda eksik değer bulunmadığı doğrulanmış; sınırlı sayıdaki takip numarası eksikliği ise işlem davranışı veya mali yapı üzerinde belirleyici bir etkisi bulunmadığından veri setinde korunmuştur.

Değişken bazlı incelemede, analitik katkı sunmayan veya klinik uygulamayı doğru biçimde yansıtmayan alanlar tespit edilmiştir. Özellikle “İşlem Yapılan Dış Sayısı” değişkeninin yalnızca 0 ve 1 değerlerinden oluşması ve birçok işlem türüyle uyumsuz biçimde doldurulması, bu değişkenin klinik gerçekliği yansıtmadığını göstermiştir. Benzer şekilde, tahlil değişkeninin tüm gözlemlerde aynı değeri alması ve klinik türü, hekim ünvanı ile hekim çalıştığı klinik alanlarının veri seti boyunca tekdüze değerler içermesi, bu değişkenlerin modelleme sürecinde ayırt edici bilgi sunmadığını ortaya koymuştur. Bu nedenle söz konusu alanlar analiz kapsamı dışında bırakılmıştır. Bu yaklaşım, veri setinin boyutunu küçültmekten ziyade bilgi yoğunluğunu artırmayı ve modelin öğrenme kapasitesini güçlendirmeyi amaçlayan seçici bir temizleme stratejisine dayanmaktadır.

Ön işleme sürecinde ayrıca kategorik değişkenlerdeki yazım farklılıkları, biçimsel tutarsızlıklar ve gereksiz karakterler giderilerek tüm kategoriler standart bir yapıya kavuşturulmuştur. Mali bileşenler arasında yer alan fatura tutarı ve SUT fiyatı gibi alanlar metinsel formatta bulunduğu için, uygun sayısal veri tiplerine dönüştürülmüş ve hesaplama elverişli hâle getirilmiştir. Hasta özelliklerine ilişkin bazı alanlarda boş görünen değerlerin, klinik uygulamada ilgili özelliğin bulunmadığını ifade ettiği değerlendirilmiş ve bu kayıtlar analitik anlam korunacak şekilde yeniden kodlanmıştır.

Bu kapsamda, veri ön işleme sürecinde özetle şu adımlar icra edilmiştir.

- Kategorik değişkenlerin modelleme sürecine uygun hale getirilmesi amacıyla label encoding ve one-hot encoding yöntemleri kullanılmıştır.
- Ağaç tabanlı algoritmalar ölçekleme gerektirmediğinden Random Forest ve Gradient Boosting modellerinde standardizasyon uygulanmamıştır.
- Mesafe tabanlı algoritmalarda (LOF, k-Means, DBSCAN) sayısal değişkenler üzerinde Min-Max normalization uygulanmamıştır.
- Tarih ve saat alanlarından işlem süresi gibi türetilmiş özelliklerin oluşturulmuştur.

Gerçekleştirilen ön işleme, temizleme ve özellik dönüştürme adımları sonucunda veri seti; eksik, tutarsız ve analitik açıdan katkı sunmayan unsurlardan arındırılmış, yapay zekâ tabanlı modelleme süreçlerinde kullanılmaya uygun, tutarlı ve denetim amaçlı güçlü bir analitik yapıya kavuşturulmuştur. Elde edilen bu yapı, izleyen aşamalarda uygulanan anomali tespiti, kümeleme ve sınıflandırma modelleri için sağlam bir veri temeli oluşturmuştur.

3.5. Kullanılan Yapay Zekâ Algoritmaları ve Modelleme Yaklaşımı

Bu çalışmada, ADMSM'den elde edilen işlem bazlı veriler üzerinde birden fazla yapay zekâ algoritmasına dayalı analitik model oluşturulmuştur. Veri setinin önceden tanımlanmış normal veya şüpheli etiketler içermemesi nedeniyle modelleme süreci öncelikle gözetimsiz (unsupervised)

yaklaşımlar üzerine kurgulanmıştır. Bu kapsamda, olağandışı işlem davranışlarının belirlenmesi amacıyla anomali tespiti ve kümeleme yöntemleri uygulanmış; potansiyel riskli olarak işaretlenen kayıtlar sonraki aşamada sınıflandırma modelleriyle değerlendirilmiştir.

Çok boyutlu, gürültülü ve hem sayısal hem kategorik değişkenler içeren veri yapısı; mesafe, yoğunluk ve ağaç tabanlı algoritmaların birlikte kullanılmasını gerekli kılmıştır. Klinik ve mali süreçlerde işlem davranışlarının heterojen bir yapı sergilemesi, kümeleme temelli yöntemlerle doğal örüntülerin ortaya çıkarılmasını mümkün kılarken, farklı türde değişkenlerin birlikte ele alınabilmesi ağaç tabanlı modellerin tercih edilmesini sağlamıştır.

Bu çerçevede altı temel algoritma kullanılmıştır. Kullanılan algoritmaların tercih edilme gerekçelerini şu şekilde özetlemek mümkündür:

- İzolasyon Ormanı (Isolation Forest) algoritmasının yüksek boyutlu verilerde seyrek anomalileri tespit etme başarısı,
- Uzaklık Faktörü (LOF) algoritmasının lokal yoğunluk farklılıklarını yakalama kapasitesi,
- k-Means algoritmasının yorumlanabilir segmentasyon sağlaması,
- DBSCAN algoritmasının yoğunluk temelli lokal yapıların belirlenmesindeki avantajı,
- Rastgele Orman (Random Forest) algoritmasının doğrusal olmayan ilişkileri modelleme başarısı,
- Gradyan Artırma (Gradient Boosting) algoritmasının hata azaltma yaklaşımı

Algoritma seçimleri yalnızca teknik tercih olarak değil, veri yapısının özellikleri ve denetim amaçları ile de ilişkilendirilmiştir.

Bu bağlamda, İzolasyon Ormanı gözlemleri rastgele bölmeler yoluyla izole ederek seyrek ve olağandışı kayıtları kısa izolasyon yolları üzerinden tespit eden, yüksek boyutlu veri setlerinde etkili bir anomali tespiti yöntemidir (Liu et al., 2008). Sağlık ve altyapı verileri üzerinde yapılan uygulamalar, bu algoritmanın uç değerleri belirlemede ölçeklenebilir ve başarılı sonuçlar üretebildiğini göstermektedir (Yarat & Orman, 2023).

Yerel Uzaklık Faktörü (LOF), bir gözlemin çevresindeki komşulara göre yerel yoğunluğunu değerlendirerek aykırılık skorları üretmekte olup, özellikle işlem süresi ve mali değişkenlerdeki lokal sapmaların belirlenmesinde kullanılmıştır (Dikbaş & Ebeğil, 2025). Buna ek olarak, k-Ortalamalar ve Yoğunluk Tabanlı Kümeleme Yöntemleri, işlem verilerindeki doğal kümeleri ve bu kümelere ayrılan kayıtları belirlemek amacıyla uygulanmıştır. Literatürde bu yöntemlerin farklı veri yoğunluklarına sahip yapılarda karşılaştırmalı olarak etkin sonuçlar verdiği belirtilmektedir (Bilgin & Çamurcu, 2005).

