



**T.C.
HARRAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

YÜKSEK LİSANS TEZİ

**SALDIRI TESPİT SİSTEMİ İÇİN VERİ KÜMESİ BİRLEŞTİRME İLE TOPLULUK
MAKİNE ÖĞRENMESİ**

**Muhammed ELHASAN
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI**

**ŞANLIURFA
2026**



**T.C.
HARRAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

YÜKSEK LİSANS TEZİ

**SALDIRI TESPİT SİSTEMİ İÇİN VERİ KÜMESİ BİRLEŞTİRME İLE TOPLULUK
MAKİNE ÖĞRENMESİ**

**Muhammed ELHASAN
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI**

DANIŞMAN: Prof. Dr. İbrahim Berkan AYDİLEK

**ŞANLIURFA
2026**



T.C.
HARRAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
YÜKSEK LİSANS TEZ ONAY SAYFASI

Prof. Dr. İbrahim Berkan AYDİLEK'in danışmanlığında, Muhammed ELHASAN'ın hazırladığı "Ensemble Machine Learning with Dataset Integration for Intrusion Detection System" konulu bu çalışma 22/07/2026 tarihinde aşağıdaki jüri tarafından Harran Üniversitesi Lisansüstü Eğitim Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

☒ Oybirliğiyle / ☐ Oy çokluğu ile

İmza

Danışman	Prof. Dr. İbrahim Berkan AYDİLEK	e-imzalıdır
Üye	Doç. Dr. Sait Ali UYMAZ	e-imzalıdır
Üye	Doç. Dr. Mehmet Emin TENKEKİ	e-imzalıdır

Bu tezin Bilgisayar Mühendisliği Ana Bilim Dalında yapıldığını ve enstitümüz kurallarına göre düzenlendiğini onaylarım.

Dr. Öğr. Üyesi Sabri AKIN
Enstitü Müdürü

Not: Bu tezde kullanılan özgün ve başka kaynaktan yapılan bildirişlerin, çizelge, şekil ve fotoğrafların kaynak gösterilmeden kullanımı 5846 sayılı Fikir ve Sanat Eserleri Kanunundaki hükümlere tabidir.

	T.C. HARRAN ÜNİVERSİTESİ TEZLİ YÜKSEK LİSANS ORJİNALLİK RAPORU VE BEYAN FORMU	Doküman No: FRM-0072
		Revizyon No: 03
		Yayın Tarihi: 10.09.2025
		Revizyon Tarihi: 12.01.2026
		Sayfa No: 1 / 1

ÖĞRENCİNİN	
Öğrenci No	235121003
Adı / Soyadı	Muhammed ELHASAN
Anabilim Dalı/Programı	Bilgisayar Mühendisliği Anabilim Dalı / Tezli Yüksek Lisans
Tez Konusu	Saldırı Tespit Sistemi için Veri Kümesi Birleştirme ile Topluluk Makine Öğrenmesi

FEN BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜNE

Yukarıda başlığı belirtilen Saldırı Tespit Sistemi için Veri Kümesi Birleştirme ile Topluluk Makine Öğrenmesi çalışmamın a) Kapak sayfası, b) Giriş, c) Ana bölümler ve d) Sonuç kısımlarından oluşan toplam 94 sayfalık kısmına ilişkin, 01/09/2026 tarihinde danışmanım tarafından **Turnitin** adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, benzerlik oranı %4'tür. Ayrıca Saldırı Tespit Sistemi için Veri Kümesi Birleştirme ile Topluluk Makine Öğrenmesi hiçbir blok kopyalama içermemektedir.

Uygulanan filtrelemeler:

- 1- Kabul/Onay ve Bildirim sayfaları hariç,
- 2- Kaynakça hariç
- 3- Alıntılar hariç/dâhil
- 4- 5 kelimeden daha az örtüşme içeren metin kısımları hariç

Yükseköğretim Kurulu Bilimsel Araştırma ve Yayın Etiği Yönergesi, Harran Üniversitesi Bilimsel Araştırma ve Yayın Etiği Yönergesini ve Harran Üniversitesi Lisansüstü Eğitim Enstitüsü tez çalışması Orijinallik Raporu alınması ve kullanılması Uygulama Esasları'nı inceledim ve bu Uygulama Esasları'nda belirtilen azami benzerlik oranlarına göre tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim. Etik ihlal tespiti halinde, Enstitü yönetim kurulunca, diplomamın iptal edilmesini kabul ediyorum.

Gereğini saygılarımla arz ederim.

01/09/2026
Muhammed ELHASAN
Hazırlayan Öğrencinin
İmzası


Yukarıda yer alan raporun ve beyanın doğruluğunu onaylarım. 01/09/2026

Prof. Dr. İbrahim Berkan AYDİLEK
Danışmanın Unvanı-Adı-Soyadı
(İmzası)



TEŞEKKÜR

Adım tezinin yazarı olsa da tezin tek başıma yazılamayacağını belirtmek gerekir.Çevremdeki birçok insan sayesinde bu mümkün olmuştur.

Öncelikle Tanrı'ya bu tür bir şey yapmam için bana güç sabır ve aydınlatma verdiği için teşekkür ederim.Ardından bana her aşamada çalışmamı tavsiye eden ve bu süreci büyük ölçüde etkileyen Prof. Dr. İbrahim Berkan AYDİLEK'e minnettarım.

Bu çalışmada anneme koşulsuz desteği ve sevgisi için teşekkür etmek istiyorum.Onun hastalığıyla mücadelesinde gösterdiği cesaret en iyi örneklerden biri.

Babama disiplin kaynağım olduğu için teşekkür ediyorum. Ayrıca bana mutluluk getiren kardeşim için de teşekkür ederim.

Son olarak bu çalışmayı yazarken yanımda olan tüm insanlara teşekkür ederim.

Zorluklarla mücadele eden insanları hatırlıyoruz-Filistinliler-Onlar sıkıntılarla karşılaştıklarında ne kadar güçlü olabileceğimizin gerçek bir örneğini veriyorlar.

İÇİNDEKİLER

ÖZET	i
ABSTRACT	ii
ÇİZELGELER DİZİNİ	iii
ŞEKİLLER DİZİNİ	iv
KISALTMALAR	v
1. GİRİŞ	1
1.1. Konunun Önemi	2
1.2. Çalışmanın Katkısı	2
1.3. Özgün Katkı	3
1.4. Araştırma Amaçları	3
1.5. Araştırma Soruları	4
1.6. Tez Yapısı	5
2. ÖNCEKİ ÇALIŞMALAR	6
2.1. Siber Güvenlikte YZ'nin Kavramsal Temelleri	6
2.2. Saldırı Tespitinde Makine Öğrenmesinin Evrimi.....	8
2.3. Modern Siber Tehdit Ortamı	11
2.3.1. Dijital Çağda Siber Suçun Evrimi.....	12
2.3.2. Geleneksel Saldırı Tespit Sistemlerinin (IDS) Sınırlılıkları.....	12
2.3.3. İleri Düzey ve YZ Destekli Saldırı Vektörleri.....	12
2.4. Önceki Çalışmaların Karşılaştırmalı Analizi	16
2.5. Araştırma Boşluğu ve Motivasyon	19
3. GEREÇ ve YÖNTEM	21
3.1. Araştırma Metodolojisine Genel Bakış	21
3.1.1. Hibrit Savunma Yaklaşımının Kavramsal Çerçevesi.....	22
3.2. Siber Güvenlikte Topluluk Öğrenmesi Kullanımının Gerekçesi.....	24
3.2.1. Derin Öğrenmenin Hariç Tutulmasının Gerekçesi.....	25
3.3. Veri Kümeleri ve Veri Kaynakları.....	26
3.3.1. Veri Kümeleri: Yapı ve İlgililik.....	26
3.3.2. Veri Bütünlüğü, Güvenilirlik ve Kıyaslama Gerekçesi	27
3.4. Veri Ön İşleme Hattı	27
3.4.1. Veri Temizleme ve Gürültü Azaltma.....	28
3.4.2. Özellik Mühendisliği ve Boyut İndirgeme	28
3.4.3. Sınıf Dengesizliğinin Ele Alınması.....	29
3.4.4. Veri Bölme Stratejisi (Eğitim, Doğrulama, Test).....	30
3.5. Makine Öğrenmesi Sınıflandırıcıları	30
3.5.1. Ağaç Tabanlı ve Klasik Modeller	31
3.5.2. Topluluk ve Boosting Tabanlı Modeller	32
3.5.3. Olasılıksal ve Mesafe Tabanlı Modeller (Temel Modeller).....	33
3.5.4. Model Çeşitliliği ve Topluluk Gürbüzlüğü.....	34
3.5.5. Topluluk Mimarisi Tasarımının Yönlendirilmesi.....	34
3.6. Topluluk Tasarım Gerekçesi ve Stratejisi	34
3.6.1. Bagging Tabanlı Öğrenme Stratejisi	35
3.6.2. Boosting Tabanlı Öğrenme Stratejisi	35
3.6.3. Topluluk Çeşitliliği ve Tasarım Gerekçesi	35

3.6.4. Topluluk Çerçevesine Kavramsal Genel Bakış	36
3.7. Deneysel Kurulum.....	36
3.7.1. Donanım ve Yazılım Ortamı	36
3.7.2. Veri Kümesi Yapılandırması	37
3.7.3. Eğitim Yapılandırması.....	37
3.8. Değerlendirme Ölçütleri.....	40
3.8.1. Accuracy (Doğruluk).....	40
3.8.2. Precision (Saldırı Sınıfı).....	41
3.8.3. Recall (Saldırı Sınıfı)	41
3.8.4. F1-Score	41
3.8.5. ROC-AUC	41
4. BULGULAR	43
4.1. Nicel Sonuçlar	43
4.1.1. En Yüksek Performanslı Modeller: Bagging, Decision Tree ve KNN.....	44
4.1.2. Boosting Modelleri: Gradient Boosting, XGBoost ve AdaBoost.....	44
4.1.3. Bagging Ailesi Modelleri: Random Forest ve Extra Trees	45
4.1.4. Doğrusal ve Olasılıksal Temel Modeller	45
4.2. Görsel Performans Analizi	45
4.2.1. Hata Matrisi Analizi	45
4.2.2. ROC Eğrisi Analizi	49
4.2.3. Karşılaştırmalı Performans Grafiği	49
4.3. Veri Kümesi Bazlı Genellenebilirlik Analizi	53
4.4. Bagging Temel Öğrenici Ablasyonu	55
4.5. Hesaplama Maliyeti Analizi	56
4.6. Rastgele Tohumlar Arasında Kararlılık	58
4.7. Bulguların Değerlendirilmesi	60
4.7.1. Bu Görevde Bagging Neden En İyi Performansı Gösteriyor?.....	60
4.7.2. IDS'te Precision-Recall Ödünleşimi	61
4.7.3. Temel Model Davranışı	62
4.7.4. Değerlendirmenin Sınırlılıkları	62
4.7.5. Özellik Atfı Analizi	62
4.8. Dağıtım Prototipi	63
4.9. Bölüm Özeti	66
5. TARTIŞMA.....	67
5.1. Katkıların Özeti	67
5.1.1. Araştırma Problemi	67
5.1.2. Metodolojik Katkılar	67
5.2. Temel Ampirik Bulgular	68
5.2.1. Ağaç Tabanlı Modellerin Üstünlüğü	68
5.2.2. Varsayılan Seçim Olarak Bagging.....	68
5.2.3. Beş Özellikli Gösterim Yeterlidir	68
5.2.4. Temel Model Başarısızlıkları Bilgilendiricidir	69
5.2.5. Veri Kümesi Bazlı Genelleme	69
5.3. Çalışmanın Sınırlılıkları	70
5.3.1. Veri Sınırlılıkları.....	70

5.3.2. Özellik Sınırlılıkları.....	70
5.3.3. Değerlendirme Sınırlılıkları	71
5.3.4. Uygulama Sınırlılıkları.....	71
5.3.5. Kapsam Sınırlılıkları	72
6. SONUÇLAR	73
6.1. Bu Tezin Başardıkları	73
6.2. Temel İçgörü: Bagging Neden XGBoost’tan Daha İyi Performans Gösterir?	73
6.3. Uygulamaya Yönelik Çıkarımlar	74
6.4. Çözümlememiş Kalanlar	74
6.5. İleriye Giden Yol	75
6.6. Son Söz.....	75
7. ÖNERİLER	76
KAYNAKLAR.....	78
ÖZ GEÇMİŞ	82

ÖZET

YÜKSEK LİSANS TEZİ

SALDIRI TESPİT SİSTEMİ İÇİN VERİ KÜMESİ BİRLEŞTİRME İLE TOPLULUK MAKİNE ÖĞRENMESİ

Muhammed ELHASAN

HARRAN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

Tez Danışmanı: Prof. Dr. İbrahim Berkan AYDİLEK

Yıl: 2026, Sayfa: 82

Saldırı Tespit Sistemlerini (Intrusion Detection Systems – IDS) destekleyen makine öğrenmesi algoritmaları; yüksek boyutlu veri, ciddi sınıf dengesizliği ve tek veri kümesine dayalı değerlendirmelerden kaynaklanan sınırlı genellenebilirlik gibi temel zorluklarla karşı karşıyadır. Bu tez, söz konusu zorlukları ele almak amacıyla topluluk öğrenmesine dayalı, çoklu veri kümesi kullanan bir saldırı tespit çerçevesi önermektedir. Bu kapsamda, yaygın olarak kullanılan üç IDS veri kümesi olan CICIDS2017, CICIDS2018 ve UNSW-NB15 ön işleme tabi tutulmuş, saldırı etiketleri birleştirilmiş ve %89,1 normal trafik ile %10,9 saldırı trafiği içeren 10.799.503 ağ akışından oluşan tek bir ikili sınıflandırma veri kümesinde birleştirilmiştir. Veri kümeleri arasında adil bir karşılaştırma yapılabilmesi ve veri kümesine özgü özelliklerin etkisinin azaltılması amacıyla yalnızca tüm veri kümelerinde ortak bulunan beş akış düzeyli özellik kullanılmıştır: Süre (Duration), Kaynak Bayt Sayısı (Source Bytes), Hedef Bayt Sayısı (Destination Bytes), Kaynak Paket Sayısı (Source Packets) ve Hedef Paket Sayısı (Destination Packets). Toplam on bir sınıflandırıcı aynı deneysel koşullar altında üç grupta değerlendirilmiştir: temel modeller (Gaussian Naive Bayes, K-En Yakın Komşu, Doğrusal SVM, Karar Ağacı ve Lojistik Regresyon), bagging tabanlı topluluk modelleri (Random Forest, Bagging ve Extra Trees) ve boosting tabanlı topluluk modelleri (AdaBoost, Gradient Boosting ve XGBoost). Tüm modeller %80/%20 tabakalı eğitim-test bölünmesiyle değerlendirilmiştir. Sonuçlar, ağaç tabanlı ve topluluk öğrenmesi yöntemlerinin doğrusal ve olasılıksal temel modellere göre belirgin biçimde daha başarılı olduğunu göstermiştir. Bagging algoritması %99,21 doğruluk, %96,38 F1-skoru ve %99,91 ROC-AUC değeri ile en yüksek genel performansı elde etmiştir. En yüksek kesinlik değeri (%97,68) XGBoost tarafından, en yüksek duyarlılık değeri (%98,65) ise Random Forest tarafından elde edilmiştir. Bununla birlikte, tek bir Karar Ağacı ve K-En Yakın Komşu modeli de en başarılı topluluk yöntemine oldukça yakın sonuçlar vermiştir. Bu durum, seçilen beş özellikli gösterimde topluluk öğrenmesinin sağladığı performans artışının sınırlı olduğunu göstermektedir. Veri kümesi bazlı değerlendirmeler, temel öğrenici ablasyon çalışmaları, hesaplama maliyeti analizleri ve beş farklı rastgele tohum üzerinde gerçekleştirilen kararlılık analizleri, Bagging yönteminin başarısında varyans azaltımının temel etken olduğunu göstermektedir. Bu çalışma, üç bağımsız veri kümesinden elde edilen küçük ancak ortak bir özellik kümesi kullanılarak etkili saldırı tespiti yapılabileceğini göstermekte ve araştırmacılar için yeniden üretilebilir çoklu veri kümesi tabanlı bir IDS çerçevesi sunmaktadır.

ANAHTAR KELİMELER: Makine öğrenmesi, Saldırı Tespit Sistemi, Topluluk Öğrenmesi, Siber Güvenlik, Ağ Trafiği Analizi

ABSTRACT

MASTER’S THESIS

ENSEMBLE MACHINE LEARNING WITH DATASET INTEGRATION FOR INTRUSION DETECTION SYSTEM

Muhammed ELHASAN

**HARRAN UNIVERSITY
GRADUATE SCHOOL
DEPARTMENT OF COMPUTER ENGINEERING**

Thesis Advisor: Prof. İbrahim Berkan AYDİLEK

Year: 2026, Page: 82

Intrusion Detection Systems (IDS) represent an important aspect of modern network security. The machine learning algorithms that support these systems must cope with three persistent problems: high-dimensional data, significant class imbalance, and over-reliance on single-dataset testing, which leads to overly optimistic accuracy measures. This thesis aims to solve these problems through an ensemble-based, multi-dataset IDS framework intended to enhance robustness and generality. Three popular benchmark IDS datasets—CICIDS2017, CICIDS2018, and UNSW-NB15—were preprocessed, label-unified, and combined into a single binary-classification corpus containing 10,799,503 network flows (89.1% benign and 10.9% attack). To perform a fair cross-dataset analysis and make the results insensitive to dataset-specific properties, the study was limited to five flow-level features common to all datasets: Duration, Source Bytes, Destination Bytes, Source Packets, and Destination Packets. Eleven classifiers were tested under identical conditions in three groups: baseline models (Gaussian Naive Bayes, K-Nearest Neighbours, Linear SVM, Decision Tree, and Logistic Regression), bagging ensembles (Random Forest, Bagging, and Extra Trees), and boosting ensembles (AdaBoost, Gradient Boosting, and XGBoost), using an 80/20 stratified train/test split with no separate validation set. Tree-based and ensemble approaches clearly outperformed the linear and probabilistic baselines. The Bagging algorithm achieved the best overall performance (99.21% accuracy, 96.38% F1-score, and 99.91% ROC-AUC), while the highest precision was obtained by XGBoost (97.68%) and the highest recall by Random Forest (98.65%). Notably, a single Decision Tree and KNN reached performance close to the best ensemble, indicating that the benefit of ensembling is minimal under such a small feature set. Per-dataset evaluation, base-estimator ablation, computational-cost measurement, and a stability analysis over five random seeds together suggest that variance reduction is the main reason for Bagging's effectiveness. This study demonstrates that robust intrusion detection can be performed using a small, universal feature set drawn from three independent datasets, and it provides researchers with a reproducible multi-dataset IDS framework.

KEYWORDS: Machine Learning, Intrusion Detection System, Ensemble Learning, Cybersecurity, Network Traffic Analysis

ÇİZELGELER DİZİNİ

Çizelge 2.1.	Geleneksel Siber Saldırıları ile YZ Destekli Saldırı Vektörlerinin Karşılaştırması	15
Çizelge 2.2.	CICIDS2018 Veri Kümesi Üzerine Önceki Çalışmaların Karşılaştırmalı Sonuçları	17
Çizelge 3.1.	On bir sınıflandırıcının hiperparametre yapılandırması; bu çalışmada açıkça belirlenen değerler ile kütüphane varsayılanları ayrı gösterilmiştir.	39
Çizelge 4.1.	On bir modelin birleştirilmiş veri kümesindeki performansı (2.159.901 test akışı), F1-score'a göre sıralanmıştır.	44
Çizelge 4.2.	Her modelin her bir veri kümesindeki F1-score'u (birleştirilmiş sütun referans için tekrarlanmıştır).	53
Çizelge 4.3.	Her modelin her bir veri kümesindeki accuracy'si (birleştirilmiş sütun referans için tekrarlanmıştır).	55
Çizelge 4.4.	Veri kümesi başına en iyi model (F1-score'a göre) ile birleştirilmiş derlemin karşılaştırması.	55
Çizelge 4.5.	Temel öğrencinin Bagging topluluğu üzerindeki etkisi (birleştirilmiş veri kümesi).	56
Çizelge 4.6.	Birleştirilmiş veri kümesinde her modelin eğitim ve çıkarım (duvar saati) süresi.	57
Çizelge 4.7.	Beş rastgele tohum arasında temel ölçütlerin ortalaması (+)(-) standart sapması (yüzde puan cinsinden).	59

ŞEKİLLER DİZİNİ

Şekil 2.1.	Saldırı Tespit Sistemlerinin İmza Tabanlıdan Topluluk Tabanlı Yaklaşımlara Kavramsal Evrimi	11
Şekil 3.1.	Önerilen Topluluk Tabanlı Çok Veri Kümeli Saldırı Tespit Çerçevesinin Genel İş Akışı (seçilen model: Bagging; birleştirilmiş derlem: 10,8M akış; bölünme: 80/20).	23
Şekil 3.2.	Eğitim ve Test için Kullanılan Tabakalı Eğitim/Test Bölünmesi (80/20).	30
Şekil 4.1.	Çok veri kümeli test kümesinde Bagging sınıflandırıcısının hata matrisi.	46
Şekil 4.2.	Çok veri kümeli test kümesinde XGBoost sınıflandırıcısının hata matrisi.	47
Şekil 4.3.	Çok veri kümeli test kümesinde tüm on bir sınıflandırıcısının hata matrisleri.	48
Şekil 4.4.	Birleştirilmiş veri kümesinde modele göre ROC-AUC.	49
Şekil 4.5.	En iyi beş sınıflandırıcı arasında accuracy, precision, recall ve F1-score karşılaştırmalı çubuk grafiği.	50
Şekil 4.6.	Tüm ölçütler arasında model performansı ısı haritası (birleştirilmiş veri kümesi).	51
Şekil 4.7.	En iyi altı modelin radar grafiği.	52
Şekil 4.8.	Üç kaynak veri kümesi arasında tüm on bir model için veri kümesi bazlı F1-score ısı haritası.	54
Şekil 4.9.	Birleştirilmiş veri kümesinde model başına eğitim ve çıkarım (duvar saati) süresi (logaritmik ölçek).	58
Şekil 4.10.	Beş rastgele tohum arasında F1-score'un kararlılığı (ortalama $\pm$ standart sapma).	60
Şekil 4.11.	CICIDS2017, CICIDS2018 ve UNSW-NB15 Veri Kümelerini Kullanan Topluluk Tabanlı Saldırı Tespiti için Etkileşimli Streamlit Prototipi.	64
Şekil 4.12.	Önerilen Streamlit Tabanlı IDS Prototipinde Topluluk Tahmin Sonuçları ve Bireysel Sınıflandırıcı Çıktıları.	64
Şekil 4.13.	Streamlit prototipi çıkarım hattı: CSV yükleme ve ön işlemeden on bir modelli çoğunluk oyu Benign/Attack kararına.	65

KISALTMALAR

AI	Yapay Zekâ (Artificial Intelligence)
APT	Gelişmiş Kalıcı Tehdit (Advanced Persistent Threat)
AUC	Eğri Altında Kalan Alan (Area Under Curve)
CHEM	Uzlaşımli Hibrit Topluluk Modeli (Consensus Hybrid Ensemble Model)
FN	False Negative (Yanlış Negatif)
FP	False Positive (Yanlış Pozitif)
TN	Gerçek Negatif (True Negative)
TP	Gerçek Pozitif (True Positive)
CNN	Evrişimli Sinir Ağı (Convolutional Neural Network)
GAN	Çekişmeli Üretici Ağ (Generative Adversarial Network)
GDPR	Genel Veri Koruma Tüzüğü (General Data Protection Regulation)
IDS	Saldırı Tespit Sistemi (Intrusion Detection System)
IPS	Saldırı Önleme Sistemi (Intrusion Prevention System)
KNN	k-En Yakın Komşu (k-Nearest Neighbours)
LIME	Yerel Yorumlanabilir Model-Bağımsız Açıklamalar (Local Interpretable Model-agnostic Explanations)
LSTM	Uzun Kısa Süreli Bellek (Long Short-Term Memory)
ML	Makine Öğrenmesi (Machine Learning)
RNN	Yinelemeli Sinir Ağı (Recurrent Neural Network)
ROC	Alıcı İşletim Karakteristiği (Receiver Operating Characteristic)
SHAP	SHapley Toplamsal Açıklamalar (SHapley Additive exPlanations)
SIEM	Güvenlik Bilgi ve Olay Yönetimi (Security Information and Event Management)
SVM	Destek Vektör Makineleri (Support Vector Machines)
XAI	Açıklanabilir Yapay Zekâ (Explainable Artificial Intelligence)
XGBoost	Aşırı Eğim Artırma Modeli (Extreme Gradient Boosting)
SMOTE	Sentetik Azınlık Aşırı Örnekleme Tekniği (Synthetic Minority Oversampling Technique)
YZ	Yapay Zekâ

1. GİRİŞ

Hızla gelişen dijital çağda siber güvenlik herkes için önemli bir konu haline geldi, birey ya da devlet kurumları olsun. Artan sayıda siber saldırı ile birlikte geleneksel güvenlik mekanizmalarının etkisiz olduğu açıkça görülüyor. Çünkü geleneksel sistemler kurallara ve imzalara dayalı çalışıyorlar ve bu sistemler dinamik siber saldırıları karşılayamıyor (Kanimozhi & Jacob, 2021).

Yeni teknolojik gelişmelerin ve uygulamalarının yanı sıra İnternet'in ortaya çıkması siber suçların artmasına ve çeşitlenmesine neden oluyor. Siber suçlar bu suçların sanal ortamda gerçekleşmesi ve bilgisayar sistemlerine sızma ve kullanıcıların verilerini çalmak gibi eylemleri içerebileceği yönünden benzersizdir. Teknolojilerin hızlı gelişimi, hedeflerin artması ve siber suçlarda sosyal mühendislik kullanımının (Sibai, 2024) artması bu suçların uluslararası niteliklerini giderek daha çok zorlaştırmaktadır.

Ayrıca yapay zekanın (AI) ve makine öğreniminin (ML) yükselişi proaktif güvenlik önlemlerinin yaratılmasına olanak sağlamaktadır (Russell & Norvig, 2021). Gerçekten de AI teknolojilerine dayalı güvenlik araçları büyük hacimli bilgileri analiz edebilir, sapmaları tespit edebilir ve alışılmadık yöntemlerle güvenlik kontrolünde görülemeyen olası tehditleri keşfedebilir. Bu güvenlik kontrol yöntemlerine tepki yerine tahmin yöntemine dayalı bir değişim getirecek ve güvenlik tehditlerine karşı tepki verilme etkinliğini artırabilecektir.

Bununla ilgili araştırmalara olan ilgi artmasına rağmen yapay zekanın mevcut teorileri ve siber güvenlik uygulaması arasında belirgin bir kopukluk var (Alkahtani vd., 2025). Şu anda yapılan araştırmalar belirli bir tür saldırıya odaklanıyor ve siber güvenlik alanının gerektirdiği esnekliği göz ardı ediyorlar.

Bu tezde, çeşitli siber tehditlerin tespit edilmesi amacıyla Karma Savunma Modeli kullanılarak verimli bir tespit yöntemi sunulmaktadır. Çalışmada en iyi sonuçların elde edilebilmesi için topluluk öğrenmesi ilkelerinin kullanılması önerilmektedir (Kilincer vd., 2022; Khan vd., 2025). Modelin verimliliği, farklı veri kümeleri, makine öğrenmesi algoritmaları ve değerlendirme ölçütleri kullanılarak test edilmiştir.

1.1. Konunun Önemi

Günümüz dünyasının birbirine bağlı olmasından dolayı dijital altyapı, günümüz toplumunun ve ekonomisinin temelini oluşturmaktadır. Bu bağımlılığın artmasıyla birlikte saldırı yüzeyi de genişlemektedir ve insanları, kurumları hatta ülkeleri fidye yazılım saldırıları, büyük veri ihlalleri, gelişmiş kalıcı tehdit (APT) ve sıfır gün saldırıları gibi farklı türde siber saldırılara açık hale getiriyor.

Bu tehditler alışlagelmiş olanlardan yeni ve dinamik olduğu için geleneksel kural tabanlı güvenlik duvarları ve imza tabanlı istilacı tespit sistemleri (IDS) bu yeni tehdidi takip edemiyor çünkü her ikisi de statik bir yapıya sahip ve bu nedenle bu tür uyarlanabilir tehditleri tespit etmek için verimsizler ve bu da daha fazla yanlış alarmlara yol açıyor.