Gözetimsiz yöntemlerle potansiyel riskli olarak işaretlenen kayıtların ikinci aşamada değerlendirilmesi amacıyla ağaç tabanlı sınıflandırıcılar kullanılmıştır.

Rastgele Orman, çok sayıda karar ağacının birlikte çalışması sayesinde karmaşık ve doğrusal olmayan ilişkileri başarıyla modelleyebilmekte ve sağlık verilerinde yüksek sınıflandırma performansı sunmaktadır (Gür et al., 2024; Eşidir, 2025).

Gradyan Artırma ise ardışık öğrenciler aracılığıyla hata terimlerini azaltarak özellikle yüksek doğruluk gerektiren sınıflandırma problemlerinde etkili bir alternatif olarak değerlendirilmiştir (Veranyurt et al., 2020).

Bu çok aşamalı modelleme yaklaşımı sayesinde, farklı algoritmaların aynı veri yapısı üzerindeki davranışları karşılaştırılabilir hâle getirilmiş; anomali tespiti, kümeleme ve sınıflandırma çıktıları birlikte değerlendirilerek denetim açısından riskli işlem örüntülerinin belirlenmesi amaçlanmıştır. Model performansının genellebilirliğini değerlendirmek ve aşırı uyum (overfitting) riskini azaltmak amacıyla veri seti eğitim (%70) ve test (%30) alt kümelerine ayrılmıştır. Eğitim seti üzerinde 5-fold çapraz doğrulama uygulanmış ve her katmanda doğruluk, F1 skoru, Duyarlılık ve

AUC değerleri hesaplanarak ortalama performans ölçütleri raporlanmıştır. Rastgele Orman modeli için ayrıca Out-of-Bag (OOB) hata oranı hesaplanarak modelin genelleme kapasitesi ek olarak değerlendirilmiştir. 5-fold çapraz doğrulama sonuçları, Rastgele Orman modelinin ortalama AUC değerinin 0.994 ± 0.003 aralığında istikrarlı olduğunu göstermiştir. OOB hata oranı %1,8 olarak hesaplanmış ve modelin eğitim verisine aşırı uyum göstermediği değerlendirilmiştir.

4. BULGULAR

ADSM'ye ilişkin 2024-2025 dönemini kapsayan ve anonimleştirilmiş 42174 işlem kaydı üzerinde yürütülen analizler sonucunda, tıbbi işlem davranışlarının klinik ve mali boyutları birlikte değerlendirilerek risk göstergeleri üretilmiştir. Çalışmada önce gözetimsiz anomali tespiti ile olağandışı kayıtlar belirlenmiş, ardından kümeleme analizi ile işlem örüntüleri daha bağlamsal biçimde yorumlanmış, son aşamada ise anomali etiketleri hedef değişken olarak kullanılarak sınıflandırma modelleri kurulmuş ve performansları karşılaştırılmıştır. Bu bölümde, elde edilen bulgular tablolar ve şekillerle desteklenerek sunulmakta; ayrıca kullanılan sınıflandırma modellerinin performansları karşılaştırılarak en başarılı model belirlenmektedir.

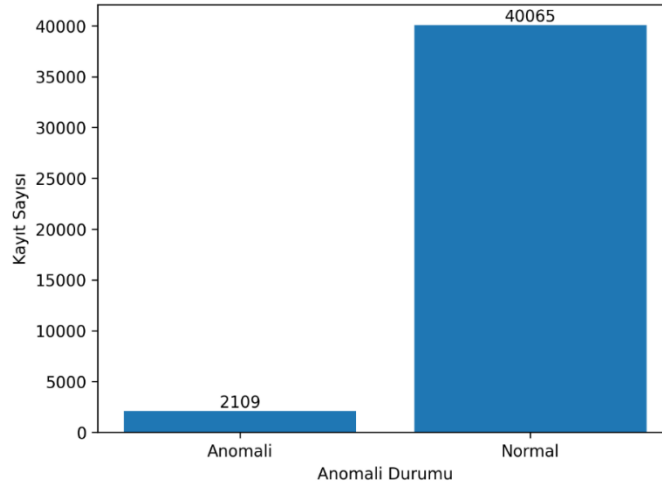
4.1. Anomali Tespiti Bulguları

ADSM'ye ait 2024-2025 dönemini kapsayan ve anonimleştirilmiş 42174 işlem kaydı üzerinde yürütülen anomali tespiti analizleri, tıbbi işlem davranışlarında olağan örüntülerden sapma gösteren kayıtların sistematik biçimde belirlenmesine imkân sağlamıştır. Önceden tanımlı normal veya şüpheli etiketlerin bulunmaması nedeniyle analiz süreci gözetimsiz öğrenme yaklaşımı çerçevesinde yürütülmüş; İzolasyon Ormanı (Isolation Forest) ve Yerel Uzaklık Faktörü (Local Outlier Factor - LOF) algoritmaları kullanılarak potansiyel riskli işlemler tespit edilmiştir. Söz konusu algoritmalarda kullanılan parametreler Tablo 2'de verilmiştir.

Tablo 2: Kullanılan parametreler

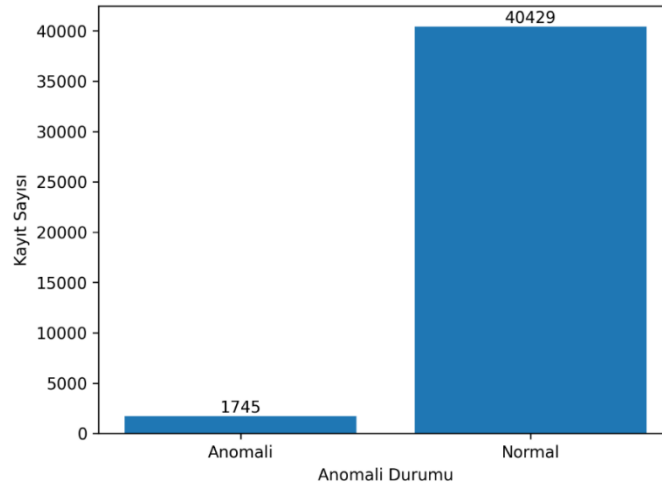
| İzolasyon Ormanı | Yerel Uzaklık Faktörü (LOF) |
|----------------------|-----------------------------|
| contamination = 0.05 | n_neighbors = 20 |
| n_estimators = 100 | |

İzolasyon Ormanı algoritması sonucunda toplam 2109 işlem kaydı anomali olarak işaretlenmiştir. Bu değer, tüm işlemlerin yaklaşık %5'ine karşılık gelmekte olup, veri seti içerisindeki nadir ve olağandışı işlem davranışlarını temsil etmektedir. Anomali olarak sınıflandırılan kayıtların özellikle işlem süresi, işlem tekrarı ve mali göstergeler açısından belirgin sapmalar içerdiği gözlemlenmiştir. Klinik olarak kısa sürede tamamlanması beklenen bazı işlemlerin olağandışı biçimde uzun sürdüğü, aynı işlem kodunun kısa zaman aralıklarında tekrarlandığı ve belirli mali tutarlarda yoğunlaşmaların olduğu kayıtlar, model tarafından tutarlı şekilde riskli olarak değerlendirilmiştir. İzolasyon Ormanı algoritması tarafından anomali ve normal olarak sınıflandırılan işlem kayıtlarının dağılımı Şekil 2'de sunulmaktadır.