Yapay Zeka (AI) topluluğunun öğrenen modeli alanındaki yeni bir yaklaşım bu sorunu çözmeye yardımcı olabilir.

Bu ciddi sonuçlara yol açabilecek artan tehdit nedeniyle akıllı bir siber savunma sistemine ihtiyaç var. Bunun nedeni siber saldırı kampanyalarının artan karmaşıklığı ve hızı etkili ve verimli bir topluluk temelli siber saldırı sistemi kaçınılmaz hale getiriyor.

1.2. Çalışmanın Katkısı

Bu çalışmanın Yapay Zeka temelli Siber Güvenlik alanına katkısı farklı veri setlerinin kullanımı yoluyla daha genel ve sağlam hale getirecek yeni bir topluluk temelli saldırı algılama yaklaşımı önermesidir. Çoğu çalışmanın tek bir veri kümesi kullanarak çeşitli algoritmaları karşılaştırdığına aksine yazarlar üç karşılaştırma veri kümesi (CICIDS2017, CICIDS2018 ve UNSW-NB15) kullanarak aynı deneyleri gerçekleştirdiler.

Deneyin ana aşamalarından biri farklı yapıları olan veri setlerini eşit şekilde işleyebilecek ön işleme boru hattının oluşturulmasıdır. Bu süreç veri setlerini ikili sınıflandırma görevi haline getirir ve tüm veri setleri için ortak olan en küçük sayıda özelliğin seçilmesini içerir.

Bu şekilde tüm veri setleri aynı sınıflandırma probleminde dönüştürülür ve herhangi bir algoritma aynı verilerden öğrenebilir ve belirli veri kümesine özgü özelliklerin etkisi

olmadan.

Ayrıca çalışma on bir makine öğrenimi modelinin karşılaştırmalı analizini sunmaktadır. Hem klasik algoritmalar hem de topluluklarının karşılık gelen algoritmaları test edilmiştir. Deney sonuçlarına göre topluluk modelleri klasik algoritmalarından daha başarılı olmuştur. Bu nedenle çalışma topluluk temelli IDS'lerin geliştirilmesine katkıda bulunmuştur.

1.3. Özgün Katkı

Bu tezin benzersizliği çeşitli klasik makine öğrenimi algoritmalarının uygulanması, test edilmesi ve analiz edilmesinden oluşmaktadır. Bu algoritmalar arasında Rastgele Orman, Karar Ağacı, Logistik Regresyon, K En Yakın Komşu, Doğrusal Destek Vektör Makinesi ve Naive Bayes ile birlikte XGBoost, AdaBoost, Bagging, Gradient Boosting ve Extra Trees gibi yenilikçi birleşik öğrenme algoritmaları bulunmaktadır. Tüm bu algoritmalar aynı veri kümesinde aynı ön işleme koşulları altında eğitildi, kaydedildi ve test edildi.

Bu çalışma başka herhangi bir araştırmadan bağımsızdır ve bu algoritmaların karşılaştırılması için bir çerçeve sunmaktadır. Bu çalışmada kullanılan performans ölçütleri Doğruluk, kesinlik, duyarlılık, F1-skoru ve ROC-AUC'dir. Kod tabanında joblib kütüphanesi kullanılmıştır ki bu da sonuçları yeniden üretilebilir ve kolay uygulanabilir hale getirir. Tüm bu algoritmaları uygulayarak bu algoritmalar arasındaki performans farkı hakkında pratik bilgiler elde edebilir ve bu hızlı klasik makine öğrenmesi algoritmaları kullanılarak saldırı tespiti için en uygun algoritma belirlenebilmektedir. Derin öğrenme algoritmalarından çok daha fazla zaman alan bu algoritmalar yerine. Ayrıca Bagging ile ilgili önemli bir çalışma bulunmaktadır.

1.4. Araştırma Amaçları

Bu tezin temel amacı; genellik, ölçeklenebilirlik ve uygulanabilirliğe odaklanarak modern siber saldırıları tespit etmek için etkili, çok veri kümeli, topluluk tabanlı bir makine öğrenmesi sistemi geliştirmektir. Özel amaçlar şunlardır:

- Geleneksel IDS'lerin ve tekil makine öğrenmesi algoritmalarının, siber saldırı bilgisi içeren büyük ve çeşitli veri kümeleriyle başa çıkmadaki eksikliklerini belirlemek.
- Üç popüler kıyaslama veritabanını (CICIDS2017, CICIDS2018 ve UNSW-NB15) birleşik

bir ikili sınıflandırma problemine dönüştürerek ortak bir saldırı tespit platformu inşa etmek.

- Veri kümeleri arasındaki yapısal farklılıkları giderecek biçimde veri temizleme, etiket standartlaştırma ve özellik birleştirme işlemlerini gerçekleştiren eksiksiz bir ön işleme şeması oluşturmak.
- Algoritmik performanstan ödün vermeden adil bir çapraz veri kümesi değerlendirmesini kolaylaştırmak için tüm veri kümeleri için asgari bir özellik kümesi tanımlamak ve kullanmak.
- Geleneksel sınıflandırıcılar ve topluluk tabanlı yöntemler (Bagging, Random Forest, Extra Trees, Gradient Boosting, AdaBoost ve XGBoost) dâhil olmak üzere on bir makine öğrenmesi algoritmasının performansını incelemek.
- Algoritmaları Accuracy, Precision, Recall, F1-Score ve ROC-AUC temelinde, dengesiz bir veri kümesinde saldırıları tespit etme yeteneklerine odaklanarak değerlendirmek.
- Her sınıflandırıcıyı tek tek her kaynak veri kümesi üzerinde eğitip test ederek bir genellenebilirlik analizi ve Bagging topluluğu içindeki temel öğreniciye yönelik bir ablasyon çalışması yürütmek.
- Geliştirilen modelin pratik yönlerini; ölçeklenebilirlik, dağıtılabilirlik ve gerçek saldırı tespit senaryolarındaki uygulanabilirlik açısından incelemek.

1.5. Araştırma Soruları

1. Birleştirilmiş üç karşılaştırmalı veri kümesinden elde edilen ve en az ortak özellik kümesine indirgenen bir derlenmiş veri kümesi, çeşitli saldırılarını tespit etmek için bireysel sınıflandırıcılara göre üstün mü?
2. Bagging tabanlı ve boosting dayalı topluluklar arasında hangisi (algılama performansı (F1-skoru) ve hesaplama maliyeti (eğitim ve çıkarma süresi) arasında) daha iyi bir denge sağlamaktadır?
3. Tek bir sınıflandırıcının derlenmiş veri kümesi üzerinde eğitilmesi her bir kaynak veri kümesi üzerinde ayrı ayrı eğitilen bir sınıflandırıcıya göre daha iyi sonuçlar verir mi?
4. Bagging tabanlı toplulukların varyans azaltma ve/veya önyargı azaltma mekanizmalarının verimliliğini ne sağlar?

5. Ölçülen performansın ne kadarı sınıf oranından, ne kadarı sınıflandırıcı kapasitesinden kaynaklanmaktadır?

1.6. Tez Yapısı

- Bölüm 1: Giriş. Tezin arka planına, amaçlarına, katkısına ve yapısına genel bir bakış sunar.
- Bölüm 2: Önceki Çalışmalar. Geleneksel IDS sistemlerine, makine öğrenmesinin siber güvenlikteki rolüne, topluluk yöntemlerine ve son gelişmelere genel bir bakış sunar.
- Bölüm 3: Materyaller ve Yöntemler. Kullanılan veri kümeleri, ön işleme, sınıflandırıcılar, topluluk yöntemleri ve değerlendirme ölçütlerine ilişkin ayrıntıları sağlar.
- Bölüm 4: Bulgular. Farklı sınıflandırıcıların performansını karşılaştırarak çalışmanın bulgularını sunar.
- Bölüm 5: Tartışma. Temel ampirik bulguları, metodolojik katkıları ve çalışmanın sınırlılıklarını yorumlar.
- Bölüm 6: Sonuç. Tezin başardıklarını ve Bagging'in üstünlüğünün esasen varyans azaltımından kaynaklandığına dair temel içgörüyü özetler.
- Bölüm 7: Öneriler. Uygulayıcılar ve araştırmacılar için pratik önerilerin yanı sıra gelecekteki çalışmalara yönelik yönergeler sunar.

2. ÖNCEKİ ÇALIŞMALAR

Bu bölüm yapay zekanın siber güvenlikte kullanımı konusunda daha önce yapılan araştırmaları ayrıntılı bir şekilde sentezlemektedir özellikle topluluk öğrenme algoritmalarının IDS'ye (Russell & Norvig, 2021) uygulanması üzerine. Yapay zekanın çeşitli araştırma alanlarında büyük ilerlemeler kaydetmesine rağmen son araştırmalar tek algoritmali modellerin modern siber tehditlerin artan karmaşıklığıyla başa çıkarken sorunlar yaşadığını kanıtlıyor (Sharafaldin vd., 2018) .Bu yüzden araştırmacılar arasında doğruluğu iyileştirmek, yanlış alarm sayısını azaltmak vesıfır gün saldırılarına karşı direnci artırmak için hibrit modeller daha yaygın hale geldi (Khan vd., 2025) (Alkahtani vd., 2025) .

İnceleme süreci konuyla ilgili teorik arka plana başlıyor; yapay zekanın temelleri (Russell & Norvig, 2021) ve yapay zekanın siber güvenlik sistemlerinin gelişimi tarihsel evrimi.Sonrasında tehditleri tespit etme,tanımlama ve bunlara yanıt verme konusunda makine öğrenimi ve derin öğrenme teknolojilerinin uygulamalarına ve uygulamalarına odaklanıyor.(Winston, 1992) tarafından belirtildiği gibi CHEM gibi hibrit topluluk öğrenme algoritmaları ve yığınlama tabanlı yöntemler diğer türlerdeki IDS'lere göre iyi sonuçlar göstermiştir.

Bu çalışmaların güçlü ve zayıf yönlerini analiz ederek bu bölüm ölçeklenebilirlik,açıklanabilirlik (Khan vd., 2025) ve gerçek zamanlı uyumluluk (Kanimozhi & Jacob, 2019a) gibi alanlarda mevcut yaklaşımlardaki boşlukları belirleyecektir.Bu şekilde önerilen hibrit siber güvenlik sistemi oluşturulacaktır.

2.1. Siber Güvenlikte YZ'nin Kavramsal Temelleri

Yapay zekanın (YZ) temel hedeflerinden biri insanların yeteneklerini geliştirmek ve karmaşık karar alma süreçlerini desteklemektir. Bununla birlikte diğer teknolojik yenilikler iş yükünü azaltırken YZ tüm insanların hizmetinde olan bağlantılı sistemlerin yönetimini geliştirmeye çalışmaktadır.YZ şimdi insanların hayatını değiştirebilecek ve eşitsizliği çözebilecek ve aynı zamanda acıları azaltabilecek devrim niteliğinde bir icat olarak görülüyor.YZ'nin uygulamaları Google, Netflix, Facebook gibi popüler web sitelerinde görülebilir. Bu da günümüz dünyası için YZ'den büyük bir fayda sağladığını göstermektedir

(Winston, 1992) .

Zeka genellikle çevreyi algılama yeteneği, bu hakkında mantıklı sonuçlar çıkarma ve hedeflere ulaşmak için rasyonel davranış sergileme yeteneği olarak tanımlanır. Makineler bilinç ve ne olduğunu anlayamayabilir fakat yapay zeka bilimi bilinçli ve ne olduğunu anlayabilen akıllı ajanlar yaratmaktadır. Bu ajanlar bile öğrenebilir, uyum sağlayabilir ve hatta belirsiz koşullarda bile kararlar alabilir. Siber güvenlikte akıllı ajanlar ağları kötü amaçlı eylemlerden korumak için önemlidir (Russell & Norvig, 2021) .

İnsan zekasına sahip makinelerin yeteneği 20. yüzyılın ortalarına doğru bilim insanları için bir nesne haline geldi. John McCarthy ve Alan Turing YZ kavramının kurucuları olarak kabul edilir. O zamandan beri YZ'yi ilerletmek için çok sayıda araştırma yapıldı. Doğal Dil İşleme 1970'lerde derin öğrenme ve yapay sinir ağları gibi kavramlarla birlikte YZ'nin ayrı bir alanı haline geldi. YZ teknolojilerinin gelişimi insan toplumunu farklı şekillerde etkilemektedir ve yeni fırsatlar doğmaktadır (Russell & Norvig, 2021) .

Son araştırmalar özellikle COVID-19 pandemisi gibi bir kriz döneminde YZ'nin siber güvenliğin artan rolünü göstermiştir. YZ verileri analiz etme ve hükümetlere uygun yanıtlar konusunda bilgi sağlama konusunda verimliliğini kanıtlamıştır. Ancak bu sistemler hackerlar tarafından hedef haline gelmiştir ve tahmin modellerini çarpıtabilir ve zararlı kararlar alabilirler. Ayrıca tıp alanında YZ destekli robotlar yeni adli sorunlar yaratmıştır ve bu sorunlar için özel dijital soruşturma yaklaşımları gerekmektedir. YZ'nin yüz ve ses tanıma gibi biyometrik güvence alanlarında kullanımı özellikle Avrupa'da temas takibi uygulamaları nedeniyle büyük gizlilik endişeleri doğurmuştur. Bu çalışmalar teknolojik gelişmelerin ulusal siber güvenlik stratejilerine nasıl entegre edileceğini ve bu entegrasyonun yasal ve etik yönlerini göz önünde bulundurmanın ne kadar önemli olduğunu göstermektedir (Feng vd., 2020).

Siber güvenlik, bağlantılı yazılım ve bilgiyi tüm türden zararlardan, yetkisiz erişilerden veya siber saldırılardan korumak için alınan önlemlerin bir setidir. Bilişim Teknolojileri sürekli değişirken yeni tehditler ortaya çıkmaktadır. YZ'nin artan gücü oyun, üretim, sağlık hizmeti, eğitim ve doğal dil işleme gibi alanlarda yenilikler ve farklı türde otomasyonlar yaratmaya olanak sağlamaktadır (Russell & Norvig, 2021).

2.2. Saldırı Tespitinde Makine Öğrenmesinin Evrimi

Bilgi teknolojilerindeki gelişmelerle birlikte siber suçlar geleneksel suçlardan farklılaşmaktadır. Suçun belli bir türü değil kullanılan teknoloji suçun farklılaşmasında rol oynamaktadır. Geleneksel suç biçimleri siber alana taşınırken kötü amaçlı yazılımlar, botnetler ve ağ ihlalleri gibi yöntemler ve araçlar geleneksel suç yöntemlerini ve araçlarını yer değiştirmiştir. Ancak suçun arkasındaki suçlu niyeti değiştirilmemiştir. Siber suçların özünü oluşturan şeyin transnasyonel, hızlı ve görünmez doğasıdır. Suçlular uzak bir yerden suç işleyebilir ve suçluların anonimliği fiziksel dünyada olduğundan daha da suçluları tespit etmeyi zorlaştırmaktadır. Suç işlemek için kullanılan araçlardaki bu değişim siber suçları dünya çapında kolay ve güvenli hale getirmektedir (Brenner, 2010) .

Bilgisayar ağlarının hızlı büyümesi ve uygulama alanları arttıkça siber saldırıların tehlikesi giderek artmaktadır. Öte yandan Intrusion Detection Sistemleri (IDS) ve Intrusion Prevention Sistemleri (IPS) bilgisayar ağlarını saldırılardan korumak için kritik öneme sahiptir. Ancak anormal durumların tespit edilmesindeki zorluk güncel ve çeşitli veri setlerinin bulunamamasıdır. Yazarlara göre 1998'den sonra oluşturulan en popüler veri setleri ya eski ya da çeşitlilik açısından sınırlı ya da yeni saldırı türlerini içermemektedir ya da yük bilgilerini içermemektedir (Kurniabudi vd., 2019) .

Ancak son araştırmalar sadece bir IDS algoritmasına güvenmenin günümüz siber güvenlik tehditlerinin karmaşıklığına karşı yeterli olmayabileceğini göstermektedir. Bu sorunu çözmek için Konsensüs Hibrit Toplu Model (CHEM) geliştirilmiştir ve Random Forest, XGBoost, Karar Ağacı ve Çok Katmanlı Algoritma (MLP) gibi farklı sınıflandırıcılar içermektedir. Ayrıca Açıklayıcı Yapay Zeka (XAI) araçları SHAP ve LIME modelin yorumlanabilirlik özelliğini geliştirmek için eklenmiştir. CHEM'in performansı birkaç karşılaştırmalı veri kümesinde, NSL-KDD ve CIC-IDS2017 gibi veri setlerinde test edilmiştir. CHEM sıfır gün saldırıları bile karşısında %98'in üzerinde doğruluk sağlamak ve yanlış alarmları önemli ölçüde azaltmaktadır. Bu sonuçlar hibrit toplu modellerin siber güvenlik sorunlarını çözmek için uyarlanabilir ve pratik bir yaklaşım olabileceğini göstermektedir (Khan vd., 2025).

Gelişen karmaşık siber saldırılar göz önüne alındığında bu çalışma gelişmiş bir

siber savunma ajanı (ACDA) geliştirmek için hibrit yapay zeka modellerinin yardımıyla bir kavramsal çerçeve tasarlamayı amaçlamaktadır.ACDA'nın amacı sembolik ve alt sembolik yapay zeka modellerini birleştirerek siber tehditlere karşı uyarlanabilir bir algılama ve tepki sağlamaktır.Sembolik yapay zeka modellerinin yorumlanabilirlik sağlayacağı beklenmektedir.Öte yandan derin sinir ağları gibi alt sembolik yapay zeka modelleri büyük miktarda veriden kalıpları tanımaya ve öğrenmeye yardımcı olacaktır.Sonuç olarak ACDA gerçek zamanlı olarak siber tehditleri algılayabilecek, anlayabilecek ve en aza insan müdahalesiyle tepki verebilecek olacaktır.Bu yüksek riskli ortamlarda,örneğin kritik altyapı veya askeri ağlarda tehditlere tepki verme süresini önemli ölçüde azaltacaktır (Loevenich vd., 2024) .

Bu araştırma makinesi öğrenimi modelleri ve optimizasyon tekniklerinin birleşimine dayalı gelişmiş bir Intrusion Detection Sistemini (IDS) önermektedir.Araştırmacılar geleneksel IDS'lerin karmaşık saldırı kalıplarını verimli bir şekilde tanımlamada sınırlamaları olduğunu belirlemişlerdir.Bu sınırlamaları aşmak için karar ağaçları ve güçlendirme gibi makine öğrenimi modelleri kullanılmıştır.Bu modellerin parametreleri Partikül Yorumlama Optimizasyonu (PSO) ve Genetik Algoritma (GA) gibi optimizasyon teknikleriyle optimize edilmiştir.Önerilen yaklaşım yanlış pozitifleri azaltmakta ve algılama doğruluğunu iyileştirmektedir.Önerilen sistemin etkinliği karşılaştırmalı veri setlerinde test edilmiştir.Sistem değişen saldırı kalıplarına verimli bir şekilde uyum sağlayabildiği kanıtlanmıştır.Bu nedenle günümüz siber güvenlik tehditlerine karşı güvenilir bir çözüm olarak değerlendirilebilir (Alkahtani vd., 2025).

Bazı makine öğrenimi ve yapay zeka tabanlı ağ saldırı tespit yöntemlerinin performansı gerçekçi "CSE-CIC-IDS2018" veri kümesi kullanılarak başka bir araştırma çalışmasında (Kanimozhi & Jacob, 2021) analiz edilmiştir. Veri kümesi Botnet,DoS,DDoS ve web saldırılarına benzer çeşitli türde modern siber saldırıları içermektedir. Bu nedenle veri kümesinin kullanımı saldırıları gerçekçi bir şekilde modellemeyi sağlar ki bu da modelin pratik uygulamalar için önemli olduğu anlamına gelir. Yazarlar AI tabanlı modellerin geleneksel sınıflandırıcılara özellikle Botnet saldırıları için ağ saldırı tespitinde daha iyi performans gösterdiğini sonucuna varmışlardır. Ayrıca önerilen modelin bulut tabanlı IDS'de

saldırıları tespit etmek için daha verimli olduğu ve iyi ayarlanmış bir olasılık sağladığı kanıtlanmıştır.

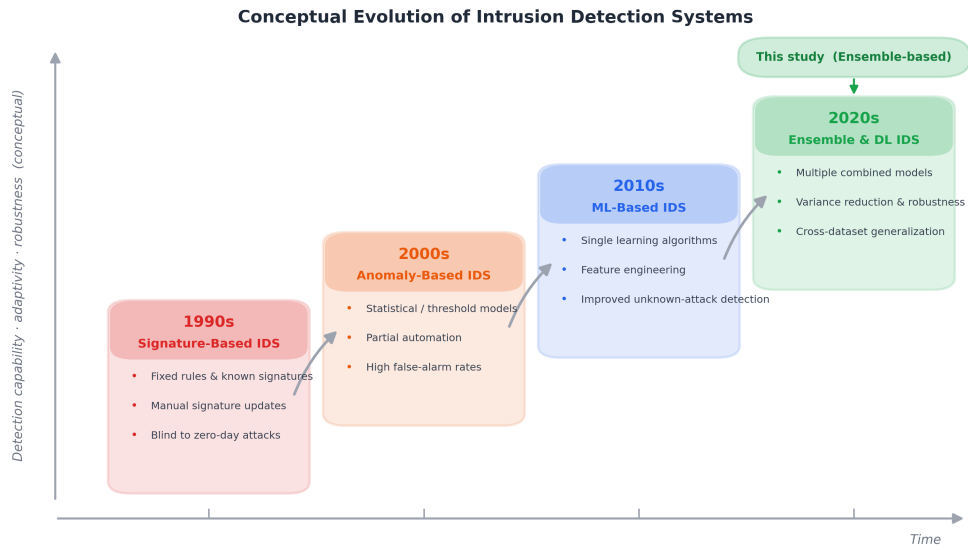
Çalışmalarında (Kilincer vd., 2022) yazarlar çeşitli siber saldırıları tespit etmek için boosting algoritmasını uygulayan bütünleştirilmiş bir saldırı tespit sistemi geliştirmiştir. Yeni bir veri kümesi olan CCIDD önerilmiş ve diğer geleneksel veri kümeleri olan NSL-KDD, UNSW-NB15 ve CSE-CIC-IDS2018 ile karşılaştırılarak verimliliği tahmin edilmiştir. Özellik seçimi için Extra Trees algoritması uygulanmıştır. Çeşitli makine öğrenimi algoritmaları sınıflandırma işlemi yapmak için kullanılmıştır. Deneyler sonucunda önerilen saldırı tespit sistemini daha verimli hale getirebilecek en iyi algoritmalar belirlenmiştir. Light Gradient Boosting Machine (LGBM) ve XGBoost en iyi algoritmalar olarak belirlenmiş ve bu algoritmalar saldırı tespit sistemini iyileştirmiştir. LGBM algoritması XGBoost algoritmasından daha yüksek doğruluk ve işleme süresi açısından daha iyi sonuçlar sağlamış bu nedenle boosting tabanlı algoritmaların saldırı tespit sistemlerinde oldukça etkili olduğu sonucuna varılabilir.

Son bir çalışmada (Liu vd., 2025) yazarlar geleneksel IDS modellerinin sınırlamalarını aşmak için karşıt örnekler kullanan yeni bir saldırı tespit sistemi sunmuştur. Özellikle yeni bir yöntem geliştirilmiştir ki bu yöntem karşıt öğrenme ve artımlı öğrenmeyi birleştirerek sızıntı tespitinin verimliliğini artırır ve aşırı uyum ve felaket unutma gibi sorunları önler.

Önerilen yöntem UNSW-NB15 ve CSE-CIC-IDS2018 gibi çeşitli veri kümelerinde CNN, DNN ve RNN gibi çeşitli makine öğrenimi modelleriyle birlikte değerlendirilmiştir. Sonuçlara göre karşıt örneklerin kullanımının saldırı tespit sisteminin verimliliğini önemli ölçüde artırdığı söylenebilir. (Kaissar vd., 2025) . Bu araştırmada yazarlar CNN tabanlı ağ saldırı tespit sistemlerinin performansını hiperparametre optimizasyonu ile geliştirmeyi amaçlamıştır. UNSW-NB15 ve CSE-CIC-IDS2018 veri kümelerinde Grid Search, Genetik Algoritma, Parçacık Kitle Optimizasyonu ve Gri Kurt Optimizasyonu gibi farklı optimizasyon algoritmaları test edilmiştir. Bu algoritmaların kullanımı CNN modelinin performansını ve verimliliğini önemli ölçüde artırdığı bulunmuştur. Performans ve verimlilik algılama, eğitim ve test zamanları açısından değerlendirilmiştir. Yazarlar Gri Kurt Optimizasyonu algoritmasının algılama performansını artırdığı ve eğitim ve test

zamanlarını azalttığını bulmuşlardır. Bu derin öğrenme tabanlı ağ saldırı tespit sistemlerinin performansını ve verimliliğini artırmak için hiperparametre optimizasyonunun önemini kanıtlamaktadır.

Bu teknolojilerin gelişim yolunu gösteren bir kavramsal gösterim Figür 2.1’de bulunabilir. Figür saldırı tespitinin statik kural tabanlı yaklaşımlardan uyarlanabilir topluluk temelli şemalara doğru gelişimini niteliksel bir şekilde görselleştirmektedir. Her nesil önceki neslin temel sorunlarını çözmektedir.



Şekil 2.1. Saldırı Tespit Sistemlerinin İmza Tabanlıdan Topluluk Tabanlı Yaklaşımlara Kavramsal Evrimi

2.3. Modern Siber Tehdit Ortamı

Zamanın gelişimi ve toplumun teknolojiye olan artan bağımlılığı ile birlikte siber tehditlerin doğasında ve bu tehditlerin türlerinde de birçok değişiklik gözlemlenmiştir. Geçmişte yalnızca yazılımlardaki güvenlik açıklarını hedef alan basit siber tehditlerle karşılaşılıyordu. Günümüzde ise hem büyük ölçekli güvenlik açıklarına hem de teknolojik, organizasyonel hatta insan kaynaklı güvenlik açıklarına yönelik çok daha karmaşık saldırılar gözlemlenmektedir. Bu nedenle siber tehditlerin doğası ve türü ile ilgili değişiklikleri ve geleneksel savunma mekanizmalarının sınırlamalarını iyi bir şekilde bilmek kritik öneme sahiptir.

2.3.1. Dijital Çağda Siber Suçun Evrimi

Siber suçlar virüsler ve bilgisayarlara yasa dışı erişim gibi basit türlerden çok daha karmaşık türlere, ransomware, gelişmiş kalıcı tehditler ve botnetlere kadar gelişti. Bu gelişme kritik altyapıların birbirine bağlanmasının artması, bulut bilişimin gelişimi ve Nesnelerin İnterneti'nin ortaya çıkmasıyla oldu. Bugünün siber suçluları sadece siber saldırılarda usta olmakla kalmıyor aynı zamanda phishing gibi psikolojik yöntemleri de kullanıyorlar.

2.3.2. Geleneksel Saldırı Tespit Sistemlerinin (IDS) Sınırlılıkları

Genel olarak IDS sistemleri ya kural tabanlı ya da imza tabanlı bir yaklaşım kullanarak istilacı saldırıları tespit etmektedir. Buna rağmen geleneksel IDS sistemleri bazı önemli zayıflıklardan muzdarip olup bunlardan bazıları şunlardır:

- Bu tür saldırılar için bilinen bir saldırı imzası olmadığı için sıfır gün saldırılarını tespit edemezler.
- Yanlış alarm sayısı artmaktadır.
- IDS sistemleri yeni saldırı türlerine uyum sağlayamazlar.

Böylece kuruluşlar geleneksel IDS sistemlerinin geniş ve uyarlanabilir saldırılara karşı çok az koruma sağladığını anlamaya başladılar. Bunun üzerine IDS'nin zayıflıkları onu akıllı IDS geliştirmek için zorunlu hale getirdi.

2.3.3. İleri Düzey ve YZ Destekli Saldırı Vektörleri

Tipik olarak geleneksel yöntemlerle yapılan saldırılar genellikle tahmin edilebilir kalıplara dayanmaktadır. Ancak günümüzde saldırganlar otomasyon ve yapay zeka teknolojilerini kullanarak etki alanlarını, verimliliklerini ve hızlarını genişletmektedirler. Bu yapay zekaya destekli saldırı kalıpları artık statik değildirler; sistemin savunma mekanizmalarını kullanarak gerçek zamanlı olarak saldırılarını değiştirerek tespit edilmekten kaçabilirler. Bu tür özellikleri geleneksel saldırı tespit sistemlerinin bu tür yapay zekaya destekli saldırı kalıplarını tespit etmesini zorlaştırmaktadır çünkü bunlar tahmin edilemez ve değişkendirler.

Yapay zekaya destekli bazı yaygın saldırı biçimleri şunlardır:

- **Uyarlanabilir Kötü Amaçlı Yazılım:** Tipik olarak zararlı yazılımlar statik imzalar kullanır ve bunlar kolayca tespit edilebilir ve anti-malware yazılımları tarafından kolayca engellenebilir. Ancak uyumlu zararlı yazılımlar imzaların dinamik olarak değişmesini kullanarak tespit edilmekten kaçınır. Zararlı yazılımlar polimorfizm ve metamorfoz kullanarak her saldırı girişimi için farklı türlerde zararlı yazılımlar üretebilir. Bu nedenle zararlı yazılımlara karşı mücadele neredeyse imkansızdır çünkü imzalar sürekli olarak değişmektedir. Ayrıca uyumlu zararlı yazılımlar anti-malware yazılımlarını tanıyabilir ve bunlardan kaçmak için yapay zekayı kullanabilirler. Bu nedenle bu tür saldırılar kalıcı hale gelir.
- **YZ Destekli Kimlik Avı (Phishing):** Siber saldırıların en iyi yöntemlerinden biri olan phishing hala çok etkili fakat yapay zekanın dahil edilmesi onu daha da güçlü hale getirmiştir. Doğal dil üretimi, duygu analizi ve kullanıcı profillemeye yöntemleri kullanılarak saldırganlar hedeflenen kişilere özel phishing e-postaları veya web sayfaları oluşturabilirler. Salırganlar e-postaları çok doğru bir şekilde hazırlayabilir, güvenilir bir kişi veya şirketin dil tarzını taklit edebilirler. Bu da hedeflenen kişinin bir e-postanın phishing saldırısı olup olmadığını ayırt etmesini zorlaştırır. Phishing saldırılarına yapay zekanın dahil edilmesinin en endişe verici özelliği aynı anda milyonlarca kişiye bu saldırıların gerçekleştirilebilmesi ve her kişiye özel olarak uyarlanabilmesidir.
- **Özerk İstismlarlar (Autonomous Exploits):** Geleneksel saldırı mekanizmalarının çoğunlukla yapılandırma ve dağıtım tabanlı olduğu aksine otomatik saldırılar az veya hiç insan etkileşimi olmadan çalışabilir. Bunun nedeni yapay zekayı ve makine öğrenmesini ağ taraması, zayıflıkları tespit etme ve saldırıları dinamik olarak yürütme için kullanmalarıdır. Güçlendirme öğrenimi veya genetik algoritmaların yardımıyla saldırılar hedef sistemin tepkisine göre yaklaşımlarını geliştirebilirler. Bu kendini geliştirme yeteneği geleneksel güvenlik önlemlerine karşı saldırıları dirençli hale getirir.
- **Çekişmeli Kaçınma (Adversarial Evasion):** Bu yapay zeka teknolojisini kullanan siber saldırıların en gelişmiş yöntemlerinden biridir. Karşıtlık makine öğrenmesinde saldırganlar yapay zeka sistemlerinin girdi verilerini manipüle ederek sınıflandırmalarında hatalara neden etmeye çalışırlar. Örneğin saldırganlar dikkatlice hazırlanmış gürültüyü ağ trafiği kalıplarına ekleyerek makine öğrenimi sistemini yanlışlıkla saldırıyı meşru

olarak sınıflandırmaya ikna edebilirler. Veya kötü amaçlı yazılımın özelliklerini hafifçe değiştirebilirler. Bu tür saldırılarda yapay zekayı kullanan saldırganların savunmayı yapay zekayı kullanarak manipüle etmeleri zayıflıklarıdır.

Bunların yanı sıra bu vektörlerin ileri düzeyde olması dışında açık kaynaklı yapay zeka kütüphanelerinin, önceden eğitilmiş modellerin ve otomatik saldırı araçlarının kolayca erişilebilir olması bireylerin bu konuda tam bir bilgiye sahip olmamalarına rağmen yapay zekayı ileri ve büyük ölçekli saldırılar gerçekleştirmek için kullanabilmelerini sağlıyor.

Bu güvenlik tehditinin seviyesini arttırmış ve güvenlik tehditinin yaygınlaştığının belli olduğu görülüyor. Bu ileri düzeyde saldırı vektörlerinin varlığı geleneksel güvenlik tekniklerinin ve tek algoritmaya dayalı savunma tekniklerinin başa çıkamadığına işaret ediyor. Bu tür sürekli gelişen siber güvenlik tehditlerine karşı koyabilmek için siber güvenlik topluluk öğrenme tekniğini benimsemeli bu teknik çeşitli makine öğrenimi paradigmalarının kullanılmasını içeriyor ve karma savunma stratejileri oluşturmayı sağlıyor. Bu sistemin bu tür sürekli gelişen yapay zeka temelli siber güvenlik tehditleriyle daha sağlam ve doğru bir şekilde başa çıkmasını sağlayacaktır.

Çizelge 2.1. Geleneksel Siber Saldırıları ile YZ Destekli Saldırı Vektörlerinin Karşılaştırması

Yön	Geleneksel Saldırıları	YZ Destekli Saldırıları
Örüntü	Statik, kural tabanlı, öngörülebilir.	Dinamik, uyarlanabilir; sistem yanıtlarından öğrenir.
Yürütme Hızı	Görece yavaş; el emeği veya önceden kodlanmış betikler.	Otomatik ve hızlı; asgari insan gözetimiyle büyük ölçekli.
Kişiselleştirme	Genel yükler (payload), birçok kurban için aynı.	NLG, duygu analizi ve profillemeye kullanılarak yüksek düzeyde uyarlanmış.
Kaçınma Teknikleri	Sınırlı; genellikle imza tabanlı sistemlerle tespit edilebilir.	İleri düzey; polimorfizm, metamorfizm, çekişmeli MÖ.
Ölçeklenebilirlik	Genişlemek için önemli insan kaynağı gerektirir.	Otomasyon yoluyla aynı anda milyonlarca hedefi vurabilir.
Örnekler	Klasik kimlik avı, bilinen kötü amaçlı yazılım imzaları, elle yapılan istismarlar.	Uyarlanabilir kötü amaçlı yazılım, YZ üretimi kimlik avı, özerk istismarlar.
Giriş Engeli	Orta ile yüksek düzeyde teknik uzmanlık gerektirir.	Açık kaynaklı YZ kütüphaneleri ve otomatik çerçevelerle düşürülmüştür.
Savunma Zorluğu	Geleneksel IDS ve antivirüs araçlarıyla tespit edilebilir.	Statik savunmaları çoğunlukla aşar; topluluk tabanlı modeller gerektirir.

Örneğin siber suçların gelişimi,saldırı tespitinin geleneksel yöntemlerindeki eksiklikler ve yapay zekâ kullanılarak gerçekleştirilebilen yeni türdeki saldırıların yükselişi tüm bunların akıllı savunma önlemlerinin uygulanması gerekliliğine işaret ediyor.Bu araştırmanın yönteminin temel varsayımı budur.

2.4. Önceki Çalışmaların Karşılaştırmalı Analizi

Farklı makine öğrenimi yaklaşımları kullanılarak IDS'ler için önemli gelişmeler sağlanmış olsa da özellikle benzer koşullarda farklı algoritmaların karşılaştırıldığı çalışmalar arasında önemli bir boşluk hala var (Leevy & Khoshgoftaar, 2020).

Çizelge 2.2. CICIDS2018 Veri Kümesi Üzerine Önceki Çalışmaların Karşılaştırmalı Sonuçları

Yazarlar	Recall	Accuracy	Precision
(Atefinia & Ahmadi, 2021)	100,00%	100,00%	100,00%
(Basnet vd., 2019)	n/a	99,00%	n/a
(Catillo vd., 2020)	98,90%	99,20%	95,00%
(Chadza vd., 2019)	n/a	97,00%	n/a
(D'hooge vd., 2020)	79,00%	96,00%	99,00%
(Ferrag vd., 2020)	98,18%	97,38%	n/a
(Lima Filho vd., 2019)	100,00%	n/a	100,00%
(Fitni & Ramli, 2020)	97,10%	98,80%	98,80%
(Gamage & Samarabandu, 2020)	98,27%	98,40%	97,79%
(Hua, 2020)	98,37%	98,37%	98,14%
(Huancayo Ramos vd., 2020)	99,99%	99,99%	100,00%
(Kanimozhi & Jacob, 2019a)	100,00%	100,00%	100,00%
(Kanimozhi & Jacob, 2019b)	100,00%	99,97%	99,96%
(Demir vd., 2020)	99,69%	99,69%	99,70%
(Choi vd., 2020)	82,25%	99,99%	81,75%
(Li vd., 2020)	100,00%	n/a	n/a
(Lin vd., 2019)	96,00%	96,20%	96,00%
(Zhang vd., 2020)	98,00%	97,90%	98,00%
(Mohamed vd., 2018)	97,23%	96,87%	96,70%
(Kurniabudi vd., 2019)	97,31%	86,92%	92,22%
(Venkatraman vd., 2019)	85,53%	81,17%	76,20%

Değerler Tabloda 2.2’de aynı ölçekte %81,17’den %100,00’e kadar çok büyük bir değişkenlik göstermektedir. Birçok çalışma mükemmel doğruluk değerleri rapor etmektedir. Bu değişkenliğin algoritmalarından kendileri kaynaklandığını düşünmek pek olası değil çünkü literatürde gözlemlenen bu değişkenliği nedenleyen açıkça tanımlanmış üç koşul vardır. Aşağıdaki yorumlar bu koşulların analiz edilirken dikkate alınması amacıyla bir eleştiri olarak değil sadece bazı koşulları belirlemeye çalışmaktadır.

İlk koşul özellik seçimidir. CIC veri setleri özel olarak oluşturulmuş bir test yatağında betiklenmiş saldırılarla oluşturulmuştur ve farklı hedef portlarına karşı farklı türlerde saldırıları üretir. Eğer port numarası, akış ID’si, zaman damgası veya uç nokta adresi gibi özellikler eğitim verilerine dahil edilirse sınıflandırıcı bu yakalama yapılandırmasını etiketlere mükemmel doğrulukla eşleyebilir ancak gerçek trafik özelliklerini göz önünde bulundurmadan. Sınıflandırıcı bu durumda genelleme yapamayacak çünkü farklı bir port’a yönelik bir saldırıyı yorumlayamayacaktır. Bu nedenle mevcut çalışma sadece beş hacimsel ve zamansal akış özelliğini dikkate almaktadır ve bu seçim bağımsız olarak 4.7.5 bölümünde yapılan atıf analizinde doğrulanmıştır.

İkinci koşul test kümesindeki sınıf dağılımıdır. Doğruluk değeri iyi-saldırı oranına bağlıdır çünkü yanlış pozitifler iyi sınıftan ve doğru pozitifler saldırı sınıfından gelir. Hatırlama değeri sadece saldırı sınıfına göre hesaplanır ve belirtilen oranla etkilenmez bu nedenle bu değer sonraki hesaplamalarda sabit bırakılabilir. Başka bir deyişle yeniden dengeleme doğruluk değerini artırabilir ancak sınıflandırıcıyı bozmaz ve bu etkinin büyüklüğü mevcut deneyim sonuçlarına dayanarak doğrudan tahmin edilebilir. Bagging sınıflandırıcısı test kümesindeki 1.924.285 iyi akıştan 9.171 yanlış pozitif üretmektedir ki bu %0,477’lik bir yanlış pozitif oranına karşılık gelmektedir. Yeniden dengelenmiş 235.616 akışlık veri kümesinde aynı sınıflandırıcı %96,63’lük bir hatırlama ve %99,51 precision sağlar. F1 skoru %98,05’tir yani bu 3,38 ve 1,67 oranında bir artışa karşılık gelmektedir ancak herhangi bir yeniden eğitim veya eşik ayarı yapılmadan ve bunun bedeli olarak doğruluğu %98,08’e düşürüldükten sonra. Bu nedenle bu etki sadece temel oranlara dayalı ölçütlere yöneliktir. Beklendiği gibi veri kümesi temelli sonuçlar bu eğilimi göstermektedir: Aynı sınıflandırıcı ve eğitim veri kümesini kullanan Logistik Regresyon modeli neredeyse

dengeli UNSW-NB15 alt kümesinde (%54,68 saldırı) %73,24 F1 ve CICIDS2018 veri kümesinde(%7,31 saldırı) %11,09 F1 elde etmektedir.

Üçüncü koşul değerlendirme kapsamıdır.Saldırı türleri oldukça farklı ayırt edici özelliklere sahiptir: Hacim temelli saldırılar gibi DoS çok farklı trafik istatistiklerine neden olur ve bu istatistikler kolayca bayt ve paket sayılarına göre sınıflandırılabilir ancak sızma ve düşük hızda web saldırıları iyi trafiği taklit eder. Bu nedenle tek sınıflı bir ölçüt tüm saldırı kümesi üzerinde bir sınıflandırıcının performansını değerlendirmek için yeterli değildir. Ayrıca tek bir veri kümesi üzerinde değerlendirme eğitim ve test trafiği arasında olası bir dağılım değişimini ortadan kaldırır ve çoklu veri kümesi üzerinde değerlendirme göre daha basit bir görevdir.

Yukarıda belirtilen tüm koşullar eklenmelidir ve bu nedenle Tabloda 2.2’de gösterilen üst sınır herhangi bir sınıflandırıcının performansını açıklamak için yeterli olacaktır ve elbette değerlendirilen modellerin performansından daha iyi olması gerekmez. %100,00 doğruluk,hatırlama ve doğruluk değerleri test kümesinde hiç yanlış sınıflandırma olmadığını gösterir ki bu bir sınıflandırıcıdan çok değerlendirme yönteminden daha net bir sonuçtur.

Diğer yandan mevcut çalışma her üç koşula da tam tersi bir yaklaşım izlemektedir: Beş özellik eksik özellik kimliklerine sahip, üç bağımsız veri kümesi ve toplamda 10.799.503 akış ve orijinal %89,1/%10,9 sınıf oranları ile hiçbir yeniden dengeleme adımı atılmadan. Bu nedenle elde edilen ölçütler Tabloda 2.2’dekilerden daha düşüktür ve bu tam olarak olması gereken şeydir- performansın dağıtım koşullarında alt sınırı.

2.5. Araştırma Boşluğu ve Motivasyon

MÖ ve YZ çerçevelerinin IDS’ye uygulanmasında önemli gelişmeler sağlansa da mevcut en son literatürde birkaç önemli boşluk var ki bunlar daha fazla araştırmayı gerektiriyor.

Öncelikle tek bir veri kümesi üzerinde eğitilip doğrulanan modellerin aşırı uygulanması onların genelleme yeteneğini sınırlıyor. Tek bir veri kümesi üzerinde eğitilip test edilen modeller o veri kümesinin belirli kalıplarını öğreniyorlar,saldırı davranışlarını değil bu da verilerin dağılımındaki değişkenlik nedeniyle pratik uygulamalarda etkisiz hale getiriyorlar.

İkincisi çeşitli MÖ algoritmaları üzerinde yapılan çok sayıda karşılaştırmalı çalışma olmasına rağmen tutarlı bir ön işleme prosedürü ve özellik temsili yok. Bu da performanslarının objektif bir şekilde karşılaştırılmasını zorlaştırıyor.

Üçüncüsü IDS'nin bir başka sorunu saldırı tespit veri setlerinin heterojenliği özellikle de yapıları isimleri ve trafik özellikleri bakımından. Bu da çapraz veri kümesi değerlendirme prosedürünü ve daha geniş bir ortam için IDS modelleri geliştirmeyi zorlaştırıyor.

Dördüncüsü topluluk öğrenmesinde bagging ve boosting yaklaşımlarına bazı olumlu sonuçlar sağlanmış olsa da bu yaklaşımlar esasen tek veri kümesi ortamında araştırılmıştır. Bagging ve boosting tekniklerinin çok veri kümesi ortamında IDS'nin sağlamlığını nasıl artırabileceği konusunda kapsamlı bir şekilde araştırılmamıştır.

Son olarak araştırmacılar veri kümesinde sınıf dengesini sağlamak için agresif bir şekilde yeniden örnekleme tekniği uyguluyorlar bu da veri dağılımını etkileyebilir. Bu nedenle değerlendirme prosedürleri veri kümesinin orijinal dağılımını koruyacak şekilde değiştirilmelidir. Bu boşlukları kapatmak amacıyla bu çalışma üç popüler IDS veri kümesi (CICIDS2017, CICIDS2018, UNSW-NB15) üzerine kurulu yeni bir çok veri kümesi IDS çerçevesi öneriyor.

3. GEREÇ ve YÖNTEM

Bu bölümde önerilen topluluk temelli istila algılama sisteminin geliştirilmesi,uygulaması ve test edilmesi için kullanılan yöntemin ayrıntılı bir açıklamasını sunmaktadır. Çok büyük hacimli ve farklı türde verilerle uğraşmayı ve çoklu veri setlerinde genelleme yeteneğinin geliştirilmesini hedeflenmektedir.

Öncelikle bu bölümde geleneksel istila algılama yaklaşımlarının sınırlamalarını ve uyarlanabilir bir yaklaşımın kullanılmasının gerekliliğini belirtmektedir. Ardından araştırma yöntemimizi özellikle veri setleriyle çalışmak,veri setinin ön işlenmesi ve özelliklerin birleştirilmesi gibi konulara dikkat çekerek açıklanmaktadır.

Araştırmamızın ana aşamalarından biri farklı karşılaştırma veri setlerini tek bir ortak platforma birleştirerek çoklu veri kümelerinde testler yapmak oldu.Veritil temizleme,etiket normalizasyonu ve özelliklerin birleştirilmesi gibi detaylar bu bölümde veriliyor.Ayrıca istila algılama problemini çözmek için kullanılan makine öğrenimi algoritmaları ve verimlilik ölçütleri de kısa bir şekilde belirtilmektedir.

3.1. Araştırma Metodolojisine Genel Bakış

Bu bölüm bu tezin tasarımını,uygulamasını ve değerlendirilmesini sağlamak için kullanılan araştırma yöntemini açıklıyor.Özellikle mevcut yaklaşım tek bir MÖ modeli olmadığının varsayımına dayanıyor ki bu model modern siber tehditlerin dinamik doğasını etkili bir şekilde kapsayabilir.

Bunun sonucunda mevcut araştırma çoklu sınıflandırıcı özelliklerini değerlendiren bir çeşitlilikte topluluk öğrenimi yönteminin uygulanmasını öneriyor.Çeşitlilikte topluluk yaklaşımı sayesinde birden fazla sınıflandırıcı özelliğinin kullanımı ile istilayı tespit sistemlerinin doğruluğu artırabilir.

Önerilen yaklaşımın temel özelliklerinden biri aynı anda birkaç referans veri kümesinin kullanılması.Deneyde tek bir istilayı tespit veri kümesi yerine araştırmacı CICIDS2017,CICIDS2018 ve UNSW-NB15 veri setlerini kullanacak.

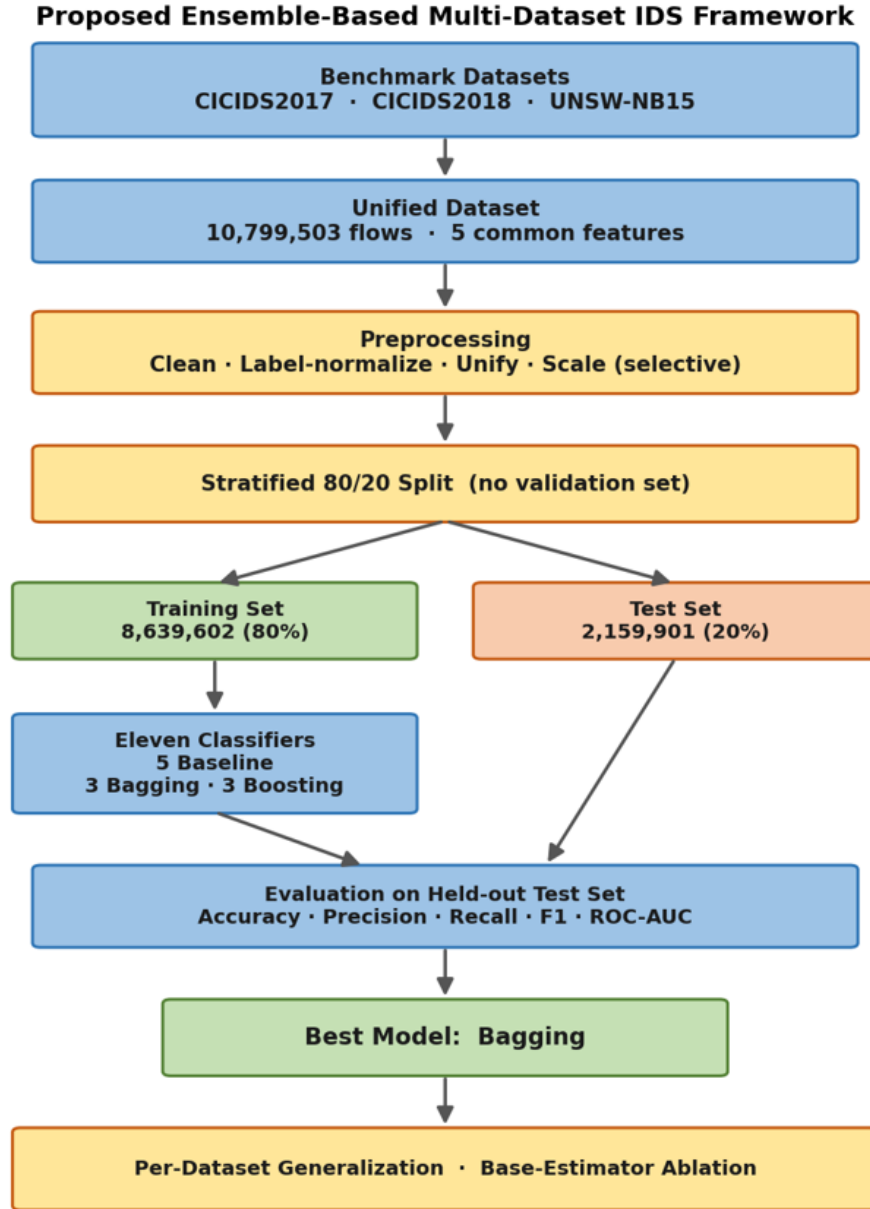
Mevcut araştırmanın yöntemi aşağıdaki adımları içeren tutarlı bir boru hattı içerir:

- Veri kümesi bütünleştirme ve normalleştirme
- Özellik ön işleme ve birleştirme
- Sınıflandırıcıların eğitimi ve değerlendirilmesi
- Algoritma performansının karşılaştırılması.

3.1.1. Hibrit Savunma Yaklaşımının Kavramsal Çerçevesi

Önerilen yöntem farklı modellerin karşılaştırmalı analizinin sonucu olan çerçeve tabanlı bir saldırı algılama yöntemine dayanmaktadır. Diğer mevcut yöntemlerin tek bir algoritmaya dayandığına karşı önerilen yöntem farklı modellerin analizini ve karşılaştırmasını içermektedir.

Görüldüğü gibi Şekil 3.1’de önerilen çerçeve genel yöntemi tek bir çizgi ile ham karşılaştırma veri setlerinden eğitilmiş ve değerlendirilmiş sınıflandırıcılara göstermektedir. Çerçeve üç bağımsız karşılaştırma veri kümesi olan CICIDS2017, CICIDS2018 ve UNSW-NB15’i birleştirerek tek bir veri kümesine başlar. Veri kümesi 10.799.503 ağ akışı kaydı ile beş ortak özelliğe sahiptir. Çerçeve geliştirilirken alınan kritik bir karar birden fazla bağımsız kaynağın kullanılması yerine tek bir veri kümesi kullanmak olmuştur. Çünkü farklı ağlardan elde edilen trafik verileri üzerine eğitilen modeller birçok ağda saldırı kalıplarını tespit edebilecek ve belirli yakalama trafiğine aşırı uyum sağlamayacaktır. Birleştirilmiş veri kümesi ön işleme yapılarak %80 eğitim kümesi ve %20 test kümesine bölünmüştür. Eğitim kümesi üzerinde on bir makine öğrenimi sınıflandırıcısı eğitilmiştir ve bu sınıflandırıcılar aynı ölçütler kullanılarak test kümesi üzerinden değerlendirilmiştir. Test edilen tüm modeller arasında en iyisi Bagging’dır.



Şekil 3.1. Önerilen Topluluk Tabanlı Çok Veri Kümeli Saldırı Tespit Çerçevesinin Genel İş Akışı (seçilen model: Bagging; birleştirilmiş derlem: 10,8M akış; bölünme: 80/20).

Bu bağlamda kavramsal çerçeve farklı sınıflandırma ailelerinin güçlerine dayanmaktadır.

Bunlar şunlardır:

- Karar ağacı tabanlı sınıflandırıcılar (Rastgele Orman, Ek Ağaçlar ve Karar Ağacı)
- Artırma yöntemleri (AdaBoost, Gradient Artırma ve XGBoost)
- Olasılıksal, mesafe temelli ve doğrusal modeller (Naive Bayes, K-En Yakın Komşu, Doğrusal SVM ve Logistik Regresyon).

Kavramsal çerçevenin ana hedefleri arasında şunlar vardır:

- Modellerin çeşitliliğinden yararlanarak tespit verimliliğini artırmak.
- Modellerin karşılaştırılması yoluyla yanlış pozitif oranlarını azaltmak.
- Heterojen veri setleri kullanarak modellerin genelleme yeteneğini geliştirmek.
- Özelliklerin bireysel veri setlerine olan bağımlılığını azaltmak için özellikleri birleştirmek.

3.2. Siber Güvenlikte Topluluk Öğrenmesi Kullanımının Gerekçesi

Güvenlik sorunlarını ele almak için topluluk öğrenimi kullanma ihtiyacı Tek model yaklaşımlarının Intrusion Detection Sistemleri (IDS) uygularken sahip olduğu eksikliklere bağlıdır. Hem geleneksel hem de makine öğrenimi algoritmaları IDS’de eğitim kümesine aşırı uyum sağlama ve yeni verilere kötü genelleme yapma gibi sorunlarla karşı karşıyadır. Tek model yöntemlerinin diğer zayıflıkları yüksek yanlış pozitif oranı ve sürekli değişen tehdit doğasına uyum sağlayamama düşük esnekliğidir.

Özellikle büyük ölçekli ve çoklu modal uygulamalarda farklı veri dağılımları ve özellikler büyük bir tehdit oluşturmaktadır. Ancak topluluk öğrenimi yukarıda bahsedilen sorunları aşmada yardımcı olur çünkü birden fazla modelin ve sonuçlarının tek bir bileşik karar alma süreci yarattığı gerçeğinden dolayıdır. Bagging ve Boosting algoritmaları algoritmaların çeşitliliğinden yararlanır ve varyans azaltarak performanslarını geliştirir. Farklı karar sınırları karmaşık ağ trafiği veri kümesindeki öğeler arasındaki ilişkiyi yakalamayı sağlar.

Yanlış sınıflandırma giriş verilerinin ciddi sonuçlara yol açabileceği Intrusion Detection Sistemlerinde topluluk öğrenimi güvenilirliği sağlamak için ek bir garanti

sunar. Dengesiz veri kümelerinde duyarlılık oranını düşürmeden doğruluk oranını artırabilme olasılığı topluluk öğreniminin en uygun yaklaşımını Intrusion Detection için yapmaktadır.

3.2.1. Derin Öğrenmenin Hariç Tutulmasının Gerekçesi

Derin mimariler açıkça değerlendirme dışına alındı. Sebep yeterli donanım kaynaklarına erişimin olmaması değil tam tersine operasyonel değerlendirme sebebidir; aslında CUDA 12.5 ile güçlendirilmiş bir RTX 4070 üzerinde (3.7.1 bölüm) deneyler yapıldı ve bu ölçekte verilerde bu çalışmada bahsedilen tekrarlayan ve konvolüsyonel modelleri eğitmek için yetkin olabildi. Bu karar gerçek maliyet özelliklerine dayanmaktadır ve çalışma süresi analizi ile gösterilmektedir.

Birleştirilmiş veri kümesi için seçilen Bagging ensemble modeli 148,9 saniye içinde eğitiliyor ve 2,4 saniyede 2.159.901 test akışı sınıflandırıyor yani yaklaşık olarak her saniyede 900.000 akışa sınıflandırıyor. XGBoost ise aynı veri kümesini eğitmek için 4,9 saniye ve aynı veri kümesini sınıflandırmak için bir saniyeden az zaman alıyor (Şekil 4.9). Bu soruna atıfta bulunmak için kullanılan derin modellerin eğitimi bile daha yavaş ve çıkarım maliyetleri katman sayısına göre doğrusal olarak değil de ağırlık sayısına göre ilerliyor.