Şekil 2: İzolasyon Ormanı Algoritması ile Tespit Edilen Anomali ve Normal Kayıtların Dağılımı

Yerel Uzaklık Faktörü algoritması kullanılarak gerçekleştirilen anomali tespiti analizinde, toplam 1745 işlem kaydı olağandışı olarak sınıflandırılmıştır. Bu kayıtlar, veri setinin yaklaşık %4,1'ine karşılık gelmekte olup özellikle yerel yoğunluk farklılıkları üzerinden ayrışan işlem davranışlarını temsil etmektedir. Tespit edilen anomalilerin büyük bölümü, benzer klinik koşullarda gerçekleştirilen işlemlerle karşılaştırıldığında beklenen işlem yoğunluğundan sapma gösteren kayıtlardan oluşmaktadır. Özellikle bazı hekimler tarafından aynı gün içerisinde gerçekleştirilen işlem sayısının diğer hekimlere kıyasla belirgin biçimde yüksek olduğu durumlar, belirli işlem türlerinde lokal yoğunlaşmalar ve işlem süresi ile işlem türü arasındaki uyumsuzluklar LOF analizinde ayırt edici bulgular olarak öne çıkmıştır. Yerel Uzaklık Faktörü algoritması tarafından anomali ve normal olarak sınıflandırılan işlem kayıtlarının dağılımı Şekil 3'te sunulmaktadır.



Şekil 3: Yerel Uzaklık Faktörü Algoritması ile Tespit Edilen Anomali ve Normal Kayıtların Dağılımı

Her iki algoritmanın sonuçları birlikte değerlendirildiğinde, önemli ölçüde örtüşen risk alanlarının bulunduğu görülmektedir. Bu alanlar; beklenenden uzun işlem süreleri, kısa zaman aralıklarında tekrarlanan işlem girişleri, SUT fiyatı ile fatura tutarı arasındaki uyumsuzluklar ve hekim bazlı olağandışı işlem yoğunlukları şeklinde özetlenebilir. İzolasyon Ormanı algoritması daha çok veri seti genelindeki küresel sapmaları ön plana çıkarırken, LOF algoritmasının lokal yoğunluk farklılıklarını yakalamada daha duyarlı olduğu gözlemlenmiştir.

Anomali tespiti aşamasında elde edilen bu bulgular, tıbbi işlem verilerinde gözlenen uyumsuzlukların tekil hatalardan ziyade çok boyutlu davranış örüntülerine dayandığını

göstermektedir. Anomali olarak işaretlenen işlem kayıtları, izleyen aşamalarda gerçekleştirilen kümeleme ve sınıflandırma analizleri için temel girdi niteliği taşımış; riskli işlem davranışlarının daha ayrıntılı biçimde incelenmesine olanak sağlamıştır.

4.2. Kümeleme Analizi Bulguları

Anomali tespiti aşamasında elde edilen sonuçların daha bağlamsal biçimde yorumlanabilmesi ve işlem davranışlarının genel örüntülerinin ortaya konulabilmesi amacıyla, veri seti üzerinde kümeleme analizleri gerçekleştirilmiştir. Bu kapsamda, farklı varsayımlara dayanan k-Ortalamlar (k-Means) ve Yoğunluk Tabanlı Uzamsal Kümeleme (DBSCAN) algoritmaları kullanılarak işlem kayıtlarının benzerlik ve yoğunluk yapılarına göre nasıl gruplandığı incelenmiştir. DBSCAN algoritmasında eps ve minPts değerleri k-distance analizi ile belirlenmiştir.

Kümeleme çıktıları, anomali tespiti sonuçlarıyla birlikte değerlendirilerek işlem davranışlarının yapısal özellikleri analiz edilmiştir.

k-Ortalamlar algoritmasında optimal küme sayısı Elbow yöntemi ve Silhouette katsayısı analizleri ile belirlenmiştir. k=3 değeri, küme içi varyansın marjinal azalma hızının düştüğü nokta ve en yüksek ortalama Silhouette değeri dikkate alınarak seçilmiştir. DBSCAN algoritmasında epsilon parametresi k-distance grafiği üzerinden belirlenmiş; minPts değeri veri boyutu ve boyutsallık dikkate alınarak ayarlanmıştır. Parametre duyarlılık analizleri sonucunda model çıktılarının istikrarlı olduğu gözlemlenmiştir.

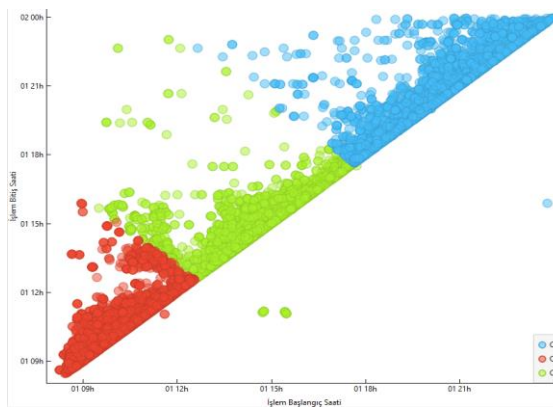
k-Ortalamlar algoritması kullanılarak gerçekleştirilen analiz sonucunda, toplam 42174 işlem kaydı üç ana kümeye ayrılmıştır. Kümelere ait kayıt sayıları ve oranlar Tablo 3'te sunulmaktadır.

Tablo 3: k-Ortalamlar Sonuçlarına Göre Kümelerin Kayıt Dağılımı

| Küme | Kayıt Sayısı | Oran (%) |
|------|--------------|----------|
| C1 | 9.159 | %21,7 |
| C2 | 15.022 | %35,6 |
| C3 | 17.993 | %42,7 |

Elde edilen küme yapısı, tıbbi işlem davranışlarının tek tip bir yapı sergilemediğini; işlem süresi, mali göstergeler ve işlem türleri bakımından farklılaşan üç temel örüntünün bulunduğunu göstermektedir. Kümelere ait kayıt sayıları ve oranlar Tablo 3'te sunulmaktadır. Kümelere ait kayıt sayıları ve oranlar Tablo 3'te sunulmaktadır. Kümelere ait kayıt sayıları ve oranlar Tablo 3'te sunulmaktadır.