Pratik açıdan bir istilayı tespit etme sisteminin dağıtımı için düşük çıkarım maliyeti bir zorunluluktur çünkü bir sistem ya trafiği satır hızında işleyemiyorsa ya akışları düşürmek zorunda kalır ya da ekstra bir gecikme getirir ve bu sistemin kullanılamaz hale gelmesine neden olur. Yeniden eğitimin maliyeti ikinci kriter çünkü ağ trafiğindeki kavram kayması yeniden eğitmeyi gerektiriyor (7. bölüm) ve dakikalar yerine saatler içinde yeniden eğitilebilmek daha sık bir yeniden eğitimin yapılmasını sağlıyor.

İncelenen literatürde derin mimarilerin bu sonuçları önemli ölçüde etkileyebileceği yönünde bir sebep bulunmamaktadır; aslında neredeyse mükemmel performans iddiasında bulunan çalışmalar neredeyse her zaman 2.4 bölümde açıklanan yöntemi takip ediyor ve bu çalışmada ulaşılan tavan yani beş sınıflandırıcının bir araya getirilmesiyle 1,1 F1 yüzde puanı civarında bir kümelenme gösteriyor bu da model kapasitesinin değil beş özellik temsiliğinin bir tıkanıklık olduğunu gösteriyor. Bu nedenle modelin karmaşıklığını genişletmek yerine temsili genişletmek tercih edilir ve bu yaklaşım gelecekte 6.5 bölümde ele alınacaktır.

3.3. Veri Kümeleri ve Veri Kaynakları

Araştırmanın geçerliliğini sağlamak için saldırı algılama üzerine yapılan çalışmada kullanılan veri kümeleri yüksek kaliteli, çeşitli ve temsili olmalıdır. Önerilen çerçevenin tüm yönlerinin analizinde tek bir veri kümesi yerine birden fazla veri kümesi kullanılacaktır. Aşağıda belirtilen üç açık kaynaklı ve popüler veri kümesi analiz edilmiştir:

- CICIDS2017 (Sharafaldin vd., 2018)
- CICIDS2018 (Sharafaldin vd., 2019)
- UNSW-NB15 (Moustafa & Slay, 2015)

Yukarıda bahsedilen veri kümeleri literatürde sıkça bahsedilmiştir, büyük bir saldırı çeşitliliğini içerir ve gerçekçi trafik kalıplarını yansıtır. Hem zararsız trafik hem de saldırıları (DoS, DDoS, brute force, saldırı ve web saldırıları) içerirler. Ön işleme sonucunda üç veri kümesi birleştirilerek tek bir veri kümesi oluşturulmuştur. Bu veri kümesinde 10.799.503 gözlem bulunmaktadır. Tüm türdeki saldırılar tek bir kategoriye-”Saldırı” olarak birleştirilmiştir. İkinci kategori ise ”Zararsız” demektir. Elde edilen veri kümesi oldukça dengesizdir: 9.621.424 zararsız akış (%89,1) ve 1.178.079 saldırı akışı (%10,9).

- Süre (Duration)
- Kaynak baytları (Source bytes)
- Hedef baytları (Destination bytes)
- Kaynak paketleri (Source packets)
- Hedef paketleri (Destination packets)

3.3.1. Veri Kümeleri: Yapı ve İlgililik

Mevcut araştırmayla ilgili literatür arasındaki temel fark mevcut araştırmada tek bir veri kümesi yerine birkaç veri kümesi kullanmasıdır. İlk veri kümesi CICIDS2017 olarak adlandırılıyor ve beş gün boyunca toplanan ağ trafiği verilerini içeriyor hem zararsız trafik hem de çeşitli saldırıları (Brute Force, DoS, DDoS, Web Saldırısı, Sızma ve Botnet). CICIDS2018 veri kümesi SQL enjeksiyonu ve XSS gibi saldırıları da içeriyor. UNSW-NB15 veri kümesi dokuz tür saldırı içeren yeni bir veri kümesi olarak kabul ediliyor (Fuzzers, Analiz, Arka Kapı, DoS, Saldırı, Genel, Tanıma, Shellcode ve Solucanlar).

Yapısal olarak yukarıda bahsedilen tüm veri kümeleri akış trafiği veri kümeleridir ve bağlantı süresi,paket boyutu ve zamanla ilgili özellikler gibi çeşitli istatistiksel özelliklere sahip olabilirler. Kullanılan veri setlerindeki temel sorun çeşitlilik-trafik isimlendirme,dağılım ve temsillerindeki farklılıklardır. Bu nedenle söz konusu veri kümeleri, asgari ortak özellik alanında birleştirilmiş; ön işleme ve dönüştürme işlemlerinin ardından ikili sınıflandırma problemine dönüştürülmüştür.

3.3.2. Veri Bütünlüğü, Güvenilirlik ve Kıyaslama Gerekçesi

Veri kümelerinin seçimi üç önemli kritere bağlıdır:Veri bütünlüğü, güvenilirlik ve karşılaştırılabilirlik.

- **Veri Bütünlüğü:** Seçilen veri kümeleri kontrol edilmiş ortamlarda üretilmiştir.CI-CIDS2017 ve CICIDS2018 verileri Kanada Siber Güvenlik Enstitüsü(CIC)tarafından üretilmiştir.Öte yandan UNSW-NB15 veri kümesi karma test ortamında üretilmiştir. Kaliteli veriler,gerçekçi trafik ve düşük gürültü makine öğrenimi deneylerini daha verimli hale getirir.
- **Güvenilirlik:** Seçilen veri kümeleri popüler ve diğer araştırmacılar tarafından iyi referans alınmıştır.Birden fazla veri kümesi kullanmak tek bir veri kümesine bağımlılığı azaltarak güvenilirliği artırabilir.
- **Karşılaştırılabilirlik:** Seçilen tüm veri kümeleri DoS,DDoS,brute force,saldırı,web uygulamaları,botnet ve normal trafik gibi farklı türdeki saldırıları içermektedir.

Tüm seçilen veri kümeleri avantajlarına sahip olsa da sadece birini kullanmak aşırı uyuma ve kötü genelleme yeteneğine neden olabilir.Bu nedenle bu çalışma aynı deneysel ortamda birden fazla veri kümesi kullanarak sonuçları gerçekçi bir şekilde karşılaştırmayı sağlamaktadır.

3.4. Veri Ön İşleme Hattı

Saldırı tespit sistemi verilerin ön işlenmesine bağlıdır çünkü ağ trafiği verilerinin kalitesi modelin performansını etkiler. Veri kümesinin farklı karşılaştırma kaynaklarından toplandığı göz önüne alındığında verilerin ön işlenmesinde standartlaştırılmış bir yönteme ihtiyaç vardır. Veri ön işleme aşamaları şunlardır: Veri temizleme,Etiket normalizasyonu,

Özellik standartlaştırması, Özelliklerin birleştirilmesi ve veri kümesinin birleştirilmesi.

3.4.1. Veri Temizleme ve Gürültü Azaltma

Ancak ağ trafiği verilerinde tutarsızlıklar, çift kayıtlar, gereksiz bilgiler ve dengesizlikler bulunabilir ki bu durum öğrenme sürecini olumsuz etkileyebilir. Bu konuda bu araştırma çalışması verilerin işlenmesinde tutarlı ve bütünlü bir strateji sunmaktadır böylece tutarlılık sağlanacak, kalite iyileştirilecek ve birden fazla veri kümesinin birleştirilmesi mümkün olacaktır. Önerilen boru hattının ana aşamaları aşağıda verilmiştir:

- **Veri Temizleme:** Çift kayıtları, geçersiz veya sonsuz değerleri ve başlık satırlarını içeren tutarsızlıkları ve geçersiz giriş verilerini ortadan kaldırmak.
- **Etiket Normalizasyonu:** Tüm veri kümelerinin tutarlılığını sağlamak için etiketleri birleştirilmiş ikili bir formata (Zararsız/Saldırı) dönüştürmek.
- **Özellik Standartlaştırma:** Veri kümelerinin tüm özelliklerini format ve isim bakımından standartlaştırmak.
- **Özellik Seçimi ve Birleştirme:** Tüm veri kümelerinde ortak olan en az özellik kümesini (Süre, Kaynak Bayt Sayısı, Hedef Bayt Sayısı, Kaynak Paket Sayısı, Hedef Paket Sayısı) seçmek.
- **Veri kümesi Birleştirme:** Tüm ön işlenmiş veri kümelerini tek bir derlemeye birleştirmek

3.4.2. Özellik Mühendisliği ve Boyut İndirgeme

Özellik Seçimi ve Birleştirme

Mevcut yöntemlerin tüm özellik alanlarını tek bir veri kümesinde ele aldığına karşı bu araştırmada farklı veri kümelerini entegre etmek için en az özellik alanı ele alınmıştır. Farklı veri kümelerinin farklı biçimlendirme ve özellik isimlendirme kuralları olduğu için sadece anlam ve biçim açısından benzer özellikler seçilmiştir. Seçilen özellikler şunlardır: (1) Süre; (2) Kaynak Baytları; (3) Hedef Baytları; (4) Kaynak Paketleri; (5) Hedef Paketleri. Bu beş akış seviyesi özelliği farklı veri kümelerinde hesaplama hızı, doğruluk ve tutarlılık arasında mükemmel bir denge sağlamaktadır. Belirli veri kümelerine özgü ve herhangi bir bilgi taşımayan değişkenler atılmıştır. Özellik listesi joblib kullanılarak kaydedilmiş ve daha sonraki analizlerde aynı özellik alanını sağlamak için kullanılmıştır.

Özellik Ölçekleme

Ağ trafiği özelliklerinin değerleri boyut açısından oldukça değişmektedir. Bu sorunu çözmek için özellik ölçeklendirmesi (standartlaştırma) Gaussian Naive Bayes, KNN, Lineer SVM ve Logistik Regresyon gibi özellik boyutuna duyarlı modeller için uygulanmıştır. StandartScaler sadece eğitim kümesine uydurulmuş ve daha sonra test kümesine uygulanarak veri sızıntısını önlemek için kullanılmıştır. Ağaç tabanlı ve topluluk modelleri çünkü boyutlara duyarlı değil ölçeklendirilmemiş özelliklerle eğitilmiştir.

Boyutsallık

Yüksek boyutlu özellik alanına karşı bu araştırmada düşük boyutlu bir özellik alanı seçilmiştir. Son özellik kümesinin minimum olduğu ve farklı veri kümelerinin entegrasyonunu amaçladığı göz önüne alındığında PCA gibi boyut azaltma teknikleri kullanılmamıştır.

3.4.3. Sınıf Dengesizliğinin Ele Alınması

Dengesiz sınıflar IDS'de (Intrusion Detection System-İhlal Algılama Sistemi) önemli bir sorun teşkil etmektedir çünkü çoğu durumda meşru trafiğin miktarı diğer herhangi bir tür aktiviteden çok daha fazla olacaktır. Bu nedenle sınıflandırıcılar çoğunluk sınıfında eğitilir ve saldırıları tespit edemezler. Bu çalışma herhangi bir yeniden örnekleme yapmadan ve orijinal veri dağılımını koruyarak dengesiz hassas yöntemlerle bu sorunu ele almaktadır. Aşağıdaki yöntemler kullanılmıştır:

- **Katmanlı Veri Bölme:** Eğitim ve test kümeleri tüm bölmelerde veri dağılımını korumak için katmanlı örnekleme yoluyla ayrılmıştır.
- **Sınıf Ağırlığı:** Ölçek hassas ve seçilen sınıflandırıcıların eğitim sürecinde sınıf ağırlığı uygulanmıştır böylece çoğunluk ve azınlık sınıfları arasındaki fark dengelenmiştir.
- **Dengesiz Hassas Metrikler:** Doğruluk, kesinlik, duyarlılık, F1-skoru ve ROC-AUC metrikleri yerine daha önemli olduğunu düşündüğümüz metrikler kullanılmıştır.
- **Topluluk Sınıflandırma Modelleri:** Topluluk sınıflandırıcıları çoklu karar sınırları kullanma yetenekleri nedeniyle kullanılmıştır.

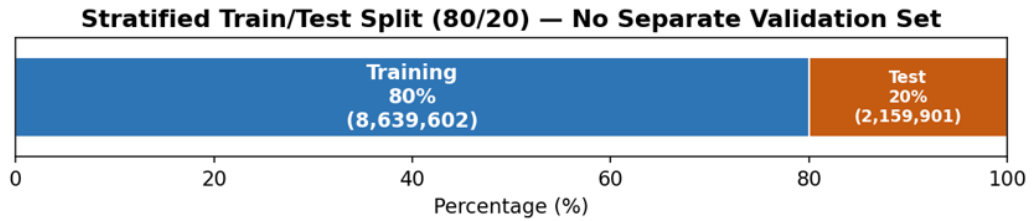
SMOTE veya rastgele alt örnekleme gibi bazı yöntemler bazı durumlarda yardımcı ola-

bilir fakat yanlış pozitifler oluşturabilir ve değerli bilgilerle birlikte ortadan kalkabilirler. Bilgi güvenliği için gerçek dünya trafiğini yansıtmak önemlidir.

3.4.4. Veri Bölme Stratejisi (Eğitim, Doğrulama, Test)

Önerilen modelin performans analizi CICIDS2017, CICIDS2018 ve UNSW-NB15 veri kümelerinin önceden işlenmiş bir kombinasyonunda yapıldı. Bu veri kümeleri 80/20 oranında (random_state = 42 ile) tren/test alt kümelerine ayrıldı. Her iki eğitim ve test setlerinde de stratifikasyon kullanılarak sınıf oranı (%89,1 zararsız ve %10,9 saldırı) korundu. Sonuç olarak eğitim alt kümesinde 8,639,602 örnek ve test alt kümesinde 2,159,901 örnek vardı.

Doğrulama olmadan bölme kasıtlı olarak yapıldı ki eğitim için mümkün olduğunca çok veri kullanılsın ve ayrı bir doğrulama alt kümesinde çoklu ayarlama nedeniyle model seçimi sorunu önlensin. Bu tezde kullanılan tüm ölçümler dokunulmamış %20 test alt kümesinde hesaplandı.



Şekil 3.2. Eğitim ve Test için Kullanılan Tabakalı Eğitim/Test Bölünmesi (80/20).

Veri Sızıntısının Önlenmesi

Veri sızıntısını önlemek için bir dizi önlem alınmıştır. Özelliklerin ölçeklendirilmesi ve dönüşümleri sadece eğitim verileri kullanılarak gerçekleştirilmiştir. Eğitim verileri üzerinde eğitilen dönüştürücüler test verilerini dönüştürmek için kullanılmıştır. Ayrıca sadece eğitim verileri kullanılarak dış değerleri sınırlamak için sınırlar da belirlenmiştir.

3.5. Makine Öğrenmesi Sınıflandırıcıları

Çalışan bir saldırı tespit sistemi oluşturmak için, mevcut araştırma; başlıca makine öğrenmesi tekniklerini yansıtan üç sınıfta gruplandırılmış aşağıdaki on bir sınıflandırıcı kümesini kullanır:

1. Grup 1 — Temel öğrenciler (5): Gaussian Naive Bayes (John & Langley, 1995); k-Nearest Neighbours (Cover & Hart, 1967), $k = 5$ ile; Linear Support Vector Machine (Cortes & Vapnik, 1995), olasılık kalibrasyonlu LinearSVC olarak uygulanmış; Decision Tree (Breiman vd., 1984), $\text{max_depth} = 20$ ve $\text{min_samples_leaf} = 10$ ile; ve Logistic Regression (Cox, 1958), $\text{class_weight} = \text{'balanced'}$ ile.
2. Grup 2 — Bagging toplulukları (3): Random Forest (Breiman, 2001), $n_estimators = 100$ ile; Bagging Classifier (Breiman, 1996), bir CART temel öğrenciyle ve $n_estimators = 50$ ile; ve Extra Trees (Geurts vd., 2006), $n_estimators = 100$ ile.
3. Grup 3 — Boosting toplulukları (3): AdaBoost (Freund & Schapire, 1997), $n_estimators = 100$ ile; Gradient Boosting (Friedman, 2001), $n_estimators = 100$, $\text{max_depth} = 6$ ve $\text{learning_rate} = 0.1$ ile; ve XGBoost (Chen & Guestrin, 2016), $n_estimators = 100$, $\text{max_depth} = 6$ ve $\text{learning_rate} = 0.1$ ile.

Yukarıda listelenen sınıflandırıcılar; karar kuralları, topluluk öğrenmesi, olasılıksal tahmin, örnek tabanlı (instance-based) öğrenme ve doğrusal sınıflandırma gibi bir dizi öğrenme paradigmasını kapsayacak biçimde seçilmiştir. Bu çeşitlilik, trafik akışlarının çok yönlü bir analizini olanaklı kılar; özellikle topluluk sınıflandırıcılarının, farklı karar sınırlarının birleşimi yoluyla güvenilirliği artırması ve aşırı uyumu azaltması beklenir.

3.5.1. Ağaç Tabanlı ve Klasik Modeller

Decision Tree (DT)

Karar Ağaçları karar kuralları ile çalışan ve bu kurallarla özellik alanını tekrarlı olarak bölen sınıflandırıcılar olarak çalışırlar. Bu özellik sayesinde karar verme süreci daha kolay anlaşılır hale gelir. DT'ler (Karar Ağaçları) yorumlanabilirliği, herhangi bir ön işleme yapmadan doğrusal olmayanlıkları ele alma yeteneği, karar verme sürecindeki şeffaflık ve verimlilikleri nedeniyle seçilmiştir. Ayrıca Karar Ağaçları aynı zamanda zayıf sınıflandırıcılar olarak da toplu modellerde kullanılırlar.

Random Forest (RF)

Rastgele Orman algoritması farklı bootstrap örnekleri ve rastgele özellik alt kümeleri üzerinde eğitilmiş Karar Ağaçlarının bir topluluğudur. Rastgele Orman algoritması tek bir

karar ağacına göre aşırı uyum sağlamaya daha az duyarlıdır ve bu nedenle daha iyi genelleme yapar. Yüksek boyutlu veriler örneğin trafik verileri gibi veriler için aşırı uyuma karşı dayanıklılığı ve gürültüye karşı dayanıklılığı ve aynı zamanda özellik önemini hesaplama yeteneği nedeniyle seçilmiştir.

3.5.2. Topluluk ve Boosting Tabanlı Modeller

Bagging: Eğitim kümesinden çıkarılan bootstrap örneklerine dayalı bağımsız temel sınıflandırıcıların eğitimi ve daha sonra tahminlerini birleştirerek istikrarlarını artırmak. Bagging aşırı uyumdan dolayı karar ağaçları gibi yöntemlerin varyansını azaltma yeteneği, gürültülü veri trafiğine karşı dayanıklılığı artırma ve karmaşık dışlayıcıları tespit etme yeteneğini geliştirmek için seçilmiştir. Burada kullanılan topluluk 50 sınırlı derinlikteki CART temel öğrencilerden oluşuyor her biri her bölünmede tüm beş özelliğe maruz kalıyor; tam yapılandırma 3.1 tablosunda sunulmuştur.

Extra Trees (Extremely Randomized Trees)

Ek Ağaçlar Rastgele Ormanlara benzer fakat ek bir rastgelelik bileşeni vardır-bir bölünme eşiği rastgele seçilir. Ek Ağaçlar gürültülü verilere karşı artan dayanıklılık, tek bir karar ağacına göre aşırı uyuma karşı azaltılmış uyum ve Rastgele Ormanlarda olduğu gibi topluluk içinde artan çeşitlilik umuduyla seçilmiştir. 5 özellik temsili için 4.1.3 bölümünde belirtildiği gibi bu umut henüz gerçekleşmemiştir; rasgele tohumlar arasındaki istikrarsızlık 4.6 bölümünde analiz edilmiştir. Yapılandırma Rastgele Orman yapılandırması gibi sınıf ağırlıklarını içermektedir; 3.1 tablosunda sunulmuştur.

AdaBoost

AdaBoost bir güçlendirme algoritmasıdır ve sınıflandırma zorluğuna göre verileri yeniden örnekleme yapar. AdaBoost azınlık saldırı sınıfına karşı artan duyarlılık elde etmek, daha hafif bir güçlendirme algoritması sağlamak ve başka bir öğrenme paradigması eklemek için seçilmiştir. Ancak 4.1.2 bölümünde belirtildiği gibi en az etkili ağaç tabanlı model olarak duyarlılık açısından en az etkili olmuştur. Bu sonuç 3.1 tablosunda belirtildiği gibi derinliği bir olan temel öğrenci kullanmasından kaynaklanmaktadır ki bu da problemin karar sınırını kapsamak için yeterli kapasiteye sahip değildir.

Gradient Boosting and XGBoost

Gradyan Güçlendirme önceki üyelerin kalıntılarını minimize etmeye çalışarak tekrarlı olarak sınıflandırıcılar oluşturur.XGBoost bir gradyan güçlendirme kütüphanesidir ve çeşitli optimizasyon algoritmaları ve paralellik uygular.XGBoost yapılandırılmış veriler üzerinde daha iyi performans,büyük veri kümelerinde yüksek verimlilik,aşırı uyumu önlemek için düzenlenme ve özellik önem analizi için seçilmiştir.İki model aynı temel ayarları paylaşmaktadır (3.1 tablosu);onlar eğitildikleri bilgisayar ayarları 3.7.1 bölümünde belirtilmiştir.

3.5.3. Olasılıksal ve Mesafe Tabanlı Modeller (Temel Modeller)

Logistic Regression

Logistik Regresyon bir lojistik fonksiyonu kullanarak olasılıkları tahmin eden bir doğrusal sınıflandırıcı olarak bilinir.Logistik Regresyon hem doğrusal hem de doğrusal olmayan sınıflandırıcılarla karşılaştırılabilecek basit bir temel model olarak seçilmiştir.

k-Nearest Neighbours (KNN)

KNN bir örnek tabanlı sınıflandırıcıdır ve özellik uzayında en yakın komşuların çoğunluk oyuna göre sınıflandırma yapar ($k=5$).Bu güçlü bir parametresiz sınıflandırıcıdır ve veri dağılımıyla ilgili herhangi bir varsayım yapmaz veparametrik ve karar ağacı sınıflandırıcılarla karşılaştırmayı sağlar.

Linear Support Vector Machine (Linear SVM)

Doğrusal SVM (LinearSVC ile olasılık ayarı kullanılmıştır) doğrusal maksimum marj hiper düzlem ayırıcısını bulmak için kullanılır. Algoritma özellikle önemli sınıf dengesizliği durumunda beş akış özelliğinde güçlü doğrusal sınıflandırıcının performansını tahmin etmek için kullanılmıştır.

Naive Bayes (GaussianNB)

Naive Bayes sınıflandırıcıları Bayes teoremine dayanan ve bağımsızlık varsayımına dayanan olasılıklı sınıflandırıcıdır.

3.5.4. Model Çeşitliliği ve Topluluk Gürbüzlüğü

Farklı türlerdeki sınıflandırıcıların kullanılması her bir modelleme yaklaşımının güçlü yönlerini kullanarak modelin sağlamlığını artırmak için faydalıdır. Ağaçlar oldukça karmaşık olmayan modeller oluşturmak için çok yeteneklidirler; Bagging ve Rastgele Orman modeli daha istikrarlı hale getirir ve modelin varyansını azaltır; Gradyan Artırma ve XGBoost modelleri zorlu gözlemler üzerine odaklanırlar; doğrusal,örnek temelli ve olasılıklı modeller (Logistik Regresyon, KNN, Doğrusal SVM ve Naif Bayes) basit karar sınırları ve verilerin diğer temsillerini elde etmemizi sağlarlar. Modellerin çeşitliliği IDS'nin performansını iyileştirmede önemli bir faktördür.

3.5.5. Topluluk Mimarisi Tasarımının Yönlendirilmesi

Temel modelin analizinden elde edilen sonuçlar hangi modellerin ve tekniklerin seçilebileceğini ve nasıl kullanılacağı konusunda yardımcı olur.Farklı sınıflandırıcıların davranışları her bir modelin güçlü ve zayıf yönlerini bize gösterir:Karar ağacı dışlayıcı değerlere ve sınıf dengesizliğine karşı dirençlidir;Bagging ve Rastgele Ormanlar gibi topluluk modelleri daha istikrarlıdır;Boosting nadir saldırıları tespit etmeye yardımcı olur ve doğrusal ve olasılıklı modeller daha basit sınıflandırma sınırları sunar.Temel model karşılaştırma tekniği en uygun modeli rastgele seçmek yerine bize kanıt sağlar ve sonuçlar topluluk modellerinin siber güvenlik verilerini nasıl ele aldığını göstermektedir.

3.6. Topluluk Tasarım Gerekçesi ve Stratejisi

Birçok teknik kısıtlama saldırgan tespit sistemlerinin çalışmasında geçerlidir. Örneğin bunlar ağ gürültüsü,veri dağıtım önyargısı,saldırıların dinamik doğası ve sınıf dengesizliği gibi unsurları içerir. Her MÖ (Çok Amaçlı) model çeşitli veri dağıtım önyargılarını algılar ancak bunların hepsi bu kısıtlamalara karşı koyamaz. Bu sorunu çözmek için mevcut araştırma topluluk odaklı bir yaklaşım üzerine kuruludur bu yaklaşım aynı test ortamı içinde çeşitli topluluk modellerini karşılaştırmayı sağlar. Yeni bir topluluk mimarisi geliştirmek (örneğin yığınlama ve karma modeller gibi)gerekmiyor önemli olan mevcut modellerin verimliliğini incelemek:

- Bootstrap Aggregating (Bagging);

- Random Forest;
- Extra Trees;
- Boosting (AdaBoost, Gradient Boosting, XGBoost).

Bu yaklaşımın amacı tüm modellerin aynı ön işleme ve verilerle eğitilmesini sağlayarak teknik tutarlılığı(adil ve tekrarlanabilirliği(aynı özellikler,bölümler ve ölçüm kümesi)sağlamak ve gerçek ortamı yansıtmaktır(çok veri kümesi senaryosu).

3.6.1. Bagging Tabanlı Öğrenme Stratejisi

Bagging veya Bootstrap Toplama birden fazla sınıflandırıcının farklı bootstrap örneklerinden eğitilmesiyle ve son tahmin için birleştirilmesiyle sağlamlığı artırır.Saldırı dağılımlarında dengesizlik durumunda sağlamlık artacaktır ve aynı zamanda ağ trafiğindeki gürültü ve heterojenlik durumunda da iyi performans gösterecektir.Çok sağlam bir sınıflandırıcı örneği rastgele özelliklerin karıştırılması ve özelliklerin ortalamasının alınmasıyla çalışan Rastgele Orman'dır.

3.6.2. Boosting Tabanlı Öğrenme Stratejisi

Boosting algoritmaları ardışık öğrenme yoluyla sınıflandırıcılar üretir. Her bir sonraki sınıflandırıcı önceki sınıflandırıcıların yaptığı hatalar üzerine eğilir. Bu nedenle artırma bu çalışmada zor tespit edilebilen istilalar için büyük önem taşıyabilir. Çalışma üç artırma modelinin düşündüğü yeteneğine odaklanıyor;önceki modellerin yaptığı yanlışları azaltmak,zor istila durumlarına karşı duyarlılığı artırmak ve yapılandırılmış verilerde olmayan bağımlılıkları ele almak.XGBoost'un büyük veri kümeleri gibi CICIDS2018 gibi veri kümelerinde istila tespitinde verimli bir sınıflandırıcı olduğu ortaya çıktı.

3.6.3. Topluluk Çeşitliliği ve Tasarım Gerekçesi

Bu çalışmada seçilen topluluk yaklaşımı sınıflandırıcıların çeşitliliğine değil karmaşıklığına vurgu yapmaktadır. Ağaç tabanlı sınıflandırıcılar,artış sınıflandırıcıları,olasılıklı sınıflandırıcılar,örnek tabanlı sınıflandırıcılar ve doğrusal sınıflandırıcılar gibi çeşitli türde sınıflandırıcıların kullanılması çözümü daha sağlam ve belirli karar sınırlarına daha az bağımlı hale getirir.Farklı saldırılara karşı genelleme yeteneğini artırır vefarklı modellerin çıktılarını karşılaştırarak yanlış pozitif sayısını azaltır ve herhangi bir derin öğrenme

teknikleri kullanmadan bir çözüm üretir.

3.6.4. Topluluk Çerçevesine Kavramsal Genel Bakış

Önerilen topluluk temelli saldırı algılama sistemi aşağıdaki dört katmandan oluşacaktır:

- **Giriş Katmanı:** Saldırı algılama veri kümesinin ön işlenmiş akışı düzeyindeki özellikleri.
- **Model Katmanı:** Ağaç,artımlı,olasılık,örnek ve doğrusal sınıflandırıcılar kullanarak eğitim ve test modelleri.
- **Değerlendirme Katmanı:** Doğruluk, kesinlik, duyarlılık, F1-Skoru ve ROC-AUC.
- **Analiz Katmanı:** Saldırı algılama probleminde sağlamlık,genelleme, istikrar ve ayrımcılık analizi.

Bu şekilde farklı makine öğrenimi yaklaşımlarıyla karşılaştırmalar yapılabilir ancak yaklaşımın şeffaflığı, ölçeklenebilirliği ve yeniden üretilebilirliği korunabilir.Topluluk seçimi saldırıların doğrusal olmayan doğasını göz önünde bulundurarak yapılır:Bagging saldırının varyansını azaltır ve artımlı yaklaşımlar bu önyargıyı azaltmaya yardımcı olur.

3.7. Deneysel Kurulum

Deneysel ayar rekabetteki algoritmalar arasında ve elde edilen sonuçlarda tekrarlanabilirlik ve adalet sağlamak üzere tasarlandı.Aynı rastgele tohumun kullanılması yoluyla eğitim-test ayrımı sabitlemek için tutarlı bir deneyim yapıldı.

3.7.1. Donanım ve Yazılım Ortamı

Deneyler, Windows 11 Pro (64-bit) çalıştıran bir Acer Predator PHN16-72 dizüstü bilgisayar üzerinde yürütülmüştür. Makine; bir Intel 13. nesil Core HX serisi işlemci (temel frekans ~2,2 GHz, 32 mantıksal çekirdek), 32 GB DDR5 bellek ve 8 GB video belleğine (CUDA 12.5) sahip bir NVIDIA GeForce RTX 4070 Laptop GPU ile donatılmıştır. XGBoost, GPU hızlandırması kullanmıştır (tree_method = 'hist', device = 'cuda'); diğer tüm modeller CPU üzerinde eğitilmiştir; mümkün olduğunda scikit-learn'de n_jobs = -1 ayarlanarak paralelleştirme uygulanmıştır. CPU süresi, n_jobs ve CUDA etkin algoritmalar için kesin olmadığından, çalışma zamanı ölçüsü olarak duvar saati (wall-clock) süresi kullanılmıştır.

Yazılım yığını, Python 3.11 ve aşağıdaki temel paketlerden oluşuyordu: temel ve topluluk sınıflandırması için scikit-learn 1.4.0, gradyan artırma için XGBoost 2.0.3, Parquet biçiminde saklanan büyük tablo verilerinin verimli işlenmesi için pandas 2.1.4 ve PyArrow 14.0.2, sayısal hesaplama için NumPy 1.26.3, model serileştirme için joblib 1.3.2 ve görselleştirme için seaborn 0.13.1 ile matplotlib 3.8.2. Dağıtım prototipi Streamlit 1.30.0 kullanılarak geliştirilmiştir.

3.7.2. Veri Kümesi Yapılandırması

Deneyisel veri kümesi Bölüm 3'te belirtilen ön işleme sürecini kullanarak oluşturulan birleştirilmiş çoklu veri kümelerinden oluşuyor. Özellikle CICIDS2017, CICIDS2018 ve UNSW-NB15 veri kümeleri birleştirilerek tek bir Parquet veri kümesi oluşturulmuştur ve bu veri seti 10.799.503 girişten oluşuyor.

Her akış girişi beş ortak özellikten (sürdürme süresi,src_bytes,dst_bytes,src_paket ve dst_paket) ve bir ikili hedef değişkeninden (Benign=0,Saldırı=1) oluşuyor. Bir ek özellik (source_dataset) yalnızca takip amaçlı kullanılıyor ve modelleme sürecinde hiçbir rolü yok. Bu veri kümesi dengesiz, 9.621.424 (%89,1) zararsız akış girişi ve 1.178.079 (%10,9) saldırı akış girişi içeriyor.

Her bölümde sınıfları dengede tutmak için sabit bir rastgele tohum (random_state=42) stratifikalı örnekleme için kullanılmış. Veri kümesi %80 eğitim ve %20 test kümeslerine bölünmüş ve doğrulama kümesi yok. Bu da 8.639.602 eğitim örneği ve 2.159.901 test örneği içerir.

3.7.3. Eğitim Yapılandırması

Tüm ağaç tabanlı sınıflandırıcılar (Karar Ağacı, Rastgele Orman, Ekstra Ağaçlar, Bagging, AdaBoost, Gradyan Artırımı ve XGBoost)özgün özellik alanıyla eğitildi çünkü ölçeklere duyarsız özellikleri vardır ve bu nedenle özellik ölçeklendirmesine ihtiyaç duymazlar.

StandartScaler yalnızca ölçek duyarlı modeller (Gaussian Naive Bayes, KNN, Doğrusal SVM ve Logistik Regresyon) için eğitim bölümüne uygulandı ve daha sonra test bölümüne veri sızıntısını önlemek için uygulandı. Çıkarımlar eğitim bölümüne uygulanan

kuantil ölçütlerine (0,1 ve 99,9 yüzdeler) göre belirlendi; çıkarımlar karşılık gelen kesme değerleriyle değiştirildi; aynı ölçütler test bölümü için de kullanıldı.

Belirtilmedikçe tüm deneyimler `random_state = 42` ile yapıldı. Tohum hem stratifiye edilmiş eğitim-test bölümlenmesi hem de her rastgele modelde (bootstrap örnekleri ve bölme) rastgeleleşmeden sorumludur; bu nedenle tüm deneyim tek bir tohum değerine bağlıdır. On bir sınıflandırıcı için tam hiperparametre kümesi Tablo 3.1’de verilmiştir; açıkça belirtilen ve varsayılan hiperparametreler arasındaki fark orada vurgulanmıştır.

Çizelge 3.1. On bir sınıflandırıcının hiperparametre yapılandırması; bu çalışmada açıkça belirlenen değerler ile kütüphane varsayılanları ayrı gösterilmiştir.

Model	Grup	Bu çalışmada açıkça belirlenen	Kütüphane varsayılanı (ayarlanmamış)
Gaussian NB	Temel	(yok — argümentsiz oluşturulmuştur)	var_smoothing=1e-9; priors=None
KNN	Temel	n_neighbors=5; algorithm='ball_tree'; n_jobs=-1	weights='uniform'; metric='minkowski'; p=2; leaf_size=30
Doğrusal SVM	Temel	CalibratedClassifierCV(cv=3) ile sarmalanmış LinearSVC(C=1.0, max_iter=2000, class_weight='balanced', random_state=42)	method='sigmoid'; ensemble=True; loss='squared_hinge'; tol=1e-4
Karar Ağacı	Temel	max_depth=20; min_samples_leaf=10; random_state=42	criterion='gini'; splitter='best'; max_features=None; class_weight=None
Logistik Regresyon	Temel	max_iter=1000; class_weight='balanced'; random_state=42	C=1.0; penalty='l2'; solver='lbfgs'; tol=1e-4
Random Forest	Bagging	n_estimators=100; max_depth=20; min_samples_leaf=10; class_weight='balanced_subsample'; random_state=42; n_jobs=-1	criterion='gini'; max_features='sqrt'; bootstrap=True
Bagging	Bagging	n_estimators=50; estimator=DecisionTreeClassifier(max_depth=20, min_samples_leaf=10, random_state=42); random_state=42; n_jobs=-1	max_samples=1.0; max_features=1.0; bootstrap=True
Extra Trees	Bagging	n_estimators=100; max_depth=20; min_samples_leaf=10; class_weight='balanced_subsample'; random_state=42; n_jobs=-1	criterion='gini'; max_features='sqrt'; bootstrap=False
AdaBoost	Boosting	n_estimators=100; random_state=42	estimator=DecisionTreeClassifier(max_depth=1); learning_rate=1.0; algorithm='SAMME.R'
Gradient Boosting	Boosting	n_estimators=100; max_depth=6; learning_rate=0.1; random_state=42	loss='log_loss'; subsample=1.0; min_samples_leaf=1; max_features=None
XGBoost	Boosting	n_estimators=100; max_depth=6; learning_rate=0.1; objective='binary:logistic'; eval_metric='logloss'; tree_method='hist'; device='auto'; random_state=42	subsample=1.0; colsample_bytree=1.0; reg_lambda=1.0; gamma=0; min_child_weight=1

Tüm diğer sütunlarda görünmeyen diğer parametreler scikit-learn 1.4.0 ve XGBoost 2.0.3 paketlerinin versiyonlarına göre varsayılan olarak bırakılmıştır. Hiçbir hiperparametre ayarı yapılmamıştır çünkü doğrulama veri kümesi yoktur(3.4.4 bölümüne bakın) ;eğer mevcut olsaydı ayar test veri kümesinde yapılacak ve bu da çalışmada sunulan ölçütleri geçersiz kılacaktı. Ağaç derinliği ve yapraklara getirilen kısıtlamalar ağaç tabanlı modellerde modelin doğruluğunu artırmak için değil de 8,6M eğitim veri kümesinde makul bir bellek kullanımı sağlamak için uygulanmıştır- herhangi bir kısıtlamadan yoksun bir ağaç kırk seviyeden daha derin olabilir ve böyle bir ağacın oluşturulması yaklaşık 15-25GB RAM gerektirir.4.1.2 bölümünde bahsedilen AdaBoost sınıflandırıcısının düşük duyarlılık skoru bu algorithmada kullanılan varsayılan derinlik değerinden kaynaklanmaktadır. Rastgele orman ve ekstra ağaç algoritmalarında her bölünmede beş özellikten ikisi seçilirken bagging’de tüm beş özellik kullanılır; bu da 4.7.5 bölümünde açıklanan farklı özellik önemine odaklanmayı açıklar. Sınıf ağırlıklandırması sadece rastgele orman ve ekstra ağaçlarda değil aynı zamanda iki doğrusal temel sınıflandırıcıda da uygulanmıştır.

3.8. Değerlendirme Ölçütleri

Uygun değerlendirme ölçütlerini seçmek son derece önemlidir.Sadece doğruluk ölçütünü hesaplamak anlamlı olmaz çünkü dengesiz bir veri kümesinde her zaman çoğunluk sınıfını(iyi sınıf) tahmin eden bir sınıflandırıcı %89 doğruluk elde edebilir fakat saldırıları tespit etmeyecektir.

3.8.1. Accuracy (Doğruluk)

Accuracy, her iki sınıftaki doğru tahminlerin genel oranını ölçer:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN) \quad (1)$$

TP, TN, FP ve FN sırasıyla gerçek pozitif, gerçek negatif, yanlış pozitif ve yanlış negatif ifade eder. Basit olmakla birlikte, bu ölçüt burada esasen bütünlük adına; yani önceki çalışmalarla bir karşılaştırma noktası olarak dâhil edilmiştir, seçimin belirleyici faktörü olarak değil.

3.8.2. Precision (Saldırı Sınıfı)

Precision, tespit edilen akıştaki saldırıların oranını yansıtır ve dolayısıyla yanlış alarm oranıyla ters orantılıdır.

$$Precision = TP / (TP + FP) \quad (2)$$

Precision önemlidir; çünkü yanlış alarmlar, ek soruşturmalar nedeniyle pahalıdır ve analistlerin kullanılan algoritmaya olan güvenini azaltır.

3.8.3. Recall (Saldırı Sınıfı)

Gerçek pozitif oranı veya duyarlılık (sensitivity) olarak da bilinen recall; modelin tespit ettiği gerçek saldırıların yüzdesi olarak tanımlanır:

$$Recall = TP / (TP + FN) \quad (3)$$

Recall kavramı belki de saldırı tespitinin en hayati yönüdür; çünkü her yanlış negatif, ciddi sonuçlara yol açabilecek gözden kaçmış bir saldırıyı temsil eder.

3.8.4. F1-Score

F1-Score, aşağıda gösterildiği gibi precision ve recall ölçütlerinin harmonik ortalaması olarak adlandırılabilir ve bu iki çelişkili ölçütü dengeler:

$$F1 = 2 \times (Precision \times Recall) / (Precision + Recall) \quad (4)$$

Harmonik ortalama, iki işlenenin küçük değerine bağlı olduğundan, F1-Score yüksek recall ancak düşük precision (veya tersi) olan algoritmaları cezalandırır. Bu nedenle F1-Score, bu bölümde karşılaştırmalar için kullanılan başlıca ölçüt hâline gelir.

3.8.5. ROC-AUC

ROC eğrisi, gerçek pozitif oranını, tüm sınıflandırma eşikleri aralığında yanlış pozitif oranına karşı çizer ve Eğri Altında Kalan Alan (ROC-AUC), bu ilişkiyi [0, 1] aralığında tek bir skaler olarak özetler. 0,5 değeri rastgele tahmini, 1,0 ise kusursuz bir sınıflandırıcıyı temsil eder. ROC-AUC eşikten bağımsız olduğundan, modelin ayırt etme gücünü ölçer.

$$AUC = \int_0^1 TPR(FPR) d(FPR) \quad (5)$$

4. BULGULAR

Bu bölümde önerilen topluluk temelli IDS sisteminin ölçekte kapsamlı bir deneysel değerlendirmesi yapılmaktadır. Deneyin yapılandırılması ve değerlendirme kriterleri daha önce Bölüm 3'ün (Alt bölümler 3.7 ve 3.8) tartışılmıştır. Bu bölüm aşağıdaki şekilde düzenlenecektir: (i) Tüm on bir modelin test kümesi üzerindeki sınıflandırma doğruluğu sonuçları; (ii) Hata matrisleri (karışıklık matrisleri), ROC grafikleri ve karşılaştırmalı grafikler aracılığıyla görsel performans analizi; (iii) Veri kümesi kullanılarak genelleme analizi; (iv) Bagging topluluk için temel öğrenici; (v) Hesaplama maliyeti analizi; (vi) Beş rastgele tohum için kararlılık analizi; ve (vii) Streamlit prototipi yardımıyla sonuçların tartışılması.

4.1. Nicel Sonuçlar

Çizelge 4.1, tüm on bir modelin 2.159.901 örneklik tutulan test kümesindeki performansını F1-score'a göre sıralanmış olarak özetlemektedir. Topluluk ve ağaç tabanlı modeller, doğrusal ve olasılıksal temel modellere açıkça üstünlük sağlamaktadır; her ölçüt sütunundaki en iyi skor, izleyen analizde belirtilmektedir.

Çizelge 4.1. On bir modelin birleştirilmiş veri kümesindeki performansı (2.159.901 test akışı), F1-score'a göre sıralanmıştır.

Model	Group	Accuracy	Precision	Recall	F1-Score	ROC-AUC
Bagging	Bagging	99,21%	96,13%	96,63%	96,38%	99,91%
Decision Tree	Baseline	99,20%	96,10%	96,60%	96,35%	99,84%
KNN	Baseline	99,13%	95,63%	96,44%	96,03%	98,83%
Gradient Boosting	Boosting	99,09%	96,16%	95,44%	95,80%	99,86%
XGBoost	Boosting	99,01%	97,68%	93,12%	95,35%	99,84%
Random Forest	Bagging	98,48%	88,67%	98,65%	93,39%	99,92%
AdaBoost	Boosting	95,90%	92,83%	67,68%	78,29%	98,11%
Extra Trees	Bagging	92,53%	63,45%	74,30%	68,44%	95,85%
Logistic Regression	Baseline	47,49%	13,77%	72,45%	23,14%	69,83%
Gaussian NB	Baseline	88,74%	6,94%	0,26%	0,50%	65,66%
SVM (Linear)	Baseline	89,10%	85,14%	0,06%	0,13%	60,05%

4.1.1. En Yüksek Performanslı Modeller: Bagging, Decision Tree ve KNN

Bagging tabanlı sınıflandırıcı en iyi performansı genel olarak gösterdi, kesinlik(96,13%) ve duyarlılık(96,63%) arasında çok güzel bir denge sağlayarak %99,21'lik bir doğruluk %96,38'lik bir F1 puanı ve 99,91'lik bir ROC-AUC ile sonuçlandı. İlginç bir nokta tek iyi yapılandırılmış bir Karar Ağacı modeli neredeyse aynı derecede iyi performans gösteriyor(96,35% F1);diğer bir KNN sınıflandırıcısı örnekler temelinde de aynı derecede iyi performans gösterdi(96,03% F1).Bu üç sınıflandırıcının performansındaki benzerlik beş boyutlu büyük veri kümesinde bulunan çoğu sinyali doğru bir şekilde yakalayabilen iyi eğitilmiş bir ağaç modeli olabileceğini gösteriyor ve kümeleme değerinin hala pozitif ama küçük olduğunu gösteriyor.

4.1.2. Boosting Modelleri: Gradient Boosting, XGBoost ve AdaBoost

Artırma teknikleri Gradient Boosting'in çok iyi performans gösterdiği alanlarda iyi performans göstermiştir (F1= 95,80%, Kesinlik=96,16%)ken XGBoost kesinlik açısından en

iyi performansı göstermiştir (97,68%) ancak düşük bir duyarlılık oranına sahiptir (93,12%) bu da onun F1 skorunu düşürmektedir (95,35%). Bu nedenle XGBoost yanlış pozitifler için cezalar uygulanan durumlar için iyi bir seçenektir. Öte yandan AdaBoost doğruluk açısından iyi performans göstermiştir (95,90%) ancak kötü bir duyarlılık oranına sahiptir (67,68%) ki bu zayıf varsayımlara sahip beş özellik üzerinde zayıf bir öğrenciyi ifade etmektedir.

4.1.3. Bagging Ailesi Modelleri: Random Forest ve Extra Trees

Random Forest tüm on bir model arasında maksimum duyarlılık değerini %98,65 ve maksimum ROC-AUC değerini %99,92 olarak elde etti; aynı zamanda nispeten ılımlı bir kesinlik değeri olan %88,67 ve F1 puanı değeri olan %93,39'u da elde etti. Modelin duyarlılık odaklı yapısı onu bir saldırıyı kesinlikle kaçırmayacağı ilk inceleme aşamasına uygun hale getiriyor ve daha sonra yanlış pozitifler filtrelenebilir. Ekstra ağaçlar ek bir bölme katmanı ekleyerek topluluğa daha düşük bir F1 puanı (%68,44) elde etti fakat bireysel olarak önemli ölçüde daha iyi performans gösterdi (Bölüm 4.3).

4.1.4. Doğrusal ve Olasılıksal Temel Modeller

Doğrusal ve olasılık temelli temel modeller birleştirilmiş derlemede en kötü F1 puanlarını alabildiler: Logistik Regresyon (%23,14), Gaussian Naive Bayes (%0,50) ve Doğrusal SVM (%0,13). Bunun nedeni açıklandığı 4.7.3 bölümündeki gibi modellerin kötü performansları modelin öğrenme içindeki zayıflıklarından daha çok birleştirilmiş veri kümesindeki çok dengesiz sınıf dağılımından kaynaklanıyor. Çünkü modeller neredeyse dengeli UNSW-NB15 veri kümesinde çok daha iyi performans gösteriyorlar. Sadece bu nedenden dolayı burada tutuluyorlar.

4.2. Görsel Performans Analizi

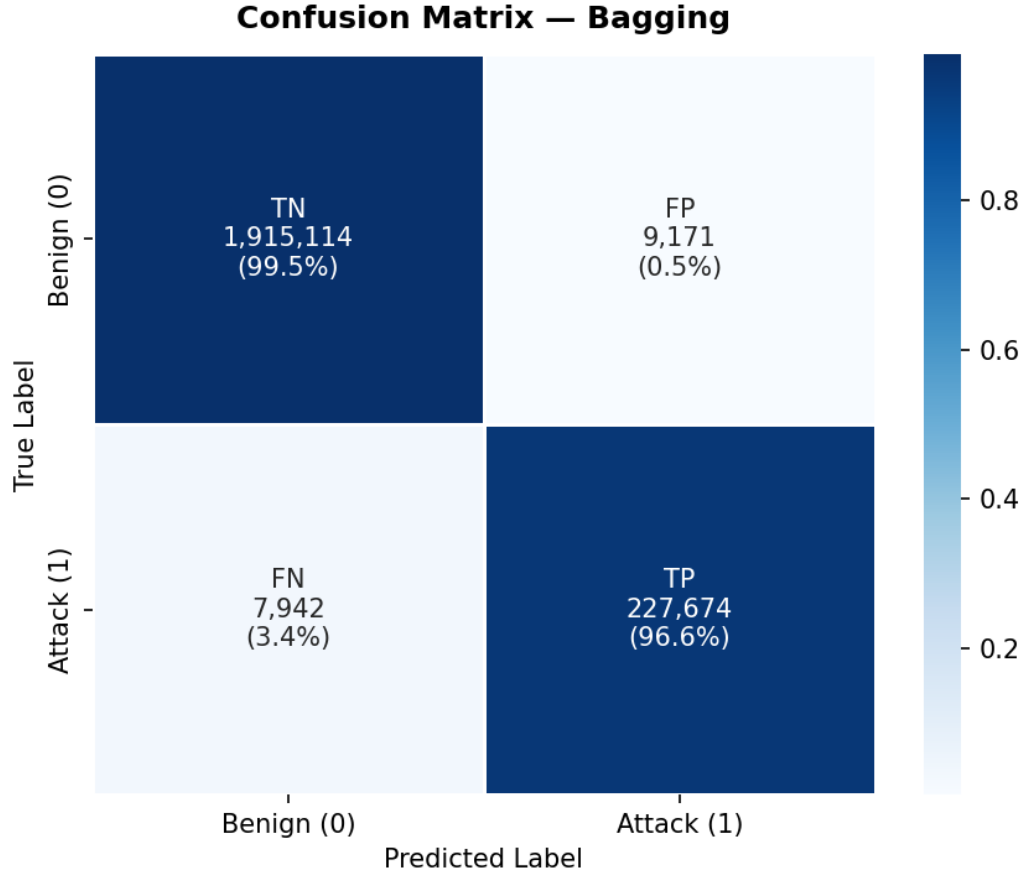
Toplu performans ölçütleri, modelin genel başarısı hakkında geniş bir bakış sağlamakta; ancak yapılan hataların niteliği hakkında bilgi vermemektedir. Bu bölümde, en başarılı modellerin karar performansının değerlendirilmesi amacıyla üç görselleştirme yöntemi tanıtılmaktadır.

4.2.1. Hata Matrisi Analizi

Bir sınıflandırıcı bir test kümesinde test edildiğinde dört olası sonuç vardır bunlar

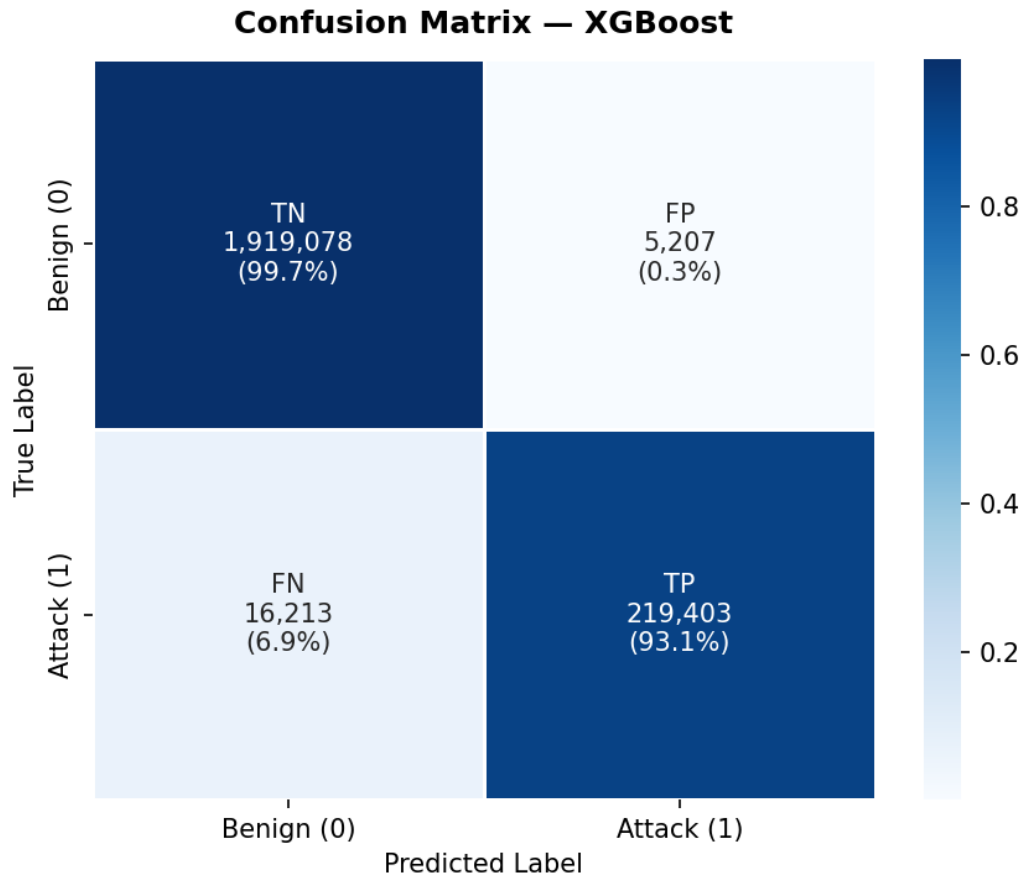
TN'ler (Gerçek Negatif),FP'ler (Yanlış Pozitif),FN'ler (Yanlış Negatif) ve TP'ler (Gerçek Pozitif) içerir. Bir istilaya karşı tespit sistemi için en sorunlu hücre FN'lerdir ki bunlar zararsız olarak yanlış sınıflandırılmış saldırıları temsil eder.

Aşağıdaki şekil (Şekil 4.1) en iyi performans gösteren sınıflandırıcının kararsızlık matrisini sunmaktadır. Test kümesindeki 235.616 saldırının 227.674'ünü (TP'ler) başarıyla sınıflandırdı ve 7.942'sini (FN'ler) kaçırarak bir duyarlılık değeri %96,63 elde etti.



Şekil 4.1. Çok veri kümeli test kümesinde Bagging sınıflandırıcısının hata matrisi.

XGBoost Sınıflandırıcısının Hata Matrisi Şekil 4.2'de gösterilmiştir. Genel doğruluk açısından bir fark yok fakat hata dağılımı oldukça farklı;yanlış pozitif sayısında yaklaşık %43'lük bir azalma (9.171'den 5.207'ye) olurken yanlış negatif sayısında iki katına çıkma (7.942'den 16.213'e) oldu.Bu kesinlikle XGBoost'un bir şeyi öğrendiğinin göstergesidir.



Şekil 4.2. Çok veri kümeli test kümesinde XGBoost sınıflandırıcısının hata matrisi.

Tüm on bir sınıflandırıcının hata matrisleri 4.3 numaralı şekilde tek bir tablo kullanarak kolay karşılaştırma için verilmiştir. Alt sol köşesi (yanlış negatifler) en belirgin şekilde AdaBoost, Gaussian NB, Lineer SVM ve Logistik Regresyon'un en iyi ağaç tabanlı sınıflandırıcılarla karşılaştırıldığında daha çok nüfuzlu olduğu görülmektedir.

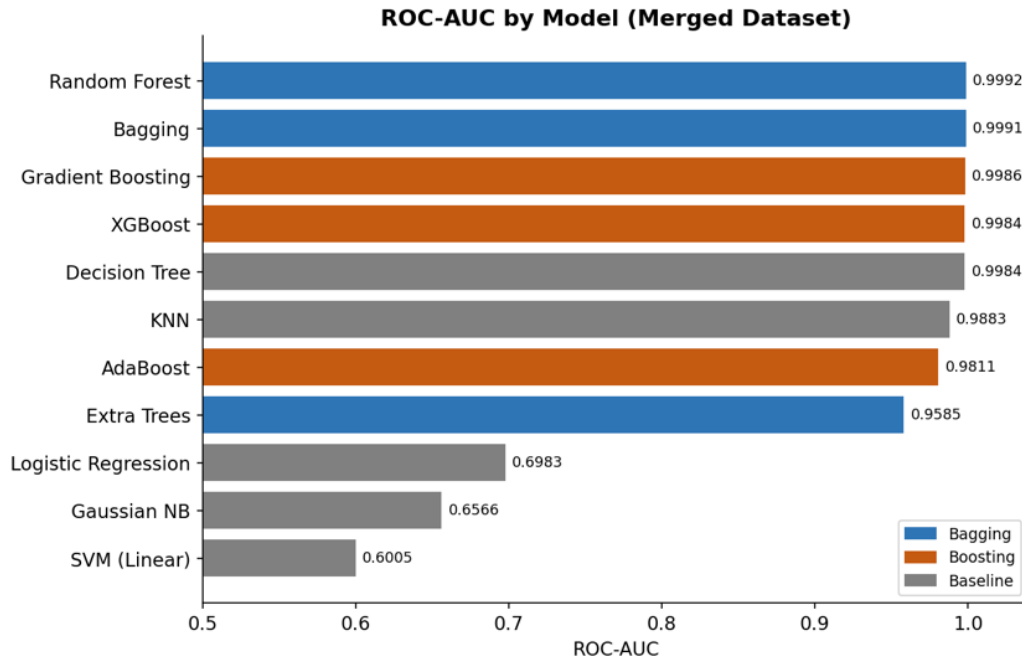
Confusion Matrices - All Eleven Classifiers (multi-dataset test set)



Şekil 4.3. Çok veri kümeli test kümesinde tüm on bir sınıflandırıcının hata matrisleri.