Kümelere ait kayıt sayıları ve oranlar Tablo 3'te sunulmaktadır. Kümelere ait kayıt sayıları ve oranlar Tablo 3'te sunulmaktadır. Kümelere ait kayıt sayıları ve oranlar Tablo 3'te sunulmaktadır.



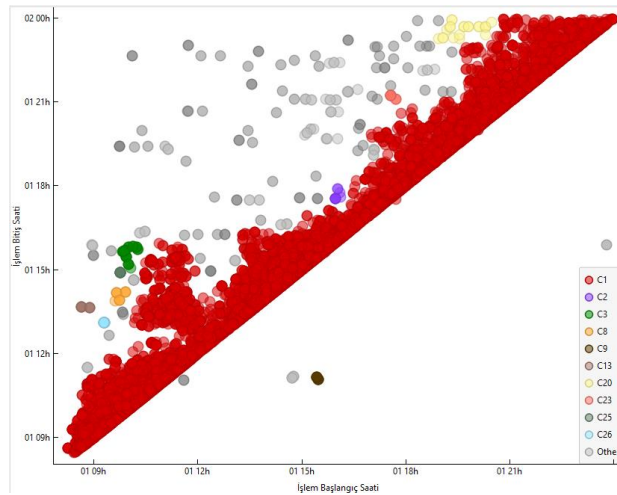
Şekil 4: Elde Edilen Kümelerin İşlem Zamanları Bağlamındaki Dağılımı

Yoğunluk Tabanlı Uzamsal Kümeleme algoritması ile gerçekleştirilen analizde ise işlem kayıtlarının büyük bölümünün tek bir baskın küme altında toplandığı görülmüştür. Sonuçlara göre oluşan küme yapısına ilişkin kayıt sayıları Tablo 4’te sunulmaktadır. Analiz sonuçları, veri setinin genel olarak yüksek yoğunluklu ve tutarlı bir işlem davranışı sergilediğini; yalnızca sınırlı sayıda kaydın küçük ve yerel kümeler içerisinde ayrıştığını göstermektedir. Bu küçük kümeler, belirli işlem türlerinde yoğunlaşma, hekim bazlı lokal işlem davranışları veya kısa zaman aralıklarında tekrar eden işlemler gibi yerel yoğunluk farklılıklarını yansıtmaktadır.

Tablo 4: Yoğunluk Tabanlı Uzamsal Kümeleme Sonuçlarına Göre Kümelerin Kayıt Dağılımı

| Küme | Kayıt Sayısı |
|-----------|---------------------|
| C1 | 41.806 |
| C2-C33 | Her biri 4-32 arası |
| Noise(-1) | 0 |

Yoğunluk Tabanlı Uzamsal Kümeleme algoritmasının herhangi bir kaydı tamamen ayırık (noise) olarak işaretlememesi, incelenen işlem verilerinin genel olarak süreklilik gösteren ve kopukluk içermeyen bir yapıya sahip olduğunu ortaya koymaktadır. Elde edilen küme yapısının işlem zamanları bağlamındaki dağılımı Şekil 5’te sunulmaktadır.



Şekil 5: Elde Edilen Kümelerin İşlem Zamanları Bağlamındaki Dağılımı

k-Ortalamalar ve Yoğunluk Tabanlı Uzamsal Kümeleme algoritmalarından elde edilen sonuçlar birlikte değerlendirildiğinde, iki yöntemin işlem verisini farklı analitik bakış açılarıyla ele aldığı görülmektedir. K-Ortalamalar algoritması, işlem davranışlarını daha yorumlanabilir ve karşılaştırılabilir ana gruplar altında toplamayı başarırken; Yoğunluk Tabanlı Uzamsal Kümeleme algoritması veri setinin genel yoğunluk ve tutarlılık yapısını ortaya koyan tamamlayıcı bir analiz sunmuştur. Bu bağlamda, k-Ortalamalar algoritması, anomali tespiti sonuçlarının kümeler içerisindeki dağılımının incelenmesi ve izleyen sınıflandırma aşamasının yapılandırılması açısından daha işlevsel bir segmentasyon sağlamıştır.

Kümeleme analizlerinden elde edilen bu bulgular, olağandışı işlem kayıtlarının hangi işlem gruplarında yoğunlaştığını ortaya koymakta ve sınıflandırma modelleri için bağlamsal bir referans oluşturmaktadır.

4.3. Sınıflandırma Modeli Bulguları

Anomali tespiti aşamasında potansiyel riskli olarak işaretlenen işlem kayıtlarının hangi klinik ve mali değişkenler üzerinden ayrıştığını belirlemek amacıyla, gözetimli sınıflandırma modelleri oluşturulmuştur. Bu aşamada, İzolasyon Ormanı algoritması tarafından üretilen anomali etiketleri (normal/şüpheli) hedef değişken olarak kullanılmış; işlem süresi, işlem kodları, mali göstergeler,

hasta profiline ilişkin bilgiler ve güvence türü gibi değişkenler açıklayıcı öznitelikler olarak modele dâhil edilmiştir.

Sınıflandırma analizlerinde Rastgele Orman ve Gradyan Artırma algoritmaları aynı veri yapısı ve aynı hedef değişken kullanılarak ayrı ayrı modellenmiş, böylece modellerin karşılaştırılabilir sonuçlar üretmesi sağlanmıştır. Her iki model için de karışıklık matrisleri, ROC eğrileri ve temel performans ölçütleri üzerinden değerlendirme yapılmıştır. Algoritmalarda kullanılan parametreler Tablo 5’te verilmiştir.

Tablo 5: Kullanılan parametreler

| Random Forest | Gradient Boosting |
|-------------------------|----------------------|
| n_estimators = 200 | n_estimators = 150 |
| max_depth = None | learning_rate = 0.05 |
| min_samples_split = 2 | max_depth = 3 |
| class_weight = balanced | |

Rastgele Orman modeli sonuçları incelendiğinde, normal ve şüpheli işlem sınıflarını ayırt etmede oldukça yüksek bir başarı sergilediği görülmektedir. Karışıklık matrisi sonuçları Şekil 6’da sunulmaktadır. Normal işlem sınıfına ait kayıtların büyük bölümü doğru biçimde sınıflandırılmış; yanlış pozitif oranının oldukça düşük seviyede kaldığı gözlemlenmiştir. Şüpheli işlem sınıfında ise yüksek sayıda doğru pozitif tespit gerçekleştirilmiş olup, modelin riskli işlemleri yakalama kapasitesinin güçlü olduğu görülmektedir.

| | | Tahmin | | Σ |
|--------|-----|--------|-------|-------|
| | | Yes | No | |
| Gerçek | Yes | 1858 | 251 | 2109 |
| | No | 66 | 39999 | 40065 |
| Σ | | 1924 | 40250 | 42174 |