4.2.2. ROC Eğrisi Analizi

ROC-AUC değerleri, herhangi bir önceden belirlenmiş eşik değeri olmaksızın modellerin ayırt etme (sıralama) yeteneğini doğrulamaktadır. Ağaç tabanlı topluluklar en iyi ve birbirine en yakın sonuçları vermektedir: Rastgele Orman (%99,92), Bagging (%99,91), Gradyan Artırma (%99,86), XGBoost (%99,84) ve Karar Ağacı (%99,84) neredeyse aynı değerlere sahiptir ve akışları saldırı olasılıklarına göre nerdeyse kusursuz biçimde sıraladıklarını göstermektedir. Bunları KNN (%98,83) ve ardından AdaBoost (%98,11) izlemektedir. Doğrusal ve olasılıksal modeller belirgin biçimde daha düşük sıralama yeteneği sergilemektedir: Logistik Regresyon (%69,83), Gaussian NB (%65,66) ve SVM (Linear) (%60,05).

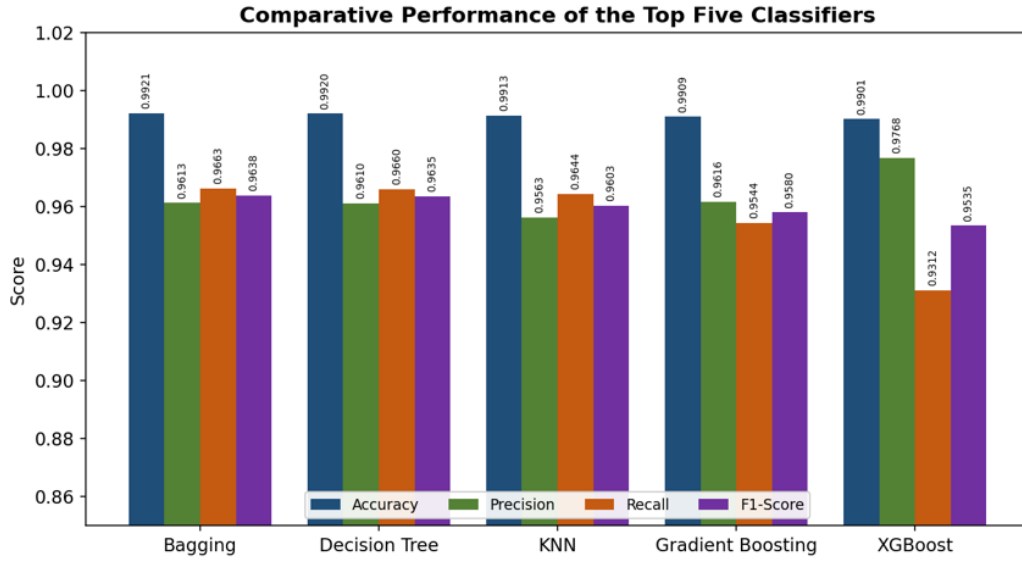


Şekil 4.4. Birleştirilmiş veri kümesinde modele göre ROC-AUC.

4.2.3. Karşılaştırmalı Performans Grafiği

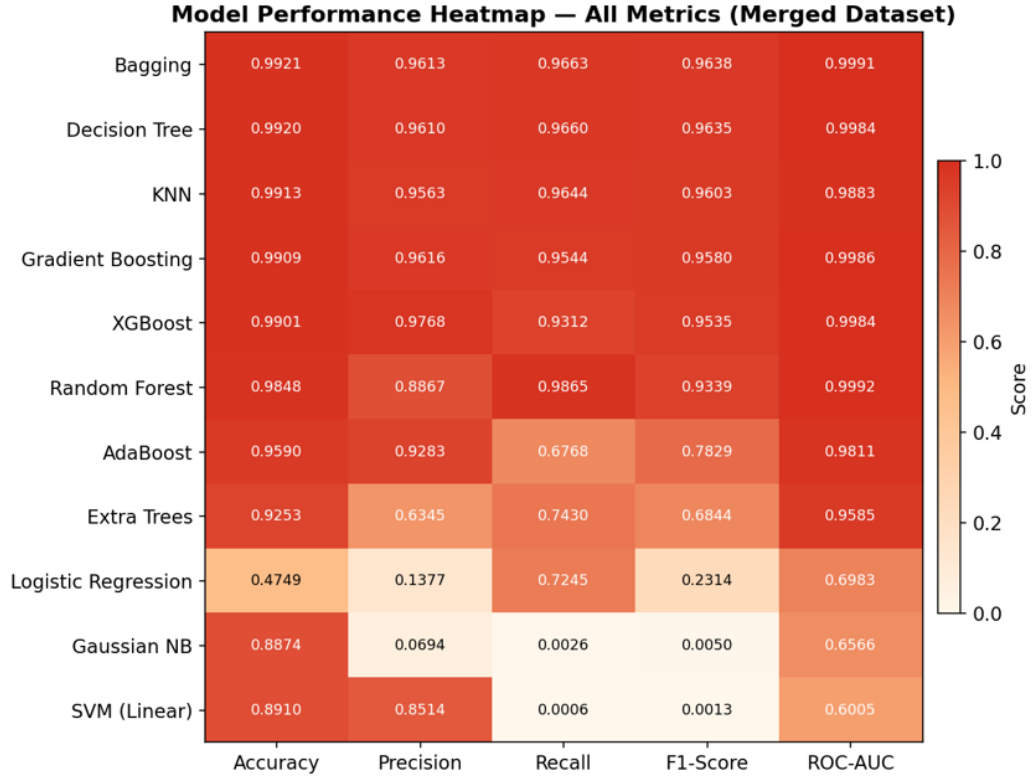
Dört önemli ölçüt yani doğruluk, kesinlik, duyarlılık ve F1-skoru en iyi beş model için şekil 4.5’de sunulmuştur. İki açık kalıp belirlenebilir; öncelikle üst sıralardaki modeller doğrulukta çok iyi performans gösteriyor ve tüm modeller 0,99’den fazla; ikincisi Bagging en yüksek dört çubuğa sahip model çünkü tüm ölçütlerde dengeli bir doğaya sahip. XGBoost

kesinlikte çok iyi performans gösteriyor fakat duyarlılıkta kötü performans gösteriyor.



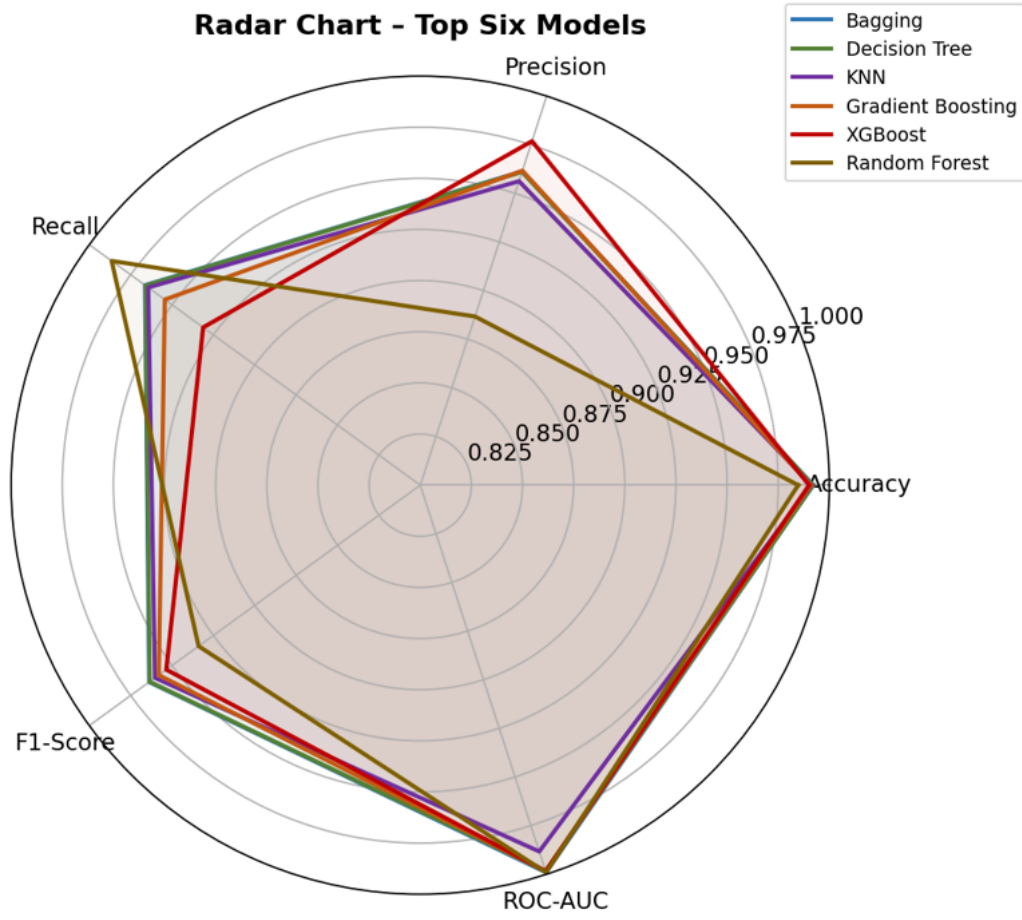
Şekil 4.5. En iyi beş sınıflandırıcı arasında accuracy, precision, recall ve F1-score karşılaştırmalı çubuk grafiği.

Beş farklı ölçütün özeti On bir model için Şekil 4.6'da gösterilmiştir ve ağaç tabanlı topluluklar ile doğrusal/olasılıklı modeller arasındaki farkı hemen ortaya koymaktadır.



Şekil 4.6. Tüm ölçütler arasında model performansı ısı haritası (birleştirilmiş veri kümesi).

Şekil 4.7 beş kriter temelinde en iyi altı algoritmanın radar grafikleridir. Bu şekil en iyi sınıflandırıcıların birbirlerine olan yakınlık düzeyini ve uzmanlıklarını (örneğin XGBoost kesinlik konusunda uzmanlaşmıştır Random Forest ise duyarlılık konusunda) göstermektedir.



Şekil 4.7. En iyi altı modelin radar grafiği.

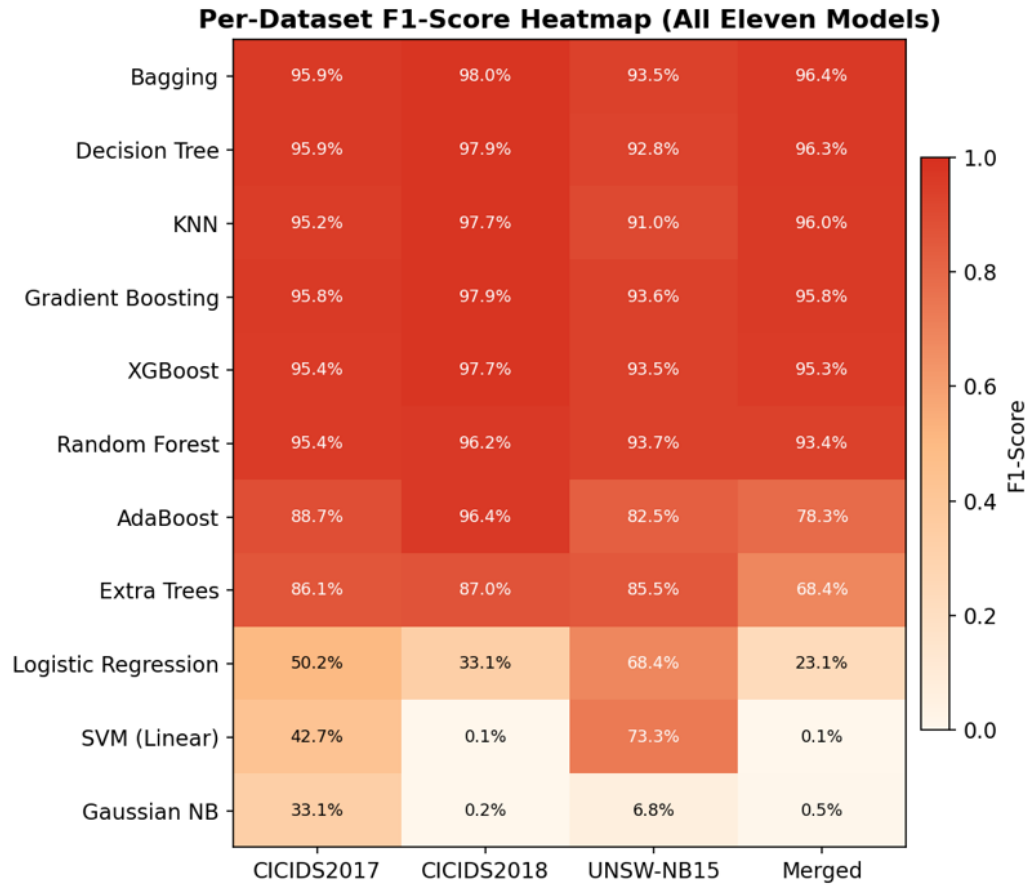
4.3. Veri Kümesi Bazlı Genellenebilirlik Analizi

Onun sınıflandırıcısının genelleme yeteneği kaynaklardan alınan her veri kümesi üzerinde sınıflandırıcının bağımsız olarak eğitilmesi ve test edilmesi ile belirlenir. Kullanılan özellikler ve 80/20'lik ayrılmış bölme sabit kalacaktır. Aşağıdaki Tablo 4.2 her modelin her veri kümesinde elde ettiği F1-skorunu gösterirken Tablo 4.3 karşılık gelen doğrulukları göstermektedir.

Çizelge 4.2. Her modelin her bir veri kümesindeki F1-score'u (birleştirilmiş sütun referans için tekrarlanmıştır).

Model	CICIDS2017	CICIDS2018	UNSW-NB15	Merged
Bagging	95,93%	98,01%	93,47%	96,38%
Decision Tree	95,92%	97,92%	92,82%	96,35%
KNN	95,24%	97,72%	90,97%	96,03%
Gradient Boosting	95,79%	97,94%	93,63%	95,80%
XGBoost	95,41%	97,69%	93,53%	95,35%
Random Forest	95,37%	96,25%	93,75%	93,39%
AdaBoost	88,67%	96,41%	82,49%	78,29%
Extra Trees	86,06%	87,03%	85,52%	68,44%
Logistic Regression	50,20%	33,05%	68,42%	23,14%
Gaussian NB	33,14%	0,22%	6,80%	0,50%
SVM (Linear)	42,69%	0,13%	73,32%	0,13%

Çevrimiçi öğrenme yöntemleri arasında Bagging hem CICIDS2017(%95,93) hem de CICIDS2018(%98,01) için en yüksek F1 skorlarını sağlasa da Rastgele Orman daha dengeli ve daha küçük boyutlardaki UNSW-NB15 veri kümesi için en iyi sonuçları(%93,75 F1) sağlamıştır. Ancak doğrusal ve olasılıklı algoritmalar arasında karşılaştırma yapıldığında bir eğilim olduğu görülmektedir. CIC veri kümeleri arasında yüksek oranda dengesiz olanlarla(Doğrusal SVM CICIDS2018 için sadece %0,13 F1 skoru alıyor)daha dengeli UNSW-NB15 veri kümesi arasında(Doğrusal SVM %73,32 F1 skoru ve Logistik Regresyon %68,42 F1 skoru) yapılan karşılaştırma bu durumu açıkça ortaya koymaktadır.



Şekil 4.8. Üç kaynak veri kümesi arasında tüm on bir model için veri kümesi bazlı F1-score ısı haritası.

Çizelge 4.3. Her modelin her bir veri kümesindeki accuracy'si (birleştirilmiş sütun referans için tekrarlanmıştır).

Model	CICIDS2017	CICIDS2018	UNSW-NB15	Merged
Bagging	98,37%	99,71%	92,97%	99,21%
Decision Tree	98,37%	99,70%	92,31%	99,20%
KNN	98,14%	99,67%	90,27%	99,13%
Gradient Boosting	98,32%	99,70%	93,17%	99,09%
XGBoost	98,17%	99,66%	93,06%	99,01%
Random Forest	98,12%	99,43%	93,31%	98,48%
AdaBoost	95,87%	99,47%	82,51%	95,90%
Extra Trees	94,70%	97,86%	84,59%	92,53%
Logistic Regression	84,56%	70,41%	64,66%	47,49%
Gaussian NB	20,60%	90,76%	45,07%	88,74%
SVM (Linear)	84,76%	92,28%	65,46%	89,10%

Çizelge 4.4. Veri kümesi başına en iyi model (F1-score'a göre) ile birleştirilmiş derlemin karşılaştırması.

Dataset	Sample Size	Best Model	Accuracy	F1-Score	ROC-AUC
CICIDS2017	2.827.876	Bagging	98,37%	95,93%	99,84%
CICIDS2018	7.889.295	Bagging	99,71%	98,01%	99,97%
UNSW-NB15	82.332	Random Forest	93,31%	93,75%	98,43%
MERGED (All 3)	10.799.503	Bagging	99,21%	96,38%	99,91%

4.4. Bagging Temel Öğrenici Ablasyonu

Çantalamayı (bagging) başarılı nedenini bulmak için BaggingClassifier'daki temel öğreniciyi 50 öğreniciyle değiştirerek diğer değişkenleri değiştirmeden bir deneme yaptık. 4.5 çizelgesinde gördüğümüz sonuçları elde ettik. Görüldüğü gibi bir karar ağacı temel öğrenici 96,38% F1-skoruyla bir topluluk oluşturmak için uygunken diğer herhangi bir temel sınıflandırıcı neredeyse sıfır F1-skoruna sahip olur. Yukarıda yapılan deneyimler bize Bagging'lamanın esasen yüksek bir önyargı/değişkenlik olan bir temel öğrenici üzerinde

kullanıldığında hiçbir sonuç vermeyeceği için bir varyans azaltma tekniği olduğunu pratik bir kanıt sağlıyor.

Çizelge 4.5. Temel öğrenicinin Bagging topluluğu üzerindeki etkisi (birleştirilmiş veri kümesi).

Temel Öğrenici	Accuracy	Precision	F1-Score	ROC-AUC
Decision Tree	99,21%	96,13%	96,38%	99,91%
Gaussian NB	88,70%	6,57%	0,51%	66,05%
Logistic Regression	89,10%	73,06%	0,12%	67,88%
Extra Tree	89,10%	100,00%	0,11%	93,00%
Decision Stump	89,09%	0,00%	0,00%	77,25%

Toplam veri kümesi boyutu; değerlendirme %20 test bölümünde yapılmıştır.

4.5. Hesaplama Maliyeti Analizi

Yalnızca belirtim kalitesi değil aynı zamanda eğitim ve test maliyetlerine bağlı olarak IDS modelinin pratik önemi de vardır. Tablo 4.6 3.7.1 bölümünde açıklandığı gibi 2.159.901 akış içeren test kümesindeki tüm akışların sınıflandırılması için modelin eğitilmesi için geçen duvar saati ve zamanını göstermektedir.

Çizelge 4.6. Birleştirilmiş veri kümesinde her modelin eğitim ve çıkarım (duvar saati) süresi.

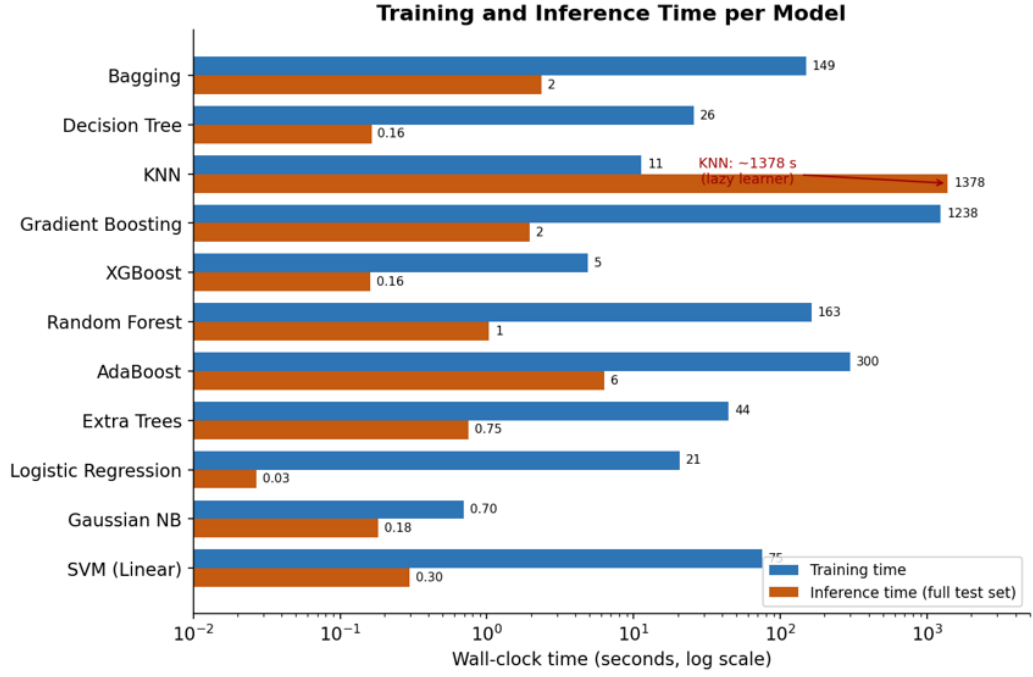
Model	Grup	Eğitim süresi (s)	Test süresi (s)	Accuracy	F1-Score
Bagging	Bagging	148,9	2,4	99,21%	96,38%
Decision Tree	Baseline	25,6	0,2	99,20%	96,35%
KNN	Baseline	11,3	1377,6	99,13%	96,03%
Gradient Boosting	Boosting	1238,4	2	99,09%	95,80%
XGBoost	Boosting	4,9	0,2	99,01%	95,35%
Random Forest	Bagging	162,7	1	98,48%	93,39%
AdaBoost	Boosting	299,5	6,3	95,90%	78,29%
Extra Trees	Bagging	44,3	0,7	92,53%	68,44%
Logistic Regression	Baseline	20,6	0,03	47,49%	23,14%
Gaussian NB	Baseline	0,7	0,2	88,74%	0,50%
SVM (Linear)	Baseline	74,8	0,3	89,10%	0,13%

Birkaç pratik içgörü elde edilebilir. Öncelikle XGBoost en hızlı algoritma olduğu kanıtlanmıştır çünkü histogram GPU uygulaması sayesinde beş saniyeden daha kısa bir sürede eğitilebilir ve test verilerini anında sınıflandırma yapabilir. Karar ağacı,Ekstra Ağaçlar,Bagging ve Rastgele Orman gibi ağaç tabanlı aileye ait algoritmalar birkaç saniye içinde tüm veri setlerini eğitebilir ve sınıflandırma yapabilir ve bu nedenle gerçek zamanlı işleme gereksinimlerini karşılayabilir. Gradient Boosting algoritmasının eğitim süresi en uzundur çünkü sıralı yapısı nedeniyle 20 dakika sürmektedir.

KNN algoritmasının en ciddi dezavantajı son derece yüksek çıkarım maliyeti olmasına rağmen algoritmanın tespit kalitesi en verimli algoritmalara yakındır(KNN'nin test veri setini sınıflandırması yaklaşık olarak 1.378 saniye sürmektedir) ; KNN tembel bir algoritmadır ve her sınıflandırma görevi için 8,6 milyon veri ögesini geçmek zorundadır. Bu nedenle KNN'nin gerçek zamanlı istila tespitinde kullanılması imkansızdır buna rağmen algoritmanın yüksek tespit kalitesi olmasına rağmen.

Bu eğitim ve çıkarım süreleri aşağıda gösterilen 4.9 numaralı grafikte logaritmik bir

ölçekte gösterilmiştir.



Şekil 4.9. Birleştirilmiş veri kümesinde model başına eğitim ve çıkarım (duvar saati) süresi (logaritmik ölçek).

4.6. Rastgele Tohumlar Arasında Kararlılık

Bildirilen sonuçların yinelenebilir olduğunu ve belirli bir rastgele bölümlemenin bir yapaylığı (artifact) olmadığını doğrulamak için, her model beş bağımsız rastgele tohum üzerinde yeniden eğitilmiş ve yeniden değerlendirilmiştir. Çizelge 4.7, beş çalıştırma üzerindeki temel ölçütlerin ortalamasını ve standart sapmasını bildirmektedir.

Kullanılan beş tohum 42, 25, 7, 13 ve 33'tür. Her durumda tohumlar tabakalı %80/%20 bölünmeyi yeniden oluşturur ve öğrencileri de tohumlar; böylece varyans, yalnızca veri bölünmesine değil, veri bölünmesine ve modelleme sürecine atfedilebilir.

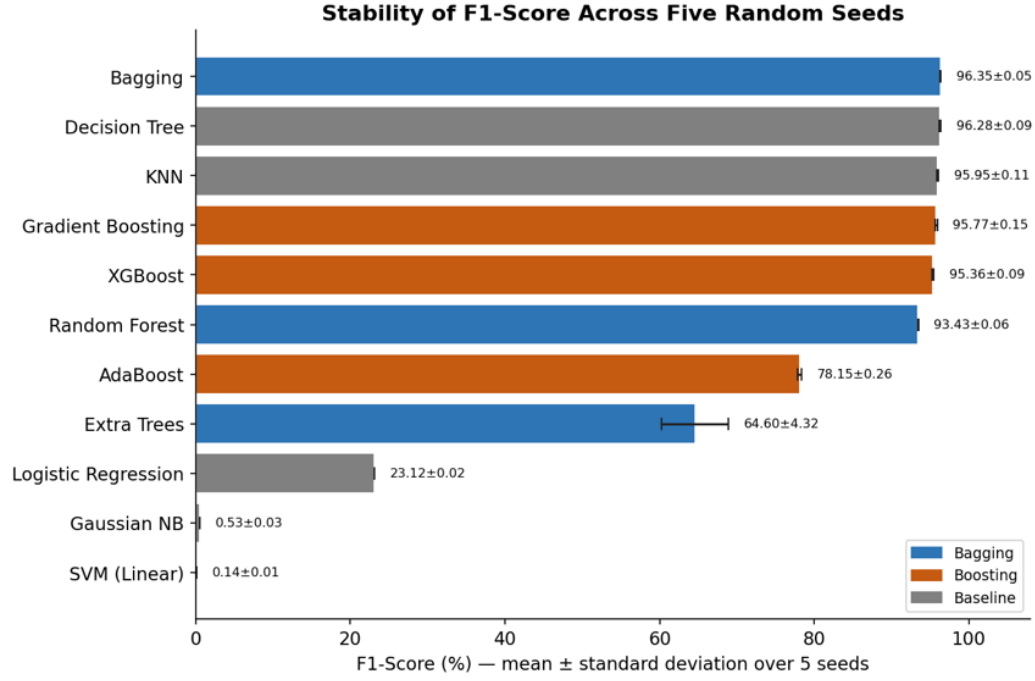
Çizelge 4.7. Beş rastgele tohum arasında temel ölçütlerin ortalaması (+)(-) standart sapması (yüzde puan cinsinden).