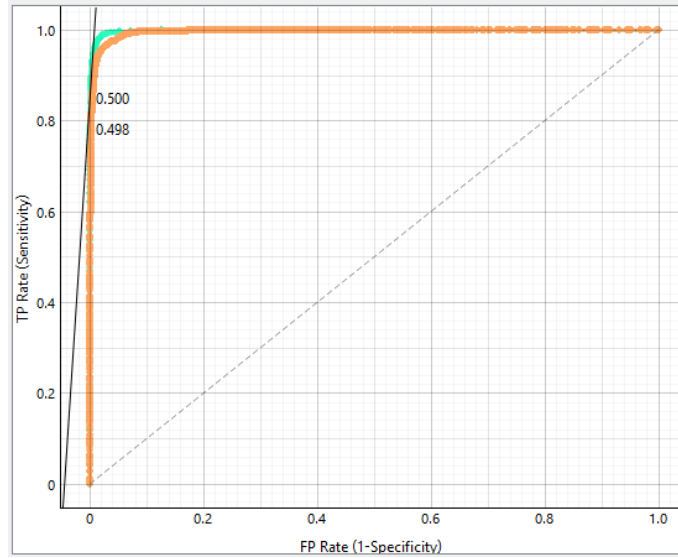
Şekil 6: Rastgele Orman Karışıklık Matris Sonuçları

Gradyan Artırma modeli için elde edilen karışıklık matrisi sonuçları Şekil 7’de sunulmaktadır. Bu model de genel olarak yüksek doğrulukla çalışmakla birlikte, özellikle şüpheli işlem sınıfında Rastgele Orman modeline kıyasla daha fazla sayıda yanlış negatif üretmiştir. Başka bir ifadeyle, bazı potansiyel riskli işlemler Gradyan Artırma modeli tarafından normal sınıfta değerlendirilmiştir. Denetim perspektifinden bakıldığında, yanlış negatiflerin yüksek olması riskli işlemlerin gözden kaçmasına yol açabileceğinden, bu durum modelin görelî bir zayıflığı olarak değerlendirilmektedir.

| | | Tahmin | | Σ |
|--------|-----|--------|-------|-------|
| | | Yes | No | |
| Gerçek | Yes | 1714 | 395 | 2109 |
| | No | 78 | 39987 | 40065 |
| Σ | | 1792 | 40382 | 42174 |

Şekil 7: Gradyan Artırma Karışıklık Matris Sonuçları

Sınıflandırma modellerinin ayırt edicilik performansları ROC eğrileri aracılığıyla değerlendirilmiştir. Rastgele Orman ve Gradyan Artırma modellerine ait ROC eğrileri Şekil 8’de sunulmaktadır. Rastgele Orman modeline ait ROC eğrisinin sol üst köşeye daha yakın konumlandığı ve daha geniş bir alan kapladığı görülmekte; bu durum modelin farklı eşik değerlerinde daha dengeli bir doğru pozitif-yanlış pozitif oranı sunduğunu göstermektedir.



Şekil 8: Rastgele Orman ve Gradyan Artırma Modellerine Ait ROC Eğrileri

Sınıflandırma analizleri kapsamında ayrıca Rastgele Orman modeli üzerinden değişken önem analizi gerçekleştirilmiştir. Değişken önem düzeylerine ilişkin sonuçlar Şekil 9’da sunulmaktadır. Analiz sonuçlarına göre, Güvence Türü, Hasta Kronik Hastalık Bilgisi, İşlem Adı ve İşlem Kodu değişkenleri modelin karar sürecine en yüksek katkıyı sağlayan öznitelikler olarak öne çıkmaktadır. Mali göstergeler ve işlem süresi değişkeni ise orta düzeyde ayırt edici etki göstermiştir. Bu bulgu, şüpheli işlem davranışlarının yalnızca mali veya zamansal göstergelerle değil, hasta profili ve işlem türleriyle birlikte oluşan çok boyutlu örüntüler üzerinden şekillendiğini gösterir.

| | # | Gain ratio | Gini |
|---|----|------------|-------|
| 1 C Güvence Türü | 7 | 0.470 | 0.021 |
| 2 C Hasta Kronik Hastalık Bilgisi | 6 | 0.038 | 0.000 |
| 3 C İşlem Adı | 57 | 0.036 | 0.029 |
| 4 N İşlem Kodu | | 0.022 | 0.005 |
| 5 C Radyolojik Görüntüleme | 2 | 0.017 | 0.002 |
| 6 N SUT Fiyatı | | 0.016 | 0.004 |
| 7 N Fatura Tutarı | | 0.015 | 0.004 |
| 8 N İşlem Dakika | | 0.009 | 0.003 |
| 9 T İşlem Başlangıç Saati | | 0.003 | 0.001 |
| 10 N Hasta ID | | 0.001 | 0.000 |
| 11 N Hekim ID | | 0.001 | 0.000 |
| 12 T İşlem Bitiş Saati | | 0.001 | 0.000 |
| 13 N Hasta Yaş | | 0.001 | 0.000 |
| 14 C Hasta Engelli / Hak Sahibi Kodu | 2 | 0.001 | 0.000 |

Şekil 9: Rastgele Orman Modeli İçin Değişken Önem Düzeyleri (Rank Analizi)

Elde edilen sınıflandırma bulguları, anomali tespiti aşamasında işaretlenen kayıtların klinik ve mali değişkenler temelinde yüksek doğrulukla ayrıştırılabildiğini göstermektedir. Bu sonuçlar, yapılacak model performans karşılaştırması ve en başarılı modelin seçimine ilişkin değerlendirmeler için güçlü bir analitik temel oluşturmaktadır.

4.4. Model Performans Karşılaştırması ve En Başarılı Modelin Seçilmesi

Bu bölümde, oluşturulan sınıflandırma modellerinin performansları karşılaştırmalı olarak değerlendirilmiş ve denetim amaçlı kullanım açısından en uygun model belirlenmiştir. Çalışmada kullanılan algoritmalar farklı öğrenme yaklaşımlarına dayandığından, doğrudan nicel performans karşılaştırması yalnızca gözetimli sınıflandırma modelleri için gerçekleştirilmiştir. Gözetimsiz anomali tespiti ve kümeleme algoritmaları, keşifsel ve destekleyici analiz araçları olarak ele alınmış; performans karşılaştırmasına dâhil edilmemiştir.

Bu kapsamda, Rastgele Orman ve Gradyan Artırma modelleri aynı veri yapısı ve aynı hedef değişken kullanılarak eğitilmiş; modellerin başarımlarını doğruluk (CA), F1 skoru, AUC ve Matthews korelasyon katsayısı (MCC) ölçütleri üzerinden değerlendirilmiştir. Veri setindeki sınıf dengesizliği nedeniyle model performansı yalnızca ROC eğrisi ile değil, Kesinlik-Duyarlılık (PR) eğrisi üzerinden de değerlendirilmiştir. Özellikle riskli işlemlerin tespiti bağlamında yanlış negatiflerin minimize edilmesi kritik olduğundan, anomali sınıfına ilişkin Duyarlılık değeri öncelikli performans göstergesi olarak ele alınmıştır.