Model	Accuracy Ort.	Accuracy SS	F1-Score Ort.	F1-Score SS	ROC-AUC Ort.	ROC-AUC SS
Bagging	99,20	0,01	96,35	0,05	99,91	0,00
Decision Tree	99,19	0,02	96,28	0,09	99,83	0,01
KNN	99,12	0,02	95,95	0,11	98,91	0,17
Gradient Boosting	99,08	0,03	95,77	0,15	99,84	0,02
XGBoost	99,01	0,02	95,36	0,09	99,84	0,01
Random Forest	98,49	0,01	93,43	0,06	99,91	0,00
AdaBoost	95,88	0,05	78,15	0,26	98,13	0,03
Extra Trees	91,27	1,51	64,60	4,32	94,26	1,05
Logistic Regression	47,45	0,05	23,12	0,02	69,76	0,07
Gaussian NB	88,69	0,08	0,53	0,03	65,55	0,33
SVM (Linear)	89,10	0,00	0,14	0,01	60,00	0,30

Algoritmalar Bagging, Karar Ağacı, KNN, Gradient Artıřtırıcılıđı, XGBoost ve Rastgele Orman son derece yüksek bir tutarlılık gösteriyor:F1 puanlarının standart sapmaları 0,15 yüzde puanından az.Bu sebeple 4.5 numaralı řekildeki sıralamalar güvenilir ve bu deneyde seçilen rastgele tohumlardan bağımsızdır.

Ek Ağaçlar ise nispeten büyük bir varyansa sahiptir ($F1=64,60\%(+)(-)4,32$).Bu özellikle bu algoritmanın ağaç oluřturma sürecinde artan rastgelelik düzeyiyle tutarlıdır ve bu da onu bir bütün olarak güvenilirliđi azaltmaktadır.Son olarak hem doğrusal hem de olasılıksal temel modeller de yeterince güvenilir ancak daha az düzeyde.Genel olarak nispeten düşük standart sapmalar kağıtta yapılan bulguların güvenilirliđine katkıda bulunuyor özellikle de Bagging algoritmasının diđerlerine göre üstün olduđu sonucuna dair.

4.10 numaralı řekil her bir model için ortalama F1 puanının deđerini ve aynı zamanda standart sapma için hata çubuklarını gösteriyor.



Şekil 4.10. Beş rastgele tohum arasında F1-score'un kararlılığı (ortalama $\pm$ standart sapma).

4.7. Bulguların Değerlendirilmesi

Bu çalışmanın bu bölümü yapılan deneylerin sonuçlarını açıklıyor ve on bir algorithmada gözlemlenen eğilimler hakkında bazı açıklamalar veriyor. Açıklama neden Bagging algoritmasının toplu veri kümelerine uygulandığında diğerlerinden daha iyi performans gösterdiğini vererek başlıyor sonra kesinlik ve duyarlılık arasındaki takas ve bu takasın anormallik tespitine olan etkisine ilişkin bir açıklama izliyor. Sonrasında doğrusal ve olasılık temelli temel algoritmaların davranışları, modelin genelleme yeteneği ve nihayet deneyin sınırlamaları hakkında açıklamalar veriliyor.

4.7.1. Bu Görevde Bagging Neden En İyi Performansı Gösteriyor?

Bagging yöntemi ve genel olarak ağaç tabanlı yaklaşımlar bu görevde birçok tamamlayıcı nedenden dolayı iyi performans gösterirler. Öncelikle bagging'de bootstrapping yoluyla oluşturulan ensemble'in heterojenliği arasındaki etkileşim pozitif bir etkileşime sahiptir. Temel ağaç bootstrapping yoluyla oluşturulan eğitim veri kümesinin tek bir kopyasını içerirken eğitim veri kümesi kendisi üç farklı veri kümesi: CICIDS2017, CICIDS2018 ve UNSW-NB15'in akışlarının birleştirilmesiyle oluşturulmuştur.

İkincisi Bagging varyans azaltma tekniğini bir önyargı azaltma yerine kullanır. Bunun nedeni çoğunlukla paket hacmi ve zamanlamasıyla ilgili beş özellikle sınırlı olmasıdır. Bu tür bir özellik kümesi bile yeterince iyi bir derinliğe sahip bir karar ağacı tarafından yeterince iyi yakalanabilir. Şimdi hata kaynağı tek ağaçtaki yüksek varyans değil artık önyargıdır. Bagging bu kısmın üstesinden gelmek için birçok ağaç üzerinde ortalama alır; bu varyansı azaltır ama önyargıyı arttırmaz. Dolayısıyla XGBoost en doğru algoritma en yüksek doğruluk ve en yüksek kesinliğe sahip olsa da bu sadece varyans küçükse ve önyargıdan daha azsa faydalıdır. Bagging XGBoost'a (95,35%) göre daha yüksek bir F1-skoru'na (96,38%) sahiptir.

Üçüncüsü tek bir karar ağacının (F1=96,35%) , Bagging (F1=96,38%) ve KNN (F1=96,03%) neredeyse eşit doğruluğu bir topluluk etkisinden çok bilgilendirici bir etkiye işaret etmektedir. Tek bir düşük varyansa sahip düzenlenileştirilmiş karar ağacı neredeyse dokuz milyon örnek üzerinde eğitildiğinde zaten yüksek kesinliğe ulaşır ve topluluk için azaltılacak çok az varyans bırakır. Bu nedenle bu özel beş özellik ve neredeyse dokuz milyon örnek durumunda topluluk faydası çok küçüktür ancak Bagging daha iyi sonuçlar verir ve en iyi kesinlik-duyarlılık takası sağlar.

Son olarak Şekil 4.5 bu açıklamayı açıkça göstermektedir. Bir karar sapı ya da olasılıkçı temel öğrenciler üzerine kurulu bir topluluk ya da doğrusal ya da olasılıkçı temel öğrenciler kullanıldığında neredeyse sıfır F1-skoru üretir. Sadece düşük önyargılı karar ağacı temel öğrenciler etkili bir topluluk performansı sağlar; bu nedenle Bagging düşük önyargılı öğrencilerin varyansını azaltarak tespit performansını artırır.

4.7.2. IDS'te Precision-Recall Ödünleşimi

Bu modellerin kesinlik ve duyarlılık oranlarındaki bu farklılıklar gerçek hayatta bu modellerin kullanımıyla ilgili sonuçlara dönüştürülebilir. Yüksek duyarlılık oranına sahip Rastgele Orman modeli ön eleme aşamasında uygulanmalıdır çünkü yanlış negatif hataları kabul edilemez ve diğer filtreler yanlış pozitifleri yönetmek zorundadır. Yüksek kesinliğe sahip XGBoost modeli tetikleyiciye bağlı olarak otomatik olarak hareket edildiğinde ve yanlış pozitifler çok büyük maliyetlere neden olduğunda faydalı olacaktır.

4.7.3. Temel Model Davranışı

Kötü sonuçlar Logistik Regresyon, Lineer SVM ve Gaussian Naive Bayes algoritmasının birleştirilmiş veri kümesinde ortaya çıktı. Bu iki tür algoritmanın doğuştan gelen sınırlamalarını ortaya koyuyorlar. Aslında bu iki şeyin birleşimiyle ilgili bir sorun var: Sorunun karar sınırlarının doğrusal olmaması ve sınıf dengesizliği gereksinimi. Bu veri kümesinin analizi ortaya çıkardı ki neredeyse dengeli UNSW-NB15 veri kümesinde Lineer SVM %73,32'lik bir F1 puanı alabilirken Logistik Regresyon %68,42'lik bir puan aldı. Ancak birleştirilmiş veri kümesinde çok kötü performans gösterdiler sadece sırasıyla %0,13 ve %23,14 puan aldılar.

4.7.4. Değerlendirmenin Sınırlılıkları

Bu analize dair üç uyarı göz önünde bulundurulmalıdır. Birincisi, beş özellikli bir veri kümesi veri kümelerinin tutarlılığı için elbette önemlidir fakat bu analizde dahil edilmeyen çok daha değerli olabilecek birçok özellik vardır (örneğin TCP bayrakları, aralık zamanları vb.). Bu nedenle sağlanan sonuçlar mümkün olan performansın alt sınırı olarak görülebilir çünkü diğer belki de değerli özelliklerini hesaba katmamaktadır. İkincisi analizdeki her bir modelde kullanılan çoğu hiperparametrenin varsayılan değerleri vardır ve onları ayarlamak modellerin sıralamasını biraz değiştirebilir fakat önemli bir şey değiştirmez. Üçüncüsü tüm sınıflandırma algoritmaları 0,5 karar eşiğinde test edilmiştir.

4.7.5. Özellik Atıf Analizi

Karar verme süreci sınıflandırıcılar açısından bağımsız olarak ele alınmıştır. Beş ağaç tabanlı model için 20.000 test akışları (2.182 saldırı ve 17.818 zararsız) kullanılarak SHAP değerleri ile temizlik temelli önem değerlerinin hesaplanmamasından dolayı olarak değil. Scikit-learn ağaç tabanlı toplulukları olasılık alanında katkıları ele alırken XGBoost ve Gradient Boosting log-odds (log-odds) alanında ele almaktadır. Bu nedenle ham değerler karşılaştırılmaz ve her bir katkı kümesi toplam katkıya göre normalleştirilmiştir.

Önem açısından en üst üç özellik tüm modeller için aynıdır: kaynak baytlar, hedef baytlar ve akış süresi ve bunlar toplam katkının %71,8'ye %89,6'sini oluşturmaktadır. Kaynak baytlar dört modelin dördünde ilk sırada yer alırken hedef baytlar Random Forest'ta

(27,3'e karşı 27,1) biraz önde yer almaktadır. Yukarıda bahsedilen üç özellik dört modelde ilk üç sırada yer almaktadır. Gradient Boosting paket süresinde bir ön sırada yer alarak dördüncü sırada (15,7'ye karşı 15,0) yer almaktadır. Paketler tüm modellerde en zayıf özelliklerdir. Özelliklerin öneminde bağımsız algoritmalara dayalı modeller arasında benzerlik elde edilen sıralama herhangi bir belirli öğrenmenin yapaylığı değil verilerin doğal bir özelliğidir.

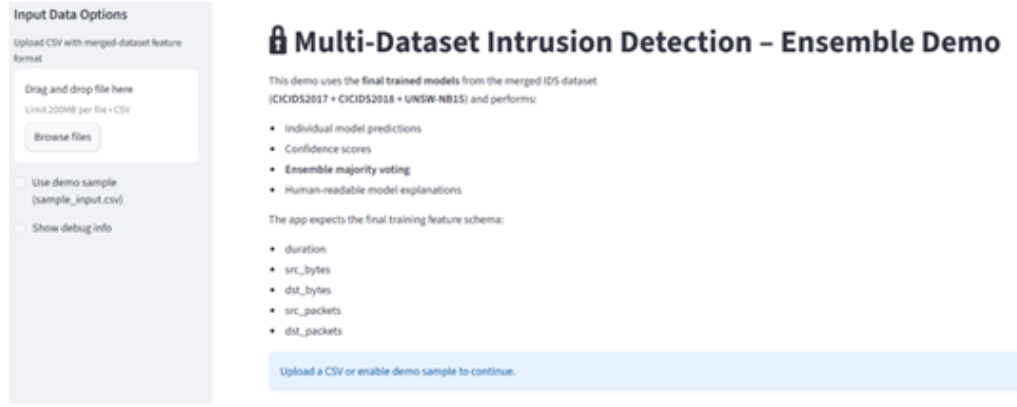
Ancak katkıların yoğunluğuna gelince modeller farklı derecede keskinlik göstermektedir. Örneğin Bagging modeli en güçlü özelliğe %40'ını, en zayıf özelliğe ise %4,5'ini ayırmaktadır bu nedenle oran 8,9'dur. Random Forest için bu oran %27,3'e %8,7 arasında değişmektedir bu nedenle oran 3,1'dir. Bu farklılıklar modellerin ayarlarına karşılık gelmektedir. Bagging için max_features parametresi 1,0 olarak ayarlanmıştır bu nedenle her temel ağaç her bölünmede en bilgilendirici tek özelliği kullanmaktadır. Random Forest varsayılan olarak karekök bootstrap örneklelemeyi kullanır bu nedenle her bölünmede beş özellikten ikisi kullanılabilir ve ağaçlar bunları değiştiren özellikleri kullanmaktadır yani paket sayısı yani bayt boyutunun ve ortalama paket boyutunun ürünü bu nedenle paket sayısı bayt sayısını değiştirmektedir.

Böylece elde edilen sonuçlar saldırının büyük hacimli ve zaman alan özelliklerine dayandığını doğrulamaktadır ve 2.4. bölümde tartışıldığı gibi kimlik gibi alanların dışlanması yönünde alınan kararı doğrulamaktadır.

4.8. Dağıtım Prototipi

Bu çerçevenin gerçekte uygulanabilir olduğunu kanıtlamak için Streamlit kullanılarak bir prototip geliştirilmiştir. Prototip tüm çıkarım boru hattının yönlerini kapsamaktadır; CSV dosyası yüklenmesinden başlayarak (uygun ise) ölçeklendirmeye ve her bir modelin tahminine ve tüm tahminlerin bir araya getirilmesine kadar.

On bir önceden eğitilmiş model joblib kullanılarak yükleniyor ve oturumun tamamında bellekte depolanıyor. Kullanıcı bir CSV dosyası yüklüyorsa veya herhangi bir örnek veri kümesinden bir satır seçiyorsa uygulama her bir modeli ayrı ayrı kullanacak, onların tahminini (mümkünse olasılıklarını) alacak ve on bir tahmin arasında ağırlıklı çoğunluk oylaması yapacaktır.



Şekil 4.11. CICIDS2017, CICIDS2018 ve UNSW-NB15 Veri Kümelerini Kullanan Topluluk Tabanlı Saldırı Tespiti için Etkileşimli Streamlit Prototipi.

Ensemble Result

Majority Vote Decision

Benign (100.0%)

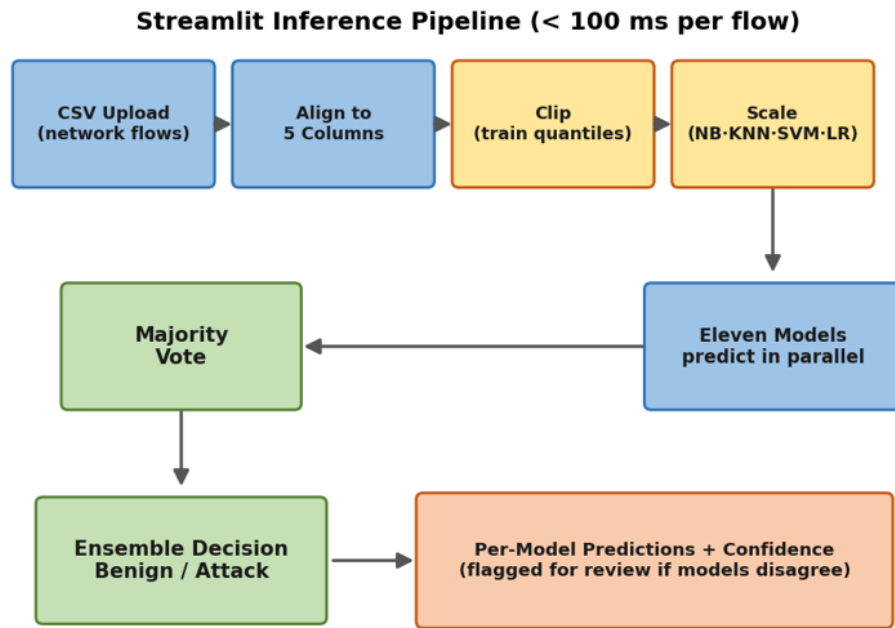
Individual Model Predictions

	Model	Prediction	Confidence	Explanation
0	DecisionTree	Benign	1.000	A simple tree-based classifier that is easy to interpret.
1	RandomForest	Benign	0.940	An ensemble of decision trees that improves generalization and robustness.
2	ExtraTrees	Benign	0.810	A highly randomized tree ensemble that is fast and often robust.
3	Bagging	Benign	1.000	Combines multiple models trained on bootstrapped samples to reduce variance.
4	AdaBoost	Benign	0.679	Boosting model that focuses more on difficult samples.
5	GradientBoosting	Benign	0.995	Sequential ensemble that corrects previous errors step by step.
6	XGBoost	Benign	0.999	A powerful gradient boosting model, strong for structured tabular data.
7	LogisticRegression	Benign	0.615	A linear classifier that works as a baseline model.
8	GaussianNB	Benign	0.525	A probabilistic model assuming Gaussian feature distributions.

Interpretation Guide

- **Prediction** → predicted class (Benign or Attack)
- **Confidence** → predicted probability when available
- **Ensemble %** → percentage of models that agreed on the final decision
- **Explanation** → short explanation of each model

Şekil 4.12. Önerilen Streamlit Tabanlı IDS Prototipinde Topluluk Tahmin Sonuçları ve Bireysel Sınıflandırıcı Çıktıları.



Şekil 4.13. Streamlit prototipi çıkarım hattı: CSV yüklemesi ve ön işlemeden on bir modelli çoğunluk oyu Benign/Attack kararına.

İki önemli karar burada vurgulanmalıdır. İlk karar ölçeklendirme sadece ölçek duyarlı sınıflandırıcıların (Gaussian Naive Bayes, k-En Yakın Komşu, Lineer SVM ve Logistik Regresyon) girdi değerlerine seçici olarak yapılacağıdır. Bu karar StandartScaler'ın örneğin eğitim aşamasında öğrenildiği ve sınıflandırıcılarla birlikte saklandığına dayanmaktadır. İkinci karar ise kesme aralıklarının kendi satırda saklanacağı ve tahmin aşamasında, dış dağılım verilerinin uzantısı durumunda patolojik davranışları önlemek için kullanılacağıdır. Prototip 100 milisaniye/kare kısıtlaması altında bitişten bitişe çıkarımı yapmayı mümkün kılıyor.

4.9. Bölüm Özeti

Bu bölüm, birleştirilmiş çok veri kümeli saldırı tespit derleminde on bir sınıflandırıcının kapsamlı bir deneysel analizini sunmuştur. Temel sonuçlar şunlardır:

- Bagging, en yüksek accuracy (%99,21) ve F1-score (%96,38) ile ve genel IDS kullanımına uygun dengeli bir precision-recall profiliyle en iyi sınıflandırıcıdır; onu yakından tek bir Decision Tree ve KNN izler.
- XGBoost en yüksek precision'ı (%97,68) ve Random Forest en yüksek recall'ı (%98,65) ile ROC-AUC'yi (%99,92) elde eder; dolayısıyla önde gelen modeller arasındaki seçim, yanlış pozitiflere karşı yanlış negatiflerin operasyonel maliyetine bağlıdır.
- Ağaç tabanlı ve topluluk modelleri, doğrusal ve olasılıksal temel modellerden belirgin biçimde daha iyi performans gösterir; ikincisi, dengeli UNSW-NB15 veri kümesindeki çok daha iyi davranışlarıyla doğrulandığı üzere, birleştirilmiş derlemde esasen şiddetli sınıf dengesizliği nedeniyle zorlanır.
- Veri kümesi bazlı deneyler, Bagging'in ağlar arasında en iyi genellemeyi yaptığını, Random Forest'in ise UNSW-NB15'te en güçlü olduğunu göstermektedir.
- Temel öğrenici ablasyonu, Bagging'in sonuçları varyans azaltma yoluyla iyileştirdiğini ve etkili olması için düşük yanlılık bir karar ağacı tabanı gerektirdiğini doğrular.

Bu sonuçlar, çalışmanın sınırlılıkları ve gelecekteki araştırmalara yönelik yönergelerle birlikte, bir sonraki bölümdeki temel bulguların formüle edilmesinde kullanılmaktadır.

5. TARTIŞMA

Bu bölüm, çalışmanın ana sonuçlarının bir özetini ve bu bulguların ağ saldırı tespiti alanındaki daha geniş çıkarımlarını sunmaktadır. Kısım 5.1’de araştırma problemi ve metodolojik katkılar yinelenirken, Kısım 5.2 ana bulguların bir özetini sunmaktadır. Kısım 5.3, deneyin sınırlılıklarına odaklanmaktadır.

5.1. Katkıların Özeti

5.1.1. Araştırma Problemi

Ağ saldırılarının tespit edilmesi Bilgisayar Güvenliği alanında ele alınması gereken önemli araştırma konularından biridir. Geleneksel teknikler,örneğin imza tabanlı teknikler ve tek başına makine öğrenimi sınıflandırıcıları gibi teknikler saldırıların dinamik doğası,değişen ağ koşulları ve günümüz ağlarının ürettiği çok büyük veri hacmi nedeniyle dezavantajlı durumdadır.Bu tezde ele alınacak temel araştırma sorunu bir makine öğrenimi tekniklerinin topluluğunun farklı veri kümeleri üzerinde ortak bir özellik alanında eğitilmesinin temel sınıflandırıcılara göre daha iyi performans gösterip göstermediği konusudur.

5.1.2. Metodolojik Katkılar

Bu tez dört önemli katkıda bulunuyor. İlk olarak çok veri kümesi deneme derlemesine dayalı bir deney tasarımı sunuyor. Bu deneyde üç iyi bilinen karşılaştırma veri kümesi olan CICIDS2017,CICIDS2018 ve UNSW-NB15,etiket normalizasyonu,özellik anlamlı hizalaması ve sadece beş akışa dayalı özelliklerden oluşan ortak bir alt küme seçilerek ön işleme yoluyla hizalanıyor. Katkı uzun süredir devam eden IDS araştırmasını engelleyen veri kümesi özgüllüğü sorunu çözmeye çalışıyor.

İkinci olarak on bir sınıflandırıcının karşılaştırmalı bir çalışmasını sunuyor. Bu sınıflandırıcılar üç kategoriye ayrılmış:Temel Sınıflandırıcılar - Gaussian Naive Bayes,K-En Yakın Komşu,Lineer Destek Vektör Makinesi,Karar Ağacı ve Logistik Regresyon; Toplama Ensemble’lar - Rastgele Orman,Bagging ve Ekstra Ağaçlar; ve Artımlı Ensemble’lar - AdaBoost,Gradient Artımlı ve XGBoost. Tüm sınıflandırıcılar aynı ön işleme ve özellik ölçeklendirme yaklaşımına göre aynı 80%/20% stratifiye bölünme üzerine eğitiliyor.

Üçüncü olarak sınıflandırıcıların veri kümesi temelli genelleme yoluyla çapraz ağ genelleme yeteneğini analiz ediyor. Her sınıflandırıcı her bir kaynak veri kümesinde eğitilip test ediliyor. Dördüncü olarak temel öğrenici atlama çalışması yapıldı. Bu çalışmanın amacı Bagging ensemble'ların yüksek performansının nedenlerini belirlemek. Ayrıca beş akışa dayalı özellik belirtimi için Streamlit arayüzünün bir prototipi geliştirildi.

5.2. Temel Ampirik Bulgular

5.2.1. Ağaç Tabanlı Modellerin Üstünlüğü

Ana empirik bulgu ağaç tabanlı ve topluluk sınıflandırıcıları tüm diğer temel ölçütlere göre beş ölçütte de daha iyi performans gösterdiği'dir. F1 puanına göre en iyi beş sınıflandırıcı -Bagging (96,38%), Karar Ağacı (96,35%), KNN (96,03%), Gradyan Artırımı (95,80%) ve XGBoost (95,35%) - hepsi ağaç tabanlı veya örnek tabanlı modellerken Logistic Regression, Gaussian Naive Bayes ve Lineer SVM kümelerinin F1 puanında performansı çok daha kötü. Bu bulgular hiperparametre ayarlaması yapılmadan elde edilmiştir.

5.2.2. Varsayılan Seçim Olarak Bagging

Tüm değerlendirilen modeller arasında Bagging en dengeli profili sergilemektedir; kesinlik (%96,13) ile duyarlılığı (%96,63) neredeyse kusursuz biçimde dengeleyerek en yüksek F1-skorunu (%96,38) elde etmektedir. Ayrıca en yüksek accuracy (%99,21) değerine de sahiptir ve ROC-AUC'si (%99,91), en yüksek değere sahip Random Forest'ın (%99,92) hemen ardından ikinci sıradadır. Bu modeli Karar Ağacı ve KNN çok yakından izlemektedir. Bir IDS'de yanlış negatif sonuçların ciddi maliyetlere yol açabilmesi nedeniyle, Bagging'in varsayılan eşik değerinde dengeli çalışması onu güvenilir bir varsayılan model hâline getirmektedir.

5.2.3. Beş Özellikli Gösterim Yeterlidir

Ağaç tabanlı modeller sadece hacim ve akış süresine bağlı beş özellik üzerinde çalışarak yüksek sonuçlar elde edebiliyorlar. Bagging sınıflandırıcısı CICIDS2017 veri kümesi için %98,37, CICIDS2018 için %99,71 ve UNSW-NB15 için %92,97 doğruluk puanlarına ulaşıyor. F1 puanları sırasıyla %95,93, %98,01 ve %93,47. Beş ağaç tabanlı sınıflandırıcı (Bagging, Karar Ağacı, Gradyan Artırma, XGBoost ve Rastgele Orman)

CICIDS2017 için %98,12-98,37, CICIDS2018 için %99,43-99,71 ve UNSW-NB15 için %92,31-93,31 doğruluk puanlarına ulaşıyorlar. Örnek tabanlı sınıflandırıcı (KNN) biraz daha kötü performans gösteriyor KNN'nin UNSW-NB15 için doğruluk puanı %90,27 (Çizelge 4.3). Bu sonuç, iyi huylu/saldırı ayırımına ilişkin bilginin önemli bir kısmının kaba hacimsel ve zamansal büyüklüklerde taşındığını göstermektedir.

Bu sonuç ayrıca 4.7.5 bölümündeki analiz sonuçlarıyla da örtüşüyor: Kaynak baytları, hedef baytları ve akış süresi tüm modellerde toplam katkının %71,8'e - %89,6'sına katkıda bulunuyorlar. Ancak bu özelliklerin "yeterliliği" koşulsuz değil: Bagging sınıflandırıcısı UNSW-NB15 için %5,4 daha düşük ve CICIDS2018 için %6,7 daha düşük doğruluk puanı alıyor. Ayrıca en iyi performans gösteren beş sınıflandırıcının F1 puanları dar bir dağılıma sahip (%95,35 - %96,38). Bu beş özellik temsiliyle ilgili bir darboğaza işaret ediyor (3.2. bölümü).

5.2.4. Temel Model Başarısızlıkları Bilgilendiricidir

Logistic Regression (F1 = %23,14), Gaussian Naive Bayes (F1 = %0,50) ve Linear SVM'in (F1 = %0,13) birleştirilmiş derlemdeki düşük performansı, genel olarak doğrusal ve olasılıksal yaklaşımlara karşı bir kanıt olarak okunmamalıdır. Veri kümesi bazlı analiz, aynı modellerin neredeyse dengeli UNSW-NB15 veri kümesinde belirgin biçimde daha iyi performans gösterdiğini ortaya koyar (Linear SVM %73,32 F1, Logistic Regression %68,42 F1). Bu nedenle davranışları, şiddetli sınıf dengesizliği ile karar sınırının doğrusal olmayan, Gauss olmayan doğasının getirdiği zorluğu betimler ve ağaç tabanlı toplulukların seçimini ampirik olarak motive eder.

5.2.5. Veri Kümesi Bazlı Genelleme

Genelleştirme deneyi veri kümesine dayalı sınıflandırıcılar için 4.3 bölümünde açıklanmıştır ve her veri kümesini aynı beş özellik temsili kullanarak ayrı ayrı işleme ve %80/%20 stratifiye bölünme içerir. Bu deneyin amacı veri kümesindeki trafik dağılımının etkisini ve sınıflandırıcının kendi etkisini ayırt etmektir. Hiçbir sınıflandırıcı üç ağda da diğerlerinden daha iyi performans göstermiyor.

Bagging hem CICIDS2017'de (95,93%) hem de CICIDS2018'de (98,01%) en iyi F1

puanlarına sahipti ve Random Forest UNSW-NB15 daha dengeli veri kümesinde %93,75'lik en iyi F1 puanını elde etti. Ancak en iyi ağaç tabanlı sınıflandırıcıların üç ağdaki puanları sadece bir puan farkla değişiyor ve Bagging her zaman en istikrarlı sınıflandırıcıdır: ya en iyi puanı alır ya da onu çok yakından takip eder. Bu istikrar birleştiriciye değil temel öğreniciye bağlıdır.

Aynı zamanda göz önünde bulundurduğumuz temsiller AdaBoost ve Extra Trees, yeterli kapasiteye sahip olmayan sınıflandırıcılar, istikrarlı değiller: UNSW-NB15'deki F1 puanları %82,49 ve %85,52'dir ki bu ağaç tabanlı sınıflandırıcıların F1 puanlarından çok daha düşüktür. Beklenildiği gibi sınıf dengesizliğinin etkisini göz önünde bulundurarak sadece yeterli kapasiteye sahip sınıflandırıcılar yani temel öğreniciler üç veri kümesinde de iyi performans gösterdi.