Kesinlik-Duyarlılık eğrisi analizinde Rastgele Orman modelinin Average Kesinlik (AP) değeri 0.987 olarak hesaplanmıştır. %5 risk diliminde Kesinlik değeri %96 olarak belirlenmiş; bu durum modelin yüksek riskli işlemleri doğru sıralama kapasitesinin güçlü olduğunu göstermektedir. Gradyan Artırma modeli için AP değeri 0.973 olarak hesaplanmıştır.

Tablo 6: Rastgele Orman ve Gradyan Artırma Performans Ölçütleri

| Model | AUC | CA | F1 | Kesinlik | Duyarlılık | MCC |
|-----------------|-------|-------|-------|----------|------------|-------|
| Rastgele Orman | 0.999 | 0.992 | 0.992 | 0.992 | 0.992 | 0.919 |
| Gradyan Artırma | 0.997 | 0.989 | 0.988 | 0.989 | 0.989 | 0.876 |

Tablo 6 incelendiğinde, Rastgele Orman modelinin tüm performans ölçütlerinde Gradyan Artırma modeline kıyasla daha yüksek değerlere ulaştığı görülmektedir. Özellikle sınıflar arası dengeyi yansıtan F1 skoru ve MCC ölçütlerinde Rastgele Orman modelinin daha istikrarlı bir performans sergilediği dikkat çekmektedir (AUC:0.999). Elde edilen yüksek sınıflandırma başarısı (AUC: 0.999), modelin gözetimsiz öğrenme aşamasında belirlenen anomali örüntülerini yüksek bir tutarlılıkla ayırt edebildiğini göstermektedir. Bu yüksek AUC değeri, veri yapısındaki belirgin ayrışmadan kaynaklanmaktadır. Bu durum, modelin hem normal hem de şüpheli işlem sınıflarını ayırt etme kapasitesinin daha güçlü olduğunu göstermektedir. Elde edilen yüksek performans değerleri, kullanılan etiketlerin gözetimsiz öğrenme yöntemleriyle oluşturulmuş olması nedeniyle dikkatli yorumlanmalıdır.

Bu aşamada Rastgele Orman ve Gradyan Artırma modellerinin temel kullanım amacı, sadece bir tahminleme mekanizması oluşturmak değil; aynı zamanda hangi değişkenlerin (özelliklerin) anomaliler üzerinde daha belirleyici olduğunu saptayarak modelin açıklanabilirliğini (explainability) artırmaktır. Böylece denetim sürecinde riskli olarak etiketlenen işlemlerin hangi kriterlere göre bu sınıfa dahil edildiği şeffaf bir şekilde ortaya konulmaktadır.

Karışıklık matrisi sonuçları da bu bulguları desteklemektedir. Rastgele Orman modeli, şüpheli işlem sınıfında daha yüksek sayıda doğru pozitif üretmiş ve riskli işlemlerin gözden kaçma olasılığını daha düşük seviyede tutmuştur. Buna karşılık Gradyan Artırma modelinde, özellikle şüpheli sınıfta yanlış negatif sayısının daha yüksek olduğu gözlemlenmiştir. Denetim perspektifinden

değerlendirildiğinde, yanlış negatiflerin artması potansiyel riskli işlemlerin inceleme dışı kalmasına yol açabileceğinden, bu durum önemli bir dezavantaj olarak değerlendirilmektedir.

Modellerin ayırt edicilik düzeyleri ROC eğrileri üzerinden de analiz edilmiştir. Rastgele Orman modeline ait ROC eğrisinin daha geniş bir alan kapladığı ve ideal sınıflandırma noktasına daha yakın konumlandığı görülmektedir. Bu bulgu, modelin farklı eşik değerlerinde daha dengeli bir doğru pozitif-yanlış pozitif oranı sunduğunu ve ayırt edicilik gücünün daha yüksek olduğunu göstermektedir.

Tüm nicel ve görsel değerlendirmeler birlikte ele alındığında, Rastgele Orman algoritmasının bu veri seti üzerinde daha dengeli, güvenilir ve denetim açısından daha uygun sonuçlar ürettiği sonucuna ulaşılmıştır. Bu nedenle, önerilen risk tabanlı denetim yaklaşımında Rastgele Orman modeli en başarılı model olarak seçilmiştir.

5. TARTIŞMA

Bu çalışma kapsamında ADSM'ye ait işlem bazlı veriler üzerinde uygulanan yapay zekâ temelli analizler, tıbbi işlem davranışlarının klinik ve mali boyutlarıyla birlikte değerlendirilmesine olanak sağlamıştır. Anomali tespiti, kümeleme ve sınıflandırma modellerinden elde edilen bulgular, işlem kayıtlarında gözlenen uyumsuzlukların tek bir değişkenle açıklanamayacak çok boyutlu bir yapıya sahip olduğunu desteklemektedir.

Anomali tespiti analizleri sonucunda, özellikle işlem süresi, işlem tekrar sıklığı ve mali göstergeler açısından olağan işlem örüntülerinden sapma gösteren kayıtların bulunduğu tespit edilmiştir. İzolasyon Ormanı algoritması tarafından belirlenen anomaliler, veri seti genelinde nadir görülen ve beklenen işlem davranışlarından ayrılan kayıtları ön plana çıkarırken; Yerel Uzaklık Faktörü algoritması, benzer işlem yoğunlukları içerisinde lokal farklılaşmalar gösteren kayıtları görünür hâle getirmiştir. Her iki algoritma tarafından anomali olarak işaretlenen kayıtların önemli bir bölümünün klinik uygulama süresi ile mali göstergeler arasındaki uyumsuzluklar açısından dikkat çekici özellikler taşıdığı görülmüştür.

Kümeleme analizleri, işlem davranışlarının genel yapısının değerlendirilmesine ve anomali tespiti sonuçlarının bağlamsal olarak yorumlanmasına katkı sağlamıştır. K-Ortalamalar yöntemi ile elde edilen kümeler, veri setinde farklı işlem davranış örüntülerinin bulunduğunu ve işlemlerin klinik ve mali özellikler açısından homojen olmayan gruplar altında toplandığını göstermektedir. Buna karşılık Yoğunluk Tabanlı Uzamsal Kümeleme yöntemi, veri setinin büyük bölümünün tek bir baskın yoğunluk yapısı sergilediğini ortaya koymuş; yalnızca sınırlı sayıda kayıtların küçük ve lokal kümeler oluşturduğunu göstermiştir. Bu bulgular, ADSM işlem süreçlerinin genel olarak tutarlı bir yapı sergilediğini, ancak belirli alt gruplarda dikkat gerektiren farklılıklar bulunduğunu göstermektedir.