5.3. Çalışmanın Sınırlılıkları

Hiçbir araştırma deneyi sınırlılıklardan bağımsız değildir. Bu kısım, mevcut sonuçların geçerliliğini, kapsamını ve genellenebilirliğini etkileyen temel sınırlılıkları beş alanda kategorize ederek belirler: veri, özellikler, değerlendirme, uygulama ve kapsam.

5.3.1. Veri Sınırlılıkları

Çevrimdışı statik testleme. Modeller hiçbir veri akışı veya kavram kayması ile karşılaşmamışlardır bu yüzden performans değerleri sadece üretim ortamında bazı olası düşüşler göz önünde bulundurularak üst sınır olarak alınabilir. Sentetik saldırı trafiği. Laboratuvar ortamında önceden tanımlanmış saldırı betikleri kullanılarak üç karşılaştırma yapılmıştır bu yüzden saldırılar gerçek dünya saldırıları kadar sofistike veya uyarlanabilir olmayabilir. Sadece ikili sınıflandırma. Tüm saldırıların tek bir pozitif sınıfa sınıflandırılması veri kümeleri arasında tutarlılık sağlamak için gerekliydi ancak bu dağıtık bir IDS'nin keşif, hizmet reddi, saldırı ve veri sızdırma gibi saldırı türlerini sınıflandırma yeteneğini ortadan kaldırıyor.

5.3.2. Özellik Sınırlılıkları

Sadece beş özellik. Sadece en az beş özellik kullanılarak veri kümeleri arasında maksimum tutarlılık sağlanacaktır. Şu anda mevcut olan (CICIDS2017'de 83 özellik, CI-

CIDS2018’de 79 özellik ve UNSW-NB15’de 49 özellik) çoğu özellik göz ardı edilecek. Protokol bayrakları, yükler ve varış aralıkları gibi daha fazla özellik daha yüksek doğruluk sağlayacaktır ancak protokol bazlı özellik mühendisliği gerektireceklerdir ve bu nedenle tek tip veri kümesi olma amacına aykırı olacaktır. Hiçbir özellik etkileşimi yok. Beş özellik herhangi bir etkileşimi olmadan kullanılıyor çünkü ağaç tabanlı sınıflandırıcılar otomatik olarak özellik etkileşimleri tespit eder. Açıkça tanımlanmış özellik etkileşimleri ağaç yapılarını daha basit hale getirecektir.

5.3.3. Değerlendirme Sınırlılıkları

Sabit eşik. Sınıflandırıcılar 0,5 değerinde bir eşik değerine göre değerlendirilmiştir. Optimum F1 skoru veren veya en az bir kesinlik kriterini sağlayan eşik değerleri özellikle XGBoost gibi mükemmel AUC skorlarına sahip sınıflandırıcılar için sıralamayı göreceli olarak etkileyebilir. Seçilen hiperparametreler. Tüm hiperparametrelerin varsayılan değerleri kullanılmıştır sadece öğrenici sayısı ve birkaç diğer hariç tutulmuştur. Test verilerine modelin aşırı uyumu önlemek için; bireysel olarak ayarlanmış hiperparametreler her bir sınıflandırıcının performansını artırabilir fakat sıralamayı değiştirmesi pek olası değildir. Bırak-bir-veri-dışında çapraz doğrulama eksikliği. Modeller tüm veri kümeleri üzerinde bireysel olarak değerlendirilmiş olmasına rağmen pratik nedenlerden dolayı bırak-bir-veri-dışında çapraz doğrulama yapılmamıştır. Bu modellerin dağılım değişimine karşı dayanıklılığını değerlendirmek için çok daha iyi bir yöntem olurdu.

5.3.4. Uygulama Sınırlılıkları

Henüz gecikme üzerine bir deney yapılmadı. Streamlit demo’dan da anlaşıldığı üzere çıkarım süresi gerçekten de çok hızlı (her çıkarım için 100ms’den az) ancak yüksek verimlilik koşullarında yapılan herhangi bir gecikme ölçümü deneyleri yok; bir terminal cihazı ya da bir yönlendirici gerçek zamanlı sınıflandırma için mikrosaniyelik bir gecikmeye ihtiyaç duyabilir. Sistemimize karşı herhangi bir kötü amaçlı saldırı yoktu. Test setinde sadece gerçek veriler vardı. Küçük değişikliklerle sınıflandırıcıdan yanlış çıktılar alıracak şekilde tasarlanmış olan kötü amaçlı örnekler kullanılmadı.

5.3.5. Kapsam Sınırlılıkları

Sadece ağ trafiği verileri. Bu belirli model sadece ağ akışlarının özelliklerini kullanıyor ve sistem günlükleri, kullanıcı bilgileri veya protokol düzeyinde veriler gibi başka herhangi bir tür veriyi içermiyor. Bu bir maliyet duyarlı öğrenme yaklaşımı değil. Tüm yanlış sınıflandırma hataları gerçekte bir saldırıyı kaçırmaktan çok daha pahalı olduğu için eşit olarak ele alınıyor.

6. SONUÇLAR

Bu çalışma, çeşitli veri kümelerini birleştiren bir derlem üzerinde eğitildiğinde topluluk ve ağaç tabanlı makine öğrenmesi modellerinin güçlü saldırı tespiti performansı sağladığını göstermektedir. CICIDS2017, CICIDS2018 ve UNSW-NB15'ten alınan 10,8 milyon akıştan oluşan birleştirilmiş bir veri kümesinde Bagging sınıflandırıcısı en iyi sonuçları elde etmiştir —%99,21 accuracy, %96,38 F1-score ve %99,91 ROC-AUC— onu yakından tek bir Decision Tree ve KNN izlemiştir.

6.1. Bu Tezin Başarıları

Bu araştırmanın en önemli katkısı farklı veri kümeleri üzerindeki saldırı algılama modellerinin performansını değerlendirme ve karşılaştırma yöntemi oluşturmaktır. Çalışma üç farklı veri kümesi olan CICIDS2017, CICIDS2018 ve UNSW-NB15'i beş akış temelli özellik üzerinden tek bir ikili sınıflandırma problemi haline getirerek bu şekilde önceki IDS çalışmalarında farklı veri kümelerinin kullanımından kaynaklanan önyargıları ortadan kaldırmayı mümkün kılmaktadır.

Deney sonuçlarına göre topluluk ve ağaç temelli makine öğrenimi algoritmaları tüm değerlendirme ölçütlerinde doğrusal ve olasılıklı karşılıklarının önüne geçmektedir. Özellikle en etkili modeller olan Bagging, Karar Ağacı, K-En Yakın Komşu, Gradient Artırma ve XGBoost oldukça yüksek F1 puanlarına sahiptir.

Bu başarının parametre ayarlaması olmadan elde edildiğini belirtmek gerekir yani sadece hiperparametre optimizasyonu ile açıklanamaz. Ayrıca önerilen algoritmanın Streamlit üzerinden bir prototipi de mevcuttur.

6.2. Temel İlgörü: Bagging Neden XGBoost'tan Daha İyi Performans Gösterir?

Sonuç olarak bu çalışma, Bagging ve ağaç tabanlı modellerin bu görevde neden çok iyi performans gösterdiğini açıklamaktadır. Yalnızca beş özellik düzeyi bulunduğundan yalnlık-varyans dengesinde varyans baskındır; yani performans iyileştirmesinin anahtarı varyans azaltımıdır. Bagging bunu bir topluluk oluşturarak gerçekleştirir; topluluktaki temel ağaçların her biri, üç farklı kaynak veri kümesinin birleşiminden alınan bootstrap örnekleri üzerinde eğitilir. İşte bu nedenle Bagging en yüksek F1-skorunu (%96,38) sağlarken,

XGBoost en yüksek kesinliği (precision, %97,68) sunmaktadır.

Temel öğrenici ablasyonu, Bagging nasıl çalıştığını anlamamızı sağlar: Bagging etkili olması için düşük yanlılıkları bir karar ağacı öğrenicisine ihtiyaç vardır; karar kütüğü (stump), doğrusal model ve olasılıksal model neredeyse hiç fark yaratmaz. Ayrıca Bagging’laştırma (F1: %96,38), tek bir Karar Ağacı (F1: %96,35) ve KNN (F1: %96,03) neredeyse aynı sonuçları vermektedir; yani beş boyutlu büyük bir veri kümesinde bile iyi yapılandırılmış tek bir ağaç zaten yeterince düşük bir varyansa sahiptir; dolayısıyla topluluk oluşturma katkıları küçük olmakla birlikte yine de olumludur. Bu deneyden IDS tasarımına yönelik pratik çıkarım şudur: en karmaşık modeli seçmek yerine, model seçiminde hem yanlış pozitiflerin hem de yanlış negatiflerin operasyonel maliyeti birlikte göz önünde bulundurulmalıdır.

Random Forest en yüksek duyarlılık (recall) ve ROC-AUC değerlerini sunarken, XGBoost en yüksek kesinliği (precision) elde etmekte; Bagging ise bu ölçütler arasında en iyi dengeyi sağlamaktadır.

6.3. Uygulamaya Yönelik Çıkarımlar

Bazı IDS tasarımcıları için öneriler elde edilen araştırmanın sonuçlarına dayanabilir. Doğruluk ve duyarlılık arasında uygun bir uzlaşma sağlamak için Bagging güçlü bir varsayılan sınıflandırıcı olarak kabul edilebilir. Zaman ve bayt/paket sayıları dahil olmak üzere beş özellik ikili sınıflandırma için yeterli olup bu da yeni özelliklerin geliştirilmesinin veri kümesi uyumluluğu ve azaltılmış hesaplama maliyetleri kadar etkili olmayacağı anlamına gelir. Sistem performansı 0,5 değerinden daha etkili olan bir ayırma eşiği ayarlamasıyla tutma yöntemi kullanılarak geliştirilebilir.

6.4. Çözümlememiş Kalanlar

Sonuçların geçerliliğini sınırlayan iki ana kısıtlama vardır. İlk olarak tüm deneysel verilerin akışı herhangi bir etkiye maruz kalmadan ve kontrol edilmiş bir ortamda işlemelidir; bu pratik performansın sınırlandırılması olarak görülebilir. İkinci olarak kullanılan ikili problemin doğasında bir kısıtlama vardır çünkü veri kümelerinin birleştirilmesi tüm saldırıları aynı pozitif sınıfa atamayı gerektirir.

6.5. İleriye Giden Yol

Dört yön, üretime hazırlığa açık bir yol sunar:

- Bir-veri-kümesini-dışarıda-bırakma (leave-one-dataset-out) değerlendirmesi; iki veri kümesi üzerinde eğitim ve üçüncüsü üzerinde test yaparak çapraz veri kümesi genellemesini doğrudan nicelemek.
- Hiyerarşik (çok sınıflı) sınıflandırma: Bir Benign'e-karşı-Attack aşaması ardından koşullandırılmış bir saldırı türü sınıflandırıcısı; ilk aşamada recall'ı korurken saldırıya özgü yanıtları olanaklı kılmak.
- Üretim benzeri koşullar altında gecikme, verim ve kavram kaymasına toleransı ölçmek için gerçek zamanlı akış (streaming) değerlendirmesi.
- SHAP veya LIME ile açıklanabilirlik; akış başına kararları ve genel özellik önemini güvenlik analistlerine saydam kılmak ve beş ortak özneliliğin ötesinde özellik kümesini genişletmek.

6.6. Son Söz

Bu çalışmada topluluk temelli öğrenmenin büyük veri kümesi durumunda bir istila tespitinde güçlü, pratik ve çok etkili bir yaklaşım olduğu gösterilmiştir. Bagging sınıflandırıcısından mükemmel ve eşit dağılmış sonuçlar, tek başına çalışan Karar Ağacı ve K-En Yakın Komşu algoritmalarından benzer sonuçlar, ayrıca Random Forest ve XGBoost algoritmalarının ek avantajları birlikte bu en iyi istila tespit algoritmasını bulmanın belirli koşullarda kesinlik-duyarlılık takas sorunu göz önünde bulundurularak özel bir maliyet fonksiyonu ile bir mühendislik zorluğu olduğunu kanıtlıyor.

Önerilen yaklaşım, empirik değerlendirme ve oluşturulan prototip elde edilen mükemmel sonuçları karmaşık yaklaşımların, derin öğrenme, hiperparametre optimizasyonu ve veri kümelerinden manuel özellik çıkarma gibi kullanmadan mümkün kılmaktadır. Yukarıda bahsedilen istila tespitindeki sorunlar mevcut bilgi ve teknoloji kullanılarak çözülebilecek gibi görünüyor.

7. ÖNERİLER

Elde edilen sonuçlara dayanarak topluluk temelli saldırı tespit sistemlerinin tasarım ve uygulanması konusunda uygulayıcılara, araştırmacılara ve kurumlara bazı önerilerde bulunabiliriz.

Uygulayıcılar açısından önerimiz Bagging'i varsayılan ikili sınıflandırıcı olarak kullanmak çünkü en yüksek F1-skoruna (XGBoost'un 95,35'ine karşı 96,38%) sahiptir ve kesinlik ve duyarlılık arasında iyi bir denge sağlar. Bu sınıflandırıcı yanlış negatiflerin mümkün olmadığı durumlarda seçilmelidir. Belirli uygulama gereksinimlerini göz önünde bulundurarak karar eşiklerinde deneyler yapmak için kalibre edilmiş bir test veri kümesi gereklidir. Beş akış düzeyli özellik, iki CIC veri kümesinde %98'in, UNSW-NB15'te ise %92'nin üzerinde accuracy sağlamaktadır; bu nedenle ikili sınıflandırma için kaba hacimsel ve zamansal ölçütler pratik bir başlangıç noktası oluşturmaktadır. Yanlış pozitiflerin çok pahalıya mal olabileceği durumlarda XGBoost en yüksek kesinliğe sahip olduğu için alternatif bir sınıflandırıcı olarak kullanılabilir. En yüksek duyarlılık ve ROC-AUC'yi sahip olan Rastgele Orman sınıflandırıcısı ön yüksek duyarlılık filtrelemesi için uygulanmalıdır.

Sistem performansının kavram drifti durumunda sürekli izlenmesi ve karşılaştırılması gerekir. Gerçek yaşam verilerinde kesinlik, duyarlılık ve F1-skorunu takip ederek ve düşüş durumunda uyarılar göndererek. Yüksek veri hacmi durumunda uygulayıcılar topluluk ağacı yöntemlerinin hızlı dağıtım, model sıkıştırma ve donanım optimizasyonu gerektirebileceğini ancak hafif oldukları için göz önünde bulundurmamalıdır. SHAP ve LIME gibi açıklama araçları ağaç topluluklarının birçok öğrencisi ile yorumlanabilirliklerine rağmen makine öğrenimini analistlerin iş akışına dahil etmelerine yardımcı olacaktır.

Araştırmacılar için ilk adım farklı veri kümeleri arasında genelleme geçerliliğini doğrulamak için "bir veri kümesi çıkar" yaklaşımı ile çapraz veri kümesi genellemesini doğrulamaktır. İki veri kümesi üzerinde öğrenmek ve bir üçüncü veri kümesi üzerinde test etmek kavram drifti durumunda karar sınırlarının genellemesini sağlar. Bir sonraki pratik olarak önemli adım çok sınıflı bir saldırı taksonomisine geçmektir. Yüksek duyarlılık ile ikili algılama ve pozitif koşullu çok sınıflı sınıflandırmanın birleştirilmesi hiyerarşik sınıflandırmada etkili olabilir. Ancak çevrimdışı değerlendirme gecikme ve ölçeklenebilirlik

garantisi vermez bu nedenle araştırmacılar bu parametreleri üretim verileri üzerinde tahmin etmelidir.

Karşıt güçlülük dayanıklılığı başka bir araştırma alanıdır. Şu anda modellerin varsayımı değiştirilemez düşmanlardır ancak saldırganların kaçınma saldırıları sorunlu olabilir çünkü ağaç tabanlı modeller gürültüye aşırı uyum sağlama eğilimindedir. Gelecekteki araştırmalar bu tür karşıt eğitim, girdi temizleme ve bileşenin çeşitliliği gibi zayıflıkları ve önlemleri üzerinde odaklanmalıdır. Gelişmiş ve veri kümesine bağımsız özellik mühendisliği alanında daha fazla araştırma yapılmalıdır: Paket başına bayt oranı, yönlü simetri, zaman özellikleri ve logaritmik sayımlar.

Son olarak kurumlar yüksek kaliteli ve temsili veri kümeleri edinerek öğrenme ve test için odaklanmalıdır. Laboratuvar veri kümeleri temel değer kazanmaları açısından önemlidir ancak kurumlar üretim trafiğinden kendi örneklerini almalıdırlar. Makine öğrenimi saldırı tespit sistemleri sürekli bakım gerektirir: Yeni tehditleri tespit etmek için modelleri yeniden eğitmek, üretimde doğrulama yapmak, görüntüleri etiketlemek ve eğitim veri kümesini geliştirmek, kesinlik ve duyarlılık eşiklerini tanımlamak. Siber güvenlik uzmanları ve makine öğrenimi mühendislerinden oluşan çok disiplinli bir ekip şarttır ve mevcut altyapı ve SIEM sistemlerinde kontrollü pilot dağıtım da gereklidir.

KAYNAKLAR

- Alkahtani, H. K., Asiri, M. M., Alanazi, R., Alshahrani, M. M., Alzahrani, F., Sorour, S., & Duhayyim, M. A. (2025). Leveraging ensemble learning with metaheuristic optimization algorithms for an intelligent cyberattack defense framework in an IoT environment. *Alexandria Engineering Journal*, 129, 103–116.
- Amarudin, Ferdiana, R., & Widyawan. (2022). Performance of intrusion detection system using bagging ensemble with SDN-base classifier. *2022 International Conference on Information Technology and Digital Applications (ICITDA)*.
- Atefinia, R., & Ahmadi, M. (2021). Network intrusion detection using multi-architectural modular deep neural network. *The Journal of Supercomputing*, 77(4), 3571–3593.
- Bacha, N. U. D., Lu, S., Rehman, A. U., Idrees, M., Ghadi, Y. Y., & Alahmadi, T. J. (2024). Deploying hybrid ensemble machine learning techniques for effective cross-site scripting (XSS) attack detection. *Computers, Materials and Continua*, 81(1), 707–748.
- Basnet, R. B., Shash, R., Johnson, C., Walgren, L., & Doleck, T. (2019). Towards detecting and classifying network intrusion traffic using deep learning frameworks. *Journal of Internet Services and Information Security*, 9(4), 1–17.
- Breiman, L. (1996). Bagging predictors. *Machine Learning*, 24(2), 123–140. <https://doi.org/10.1007/BF00058655>
- Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32. <https://doi.org/10.1023/A:1010933404324>
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (1984). *Classification and Regression Trees*. Wadsworth International Group.
- Brenner, S. W. (2010). *Cybercrime: Criminal Threats from Cyberspace*. Praeger.
- Catillo, M., Rak, M., & Villano, U. (2020). 2L-ZED-IDS: A two-level anomaly detector for multiple attack classes. *Workshops of the International Conference on Advanced Information Networking and Applications*, 687–696.
- Chadza, T., Kyriakopoulos, K. G., & Lambbotharan, S. (2019). Contemporary sequential network attacks prediction using hidden Markov model. *2019 17th International Conference on Privacy, Security and Trust (PST)*, 1–3.
- Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. <https://doi.org/10.1145/2939672.2939785>
- Choi, E., Kim, H., Kim, J., & Shim, M. (2020). CNN-based network intrusion detection against denial-of-service attacks. *Electronics*, 9(6), 916.
- Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20(3), 273–297. <https://doi.org/10.1007/BF00994018>

- Cover, T. M., & Hart, P. E. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21–27. <https://doi.org/10.1109/TIT.1967.1053964>
- Cox, D. R. (1958). The regression analysis of binary sequences. *Journal of the Royal Statistical Society: Series B (Methodological)*, 20(2), 215–242. <https://doi.org/10.1111/j.2517-6161.1958.tb00292.x>
- de Lima Filho, F. S., Silveira, F. A. F., de Medeiros Brito Junior, A., Vargas-Solar, G., & Silveira, L. F. (2019). Smart detection: An online approach for DoS/DDoS attack detection using machine learning. *Security and Communication Networks*, 2019.
- Demir, O., Karatas, G., & Sahingoz, O. K. (2020). Increasing the performance of machine learning-based IDSs on an imbalanced and up-to-date dataset. *IEEE Access*, 8, 32150–32162.
- D’hooge, L., Wauters, T., Volckaert, B., & Turck, F. D. (2020). Inter-dataset generalization strength of supervised machine learning methods for intrusion detection. *Journal of Information Security and Applications*, 54.
- Feng, X., Feng, Y., & Dawam, E. S. (2020). Artificial intelligence cyber security strategy. *2020 IEEE International Conference on Dependable, Autonomic and Secure Computing, International Conference on Pervasive Intelligence and Computing, International Conference on Cloud and Big Data Computing, International Conference on Cyber Science and Technology Congress (DASC/PiCom/CBD-Com/CyberSciTech)*, 328–333.
- Ferrag, M. A., Maglaras, L., Moschoyiannis, S., & Janicke, H. (2020). Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. *Journal of Information Security and Applications*, 50.
- Fitni, Q. R. S., & Ramli, K. (2020). Implementation of ensemble learning and feature selection for performance improvements in anomaly-based intrusion detection systems. *2020 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*, 118–124.
- Freund, Y., & Schapire, R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 55(1), 119–139.
- Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *The Annals of Statistics*, 29(5), 1189–1232.
- Gamage, S., & Samarabandu, J. (2020). Deep learning methods in network intrusion detection: A survey and an objective comparison. *Journal of Network and Computer Applications*, 169.
- Geurts, P., Ernst, D., & Wehenkel, L. (2006). Extremely randomized trees. *Machine Learning*, 63(1), 3–42. <https://doi.org/10.1007/s10994-006-6226-1>

- Hua, Y. (2020). An efficient traffic classification scheme using embedded feature selection and LightGBM. *2020 Information Communication Technologies Conference (ICTC)*, 125–130.
- Ilhan Firat Kilincer, Ertam, F., & Sengur, A. (2022). A comprehensive intrusion detection framework using boosting algorithms. *Computers and Electrical Engineering*, 100.
- John, G. H., & Langley, P. (1995). Estimating continuous distributions in Bayesian classifiers. *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence (UAI)*, 338–345.
- Kaissar, A., Nassif, A. B., Soudan, B., & Injadat, M. (2025). Enhancing CNN-based network intrusion detection through hyperparameter optimization. *Intelligent Systems with Applications*, 26.
- Kanimozhi, V., & Jacob, T. P. (2019a). Artificial intelligence-based network intrusion detection with hyper-parameter optimization tuning on the realistic cyber dataset CSE-CIC-IDS2018 using cloud computing. *2019 International Conference on Communication and Signal Processing (ICCSP)*, 33–36.
- Kanimozhi, V., & Jacob, T. P. (2019b). Calibration of various optimized machine learning classifiers in network intrusion detection system on the realistic cyber dataset CSE-CIC-IDS2018 using cloud computing. *International Journal of Engineering Applied Sciences and Technology*, 4(6).
- Kanimozhi, V., & Jacob, T. P. (2021). Artificial intelligence outflanks all other machine learning classifiers in network intrusion detection system on the realistic cyber dataset CSE-CIC-IDS2018 using cloud computing. *ICT Express*, 7(3), 366–370.
- Khan, S., Ahmed, U., Jiangbin, Z., & Sadiq, M. T. (2025). Consensus hybrid ensemble machine learning for intrusion detection with explainable AI. *Journal of Network and Computer Applications*, 235.
- Kurniabudi, Purnama, B., Sharipuddin, Darmawijoyo, Stiawan, D., Samsuryadi, Heryanto, A., & Budiarto, R. (2019). Network anomaly detection research: A survey. *Indonesian Journal of Electrical Engineering and Informatics (IJEI)*, 7(1).
- Leevy, J. L., & Khoshgoftaar, T. M. (2020). A survey and analysis of intrusion detection models based on CSE-CIC-IDS2018 big data. *Journal of Big Data*, 7(1).
- Lin, P., Ye, K., & Xu, C.-Z. (2019). Dynamic network anomaly detection system by using deep learning techniques. *International Conference on Cloud Computing*, 161–176.
- Liu, K., Wu, X., Jin, Z., Chen, X., & Zhou, J. (2025). Boosting incremental intrusion detection system with adversarial samples. *Expert Systems with Applications*, 271. <https://doi.org/10.1016/j.eswa.2025.126632>
- Loevenich, J. F., Adler, E., Mercier, R., Velazquez, A., & Lopes, R. R. F. (2024). Design of an autonomous cyber defence agent using hybrid AI models. *2024 International*

- Conference on Military Communication and Information Systems (ICMCIS)*.
<https://doi.org/10.1109/ICMCIS61231.2024.10540988>
- Mohamed, H., Hefny, H., & Alsawy, A. (2018). Intrusion detection system using machine learning approaches. *Egyptian Computer Science Journal*, 42(3).
- Moustafa, N., & Slay, J. (2015). UNSW-NB15: A comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). *2015 Military Communications and Information Systems Conference (MilCIS)*, 1–6. <https://doi.org/10.1109/MilCIS.2015.7348942>
- Ramos, K. S. H., Monge, M. A. S., & Vidal, J. M. (2020). Benchmark-based reference model for evaluating botnet detection tools driven by traffic-flow analytics. *Sensors*, 20(16).
- Russell, S., & Norvig, P. (2021). *Artificial Intelligence: A Modern Approach* (Fourth Edition, Global Edition, tech. rep.). Pearson.
- Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. *Proceedings of the 4th International Conference on Information Systems Security and Privacy (ICISSP)*, 108–116. <https://doi.org/10.5220/0006639801080116>
- Sharafaldin, I., Lashkari, A. H., Hakak, S., & Ghorbani, A. A. (2019). Developing Realistic Distributed Denial of Service (DDoS) Attack Dataset and Taxonomy. *2019 International Carnahan Conference on Security Technology (ICCST)*, 1–8. <https://doi.org/10.1109/CCST.2019.8888419>
- Sibai, F. N. (2024). *AI crimes: A classification* (tech. rep.). Tech. rep.
- Venkatraman, S., Alazab, M., & Vinayakumar, R. (2019). A hybrid deep learning image-based analysis for effective malware detection. *Journal of Information Security and Applications*, 47, 377–389. <https://doi.org/10.1016/j.jisa.2019.06.006>
- Winston, P. H. (1992). *Artificial Intelligence*. Addison-Wesley.
- Wu, L., Li, X., Zhang, Q., & Chen, W. (2020). Building auto-encoder intrusion detection system based on random forest feature selection. *Computers & Security*, 95.
- Zhang, H., Zhao, F., Peng, J., Zhuang, X., & Na, S.-G. (2020). A semi-self-taught network intrusion detection system. *Neural Computing and Applications*, 32, 17169–17179.

ÖZ GEÇMİŞ

KİŞİSEL BİLGİLER

Ad Soyadı : MUHAMMED ELHASAN
Öğrenci No : 235121003
E-posta : 235121003@ogrenci.harran.edu.tr
Telefon : 05453130300
Doğum Tarihi : 1999-06-23
ORCID : 0009-0003-3122-7646
ABD (TR) : Bilgisayar Mühendisliği Anabilim Dalı
ABD (EN) : Department of Computer Engineering
:

T.C
YÜKSEKÖĞRETİM KURULU
ULUSAL TEZ MERKEZİ

TEZ VERİ GİRİŞ FORMU

Referans No	10824829
Yazar Adı / Soyadı	MUHAMMED ELHASAN
Orcid	0009-0003-3122-7646
T.C.Kimlik No	15033083036
Telefon	5453130300
E-Posta	mhmad2010122@gmail.com
Tezin Dili	Türkçe
Tezin Özgün Adı	Saldırı Tespit Sistemi için Veri Kümesi Birleştirme ile Topluluk Makine Öğrenmesi
Tezin Tercümesi	Ensemble Machine Learning with Dataset Integration for Intrusion Detection System
Konu	Bilgisayar Mühendisliği Bilimleri-Bilgisayar ve Kontrol = Computer Engineering and Computer Science and Control
Üniversite	HARRAN ÜNİVERSİTESİ
Enstitü / Hastane	FEN BİLİMLERİ ENSTİTÜSÜ
Ana Bilim Dalı	BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
Bilim Dalı	Bilgisayar Mühendisliği Bilim Dalı
Tez Türü	Yüksek Lisans
Yılı	2026
Sayfa	94
Tez Danışmanları	PROF. DR. İBRAHİM BERKAN AYDİLEK
Dizin Terimleri	
Önerilen Dizin Terimleri	

01.09.2026

İmza:.....