Sınıflandırma modellerinden elde edilen sonuçlar, anomali olarak işaretlenen kayıtların klinik ve mali değişkenler temelinde yüksek doğrulukla ayrıştırılabildiğini göstermektedir. Rastgele Orman ve Gradyan Artırma modelleri, anomali ve normal işlem sınıflarını ayırt etmede başarılı performans sergilemiş; ancak performans metrikleri ve hata dağılımları açısından modeller arasında farklılıklar gözlemlenmiştir. Özellikle Rastgele Orman modelinin, anomali sınıfında daha dengeli bir sınıflandırma performansı sunduğu ve daha az sayıda riskli kaydı normal sınıfta değerlendirdiği görülmüştür.

Değişken önem analizleri, anomali davranışlarının yalnızca operasyonel göstergelerle sınırlı olmamakla birlikte; hasta profili, güvence türü ve işlem türü gibi değişkenlerin de sınıflandırma sürecinde belirleyici rol oynamaktadır. Güvence türü, hasta kronik hastalık bilgisi, işlem kodları ve mali göstergelerin öne çıkması, klinik uygulamalar ile mali yapı arasındaki ilişkinin riskli işlem davranışlarının oluşumunda etkili olduğunu göstermektedir.

Ayrıca, model tarafından anomali olarak işaretlenen kayıtların bir alt kümesi üzerinde gerçekleştirilen uzman incelemesi, bu kayıtların önemli bir kısmının denetim açısından anlamlı risk göstergeleri içermektedir.

Genel olarak değerlendirildiğinde, elde edilen bulgular tıbbi işlem kayıtlarında gözlenen uyumsuzlukların rastlantısal değil, belirli klinik ve mali örüntüler etrafında şekillendiğini göstermektedir. Anomali tespiti, kümeleme ve sınıflandırma analizlerinin birlikte kullanılması, işlem bazlı risklerin daha görünür hâle gelmesini sağlamış ve denetim süreçleri açısından anlamlı bir analitik çerçeve sunmuştur.

Bu noktada çalışma kapsamındaki sınırlılıkları;

- Verilerin yalnızca tek bir ADSM'den elde edilmesi (bu yüzden model, spesifik örüntüler öğrenmiş olma handikabına sahiptir),
- Kurumsal işleyiş farklılıklarının modele özgü örüntüler üretmiş olabileceği,
- Gerçek doğrulanmış usulsüzlük etiketlerinin bulunmaması,
- Uzman doğrulama örnekleminin sınırlı olması,
- Zaman serisi veya çok merkezli doğrulama yapılmamış olması,
- Modelin farklı kurumlarda yeniden test edilmesi gerekliliği olarak özetlemek mümkündür.

Bu çalışmada kullanılan risk etiketleri, doğrulanmış denetim bulgularına veya usulsüzlük kayıtlarına değil, gözetimsiz öğrenme yöntemleriyle belirlenen risk örüntülerine dayanmaktadır. Bu nedenle elde edilen performans sonuçları, modelin doğrulanmış usulsüzlükleri tespit etme başarısını değil, veri içerisindeki potansiyel riskli davranışları ayırt etme yeteneğini göstermektedir. Gözetimsiz öğrenme yöntemleri riskli görünen kayıtların belirlenmesinde önemli avantajlar sağlamakla birlikte, tespit edilen her riskli kaydın gerçek bir usulsüzlük veya denetim bulgusu olduğu söylenemez. Bu nedenle bulgular dikkatli değerlendirilmeli ve model çıktıları nihai karar verme aracı olarak değil, denetim süreçlerinde önceliklendirme ve karar destek amacıyla kullanılmalıdır.

6. SONUÇ

Bu çalışmada, kamu sağlık kurumlarında yürütülen tıbbi işlemlere ait büyük hacimli veriler kullanılarak yapay zekâ destekli ve risk tabanlı bir denetim modeli geliştirilmiştir. Başka bir ifadeyle çalışmada, kamu sağlık kurumlarında tıbbi işlem verilerinin risk temelli değerlendirilmesine yönelik veri odaklı yaklaşımların önemi ortaya konulmuştur.

ADSM'ye ait 2024-2025 dönemini kapsayan ve anonimleştirilmiş işlem kaydı üzerinde gerçekleştirilen analizler, tıbbi işlem davranışlarının klinik ve mali boyutlarıyla birlikte değerlendirilmesine imkân tanımıştır. Çalışma kapsamında gözetimsiz anomali tespiti, kümeleme ve gözetimli sınıflandırma yaklaşımları bütüncül bir analitik yapı içerisinde ele alınmıştır.

Anomali tespiti aşamasında uygulanan İzolasyon Ormanı ve Yerel Uzaklık Faktörü algoritmaları, olağan işlem davranışlarından sapma gösteren kayıtların sistematik biçimde belirlenmesini sağlamıştır. Bu yöntemler, işlem süresi, işlem tekrarı ve mali göstergeler açısından dikkat çeken sapmaları görünür hâle getirerek potansiyel risk taşıyan işlemlerin ön plana çıkarılmasına katkı sunmuştur. Kümeleme analizleri, işlem verilerinin genel olarak tutarlı bir yapı sergilediğini, ancak belirli alt gruplarda klinik ve mali açıdan farklılaşan işlem örüntülerinin bulunduğunu ortaya koymuştur. Gözetimsiz yöntemlerle elde edilen anomali etiketleri kullanılarak oluşturulan ağaç tabanlı sınıflandırma modelleri, riskli ve normal işlem kayıtlarının yüksek doğrulukla ayrıştırılabildiğini göstermiştir. Performans karşılaştırmaları sonucunda Rastgele Orman modelinin doğruluk, F1 skoru ve AUC gibi ölçütlerde daha dengeli ve güçlü sonuçlar ürettiği belirlenmiş; bu model çalışmada en başarılı sınıflandırma yaklaşımı olarak öne çıkmıştır. Değişken önem analizleri, anomali davranışlarının işlem türü, hasta profili, güvence yapısı ve mali göstergelerle birlikte şekillendiğini ortaya koymuştur.

Elde edilen bulgular, tıbbi işlem verilerinde gözlenen uyumsuzlukların tekil hatalardan ziyade çok boyutlu ve yapısal örüntülerden kaynaklandığını göstermektedir. Geliştirilen yapay zekâ destekli risk tabanlı denetim modeli, kamu sağlık kurumlarında denetim süreçlerinin daha hedefli, veri temelli ve önceliklendirilmiş biçimde yürütülmesine katkı sağlayabilecek nitelikte bir analitik çerçeve sunmaktadır. Bu yaklaşım, denetim kaynaklarının etkin kullanımını destekleyerek riskli işlemlerin

erken aşamada belirlenmesine olanak tanımaktadır. Çalışma kapsamında önerilen model, kamu sağlık kurumlarında tıbbi işlem verilerinin risk temelli değerlendirilmesine yönelik uygulanabilir ve bütüncül bir yöntem ortaya koymaktadır. Yapay zekâ destekli analitik yaklaşımların denetim süreçlerine entegrasyonu, sağlık hizmetlerinde mali disiplin, şeffaflık ve hesap verebilirliğin güçlendirilmesine yönelik önemli bir potansiyel sunmaktadır.

Ayrıca önerilen model, denetçilerin yerini alan otomatik bir karar sistemi değil; riskli işlemlerin önceliklendirilmesini destekleyen veri temelli bir karar destek aracı niteliğindedir. Bu çalışmanın önemli sınırlılıklarından biri, model performansının doğrulanmış denetim bulguları yerine gözetimsiz öğrenme yöntemleriyle oluşturulan risk etiketleri üzerinden değerlendirilmiş olmasıdır. Bu nedenle elde edilen sonuçlar, potansiyel risk örüntülerinin belirlenmesine yönelik bir değerlendirme olarak ele alınmalıdır. Gelecekte gerçek denetim bulguları, disiplin soruşturmaları veya doğrulanmış usulsüzlük kayıtlarından oluşan veri setleri kullanılarak yapılacak çalışmalar, modelin farklı koşullardaki performansının daha kapsamlı biçimde değerlendirilmesine ve elde edilen bulguların desteklenmesine katkı sağlayacaktır. Son olarak, gelecekte yapılabilecek çalışmalar için çok kurumlu veri setleri, gerçek denetim sonuçları ile doğrulama, yarı gözetimli öğrenme ve açıklanabilir yapay zeka (Explainable AI) yöntemleri önerilmektedir.

KAYNAKÇA

- Albayrak, M. (2022). Türkiye’de sağlık harcamaları ile yolsuzluk arasındaki dinamik ilişkinin ARDL yaklaşımıyla incelenmesi. *Pamukkale Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, (51), 339–355.
- Arslan Kurtuluş, S. (2021). Sağlık hizmetlerinde kontrol. *Cumhuriyet Üniversitesi Fen-Edebiyat Fakültesi Sosyal Bilimler Dergisi*, 45(2), 77–100.
- Bilgin, T. T., & Çamurcu, Y. (2005). DBSCAN, OPTICS ve K-Means kümeleme algoritmalarının uygulamalı karşılaştırılması. *Politeknik Dergisi*, 8(2), 139–145.
- Cumhurbaşkanlığı Strateji ve Bütçe Başkanlığı. (2025). 2025 yılı genel bütçeli idareler ekonomik kod icmalı ile 2026–2027 gider tahminleri. Ankara, Türkiye.
- Demir, H., & Kurt, M. E. (2017). Yolsuzluk ve katastrofik sağlık harcamaları. *Dicle Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 7(14), 363–384.
- Dikbaş, Ü., & Ebeğil, M. (2025). Makine öğrenmesi ile kentsel kiralık bisiklet verilerinde anomali tespiti: Konya ili örneği. *The Journal of International Scientific Researches*, 10(3), 496–509.
- Efe, A., & Tunçbilek, M. (2023). Yapay zekâ algoritmaları ile dönüşen denetim araçları üzerine bir değerlendirme. *Denetişim*, (27), 72–102.
- Eşidir, K. A. (2025). Makine öğrenimi modelleri ile yetişkin eğitimi analizi: Modellerin karşılaştırmalı performansı. *Elektronik Sosyal Bilimler Dergisi*, 24(2), 946–964.
- Güley, A. Ö., & Bozdemir, E. (2024). Sağlık Bakanlığı döner sermaye işletmeleri Sayıştay denetim raporlarının muhasebesel incelenmesi. *Denetişim*, (30), 172–187.
- Gür, Y. E., Eşidir, K. A., & Şimşek, A. İ. (2024). Sağlık istatistiklerinin veri madenciliği teknikleri ile analizi: Makine öğrenmesi algoritmaları kullanılarak genel sağlık durumunun sınıflandırılması. *Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, 24(6), 1364–1381.
- Karagöz, N., & Yaman, E. (2019). Sağlık kuruluşlarında uygulanan iç denetim ve dış denetimin çalışanlar üzerindeki etkileri. *Uluslararası Sağlık Yönetimi ve Stratejileri Araştırma Dergisi*, 5(2), 165–183.

- Kılınç, S., & Boztosun, D. (2024). Sağlık işletmelerinde medikal muhasebe perspektifinde faturalandırılmayan işlemlere yönelik tahakkuklar. *Denetişim*, (29), 35–46.
- Koçak, A., & Ergün, P. M. A. (2023). Sağlıkta veri kalitesi ve veri madenciliği uygulamaları. *Disiplinlerarası Yenilik Araştırmaları Dergisi*, 3(1), 23–30.
- Korkmaz, E., & Coşkun, G. (2022). Pandemi döneminde sağlık sektöründe uzaktan iç denetim: Nitel bir araştırma. *Hacettepe Sağlık İdaresi Dergisi*, 25(4), 761–774.
- Koyuncugil, A., & Özgülbaş, N. (2010). Veri madenciliği: Tıp ve sağlık hizmetlerinde kullanımı ve uygulamaları. *Bilişim Teknolojileri Dergisi*, 2(2).
- Küçük, A. (2023). Sağlık hizmet ödemelerinde usulsüzlük türleri ve mücadele stratejileri. *Sayıştay Dergisi*, 33(127), 585–607.
- Liu, F. T., Ting, K. M., & Zhou, Z.-H. (2008). Isolation forest. In *Proceedings of the 2008 IEEE International Conference on Data Mining (ICDM)* (pp. 413–422). Pisa, Italy.
- Mandalas, M., & Can, E. N. (2020). Kurumsal risk yönetimi kapsamında risk odaklı iç denetim analizi: Marmara bölgesindeki devlet üniversitelerine yönelik bir araştırma. *Denetişim*, (20), 78–95.
- Nazlıoğlu, B. (2020). Sağlık işletmelerinde muhasebe sistemi perspektifinde medikal muhasebe. *Ankara Hacı Bayram Veli Üniversitesi İktisadi ve İdari Bilimler Fakültesi Dergisi*, 22(2), 533–550.
- Sağlık Bakanlığı Teftiş Kurulu Başkanlığı. (2024). Risk esaslı denetim sistemi koordinatörlüğü yönergesi. Ankara, Türkiye.
- T.C. Resmî Gazete. (2003). 5018 sayılı kamu malî yönetimi ve kontrol kanunu (No. 25326, 24 Aralık 2003).
- Türedi, H., Zor, Ü., & Gürbüz, F. (2015). Risk odaklı iç denetim. *Muhasebe ve Finansman Dergisi*, (66), 1–20.
- Veranyurt, Ü., Deveci, A., Esen, M. F., & Veranyurt, O. (2020). Makine öğrenmesi teknikleriyle hastalık sınıflandırması: Rastgele Orman, K-Nearest Neighbour ve AdaBoost algoritmaları uygulaması. *Uluslararası Sağlık Yönetimi ve Stratejileri Araştırma Dergisi*, 6(2), 275–286.
- Yarat, S., & Orman, Z. (2023). Elektrik güç dağıtımında akıllı sayaç verileri için anomali tespiti ve tahminleme. *Bilgisayar Bilimleri ve Teknolojileri Dergisi*, 3(2), 72–85.