INO TECH
Inspiring Technologies and Innovations

Inspiring Technologies and Innovations

December 2025, Volume: 4 Issue: 2

Research Article

Kurumsal Ağlarda İç Ağdan Gelen Atak ve Tehditlerin Makine Öğrenimi Yöntemleri ile Sınıflandırılması ve Anomali Tespiti

Sercan AKÇALI[a], Oğuzhan KENDİRLİ[b]

[a]Düzce Üniversitesi, Fen Bilimleri Enstitüsü, Siber Güvenlik Anabilim Dalı, Düzce, Türkiye
[b]Düzce Üniversitesi, Dr. Engin Pak Cumayeri Meslek Yüksekokulu, Elektronik ve Otomasyon Bölümü, Düzce, Türkiye

ORCID[a]: 0009-0004-2799-0822
ORCID[b]: 0000-0001-7134-2196

Sorumlu Yazar e-mail: akcalisercan@gmail.com

https://doi.org/10.5281/zenodo.17972846

Received	: 29.05.2025	Accepted	: 08.07.2025	Pages	: 33-43

ÖZET: Kurumsal ağlar hem iç hem de dış ağdan gelebilecek siber tehditlere karşı hassas bir yapıdadır. Günümüzün dijital dünyasında, kurumlar her geçen gün daha fazla sayıda ağ bağlantısına ve cihazlara sahip olmakta, bu durum beraberinde çeşitli güvenlik tehditlerini getirmektedir. Kurumsal ağlarda çoğu zaman odak noktası dış kaynaklı siber saldırılar ve tehditler olurken, kurum içinden gelebilecek saldırılar göz ardı edilebilmektedir. Oysaki, kurum iç ağında meydana gelebilecek bir güvenlik ihlali, en az kurum dış ağından gelen bir saldırı kadar yıkıcı etkilere, maddi kayıp ve itibar zedelenmesine neden olabilir. Kurumların hem merkez hem de ona bağlı diğer lokasyonlarında kullanılan Router, Switch, Printer, CCTV (Close Circuit TeleVision) cihazları ve kullanıcı cihazları gibi çok sayıda uç nokta, ağ güvenliğini karmaşık hale getirmektedir. Ayrıca hem yerel kullanıcıların hem de misafirlerin bağlandığı kurum kablosuz ağları, pandemi sonrası yaygınlaşan uzaktan çalışma ile birlikte VPN (Virtual Private Network) bağlantısıyla kurum kaynaklarına uzaktan erişim ihtiyacının artması, IPSec (Internet Protocol Security) bağlantısı üzerinden gerçekleşen kurumlar arası bağlantılar potansiyel güvenlik açıklarını artırmaktadır. Bu çalışmada, temsili kurumsal bir ağ topolojisi tasarlanmış olup kurum için güvenilir kabul edilen ağ bağlantıları üzerinden bilinçli veya bilinçsiz bir şekilde gerçekleştirilen atak ve tehditlerin makine öğrenmesi teknikleri kullanılarak tespit ve sınıflandırması incelenmektedir. Bu çalışma, kurumların kendi güvenliklerini sağlama konusundaki farkındalıklarını artırmayı ve iç ağ güvenliğini sağlamada daha proaktif bir yaklaşım geliştirmelerini hedeflemektedir. Çalışmamızda, Next Generation Firewall sisteminden alınmış gerçek zamanlı Vulnerability, Antivirus, Spyware threat loglarına dair veri seti kullanılmıştır. Bu veri seti, Random Forest, Logistic Regression, XGBoost (Extreme Gradient Boosting), MLP (Multi Layer Preception), SVC (Support Vector Classification) makine öğrenmesi algoritmaları kullanarak doğruluk, F1 skoru, kesinlik, duyarlılık kriterleri ile değerlendirilmiştir. Sonuç olarak, XGBoost algoritması, en yüksek doğruluk oranı %99,17, F1 skoru %97,89, kesinlik %97,11 ve duyarlılık %98,69 değerleriyle genel olarak en iyi performansı sergilemiştir. Yanlış pozitifleri minimize etme açısından MLP algoritması en yüksek kesinlik oranına %98,84 ulaşmıştır.

ANAHTAR KELİMELER: IPSec, Kablosuz Ağ, Kurumsal Ağlar, Makine Öğrenmesi, Saldırı Tespit Sistemi, VPN

CLASSIFICATION AND ANOMALY DETECTION OF ATTACKS AND THREATS COMING FROM THE INTERNAL NETWORK IN CORPORATE NETWORKS USING MACHINE LEARNING METHODS

ABSTRACT: Corporate networks are inherently vulnerable to cyber threats that may originate from both internal and external sources.In today's digital world, organizations are increasingly connected through a growing number of network connections and devices, which introduces various security threats. While the primary focus in corporate networks is often on cyberattacks and threats from external sources, threats that may arise from within the organization can sometimes be overlooked. However, a security breach occurring within the internal network of an organization can be just as devastating causing financial losses and reputational damage as one originating from outside the network.Numerous endpoints such as routers, switches, printers, CCTV (Closed-Circuit Television) devices, and user devices used both at the organization's headquarters and at affiliated locations make network security more complex. In addition, corporate wireless networks accessed by both local users and guests, the increasing need for remote access to corporate resources via VPN (Virtual Private Network) due to the rise of remote work after the pandemic, and inter-organizational connections via IPSec (Internet Protocol Security) all contribute to a growing number of potential security vulnerabilities.In this study, a representative corporate network topology was designed, and attacks or threats whether intentional or unintentional through network connections considered secure within the organization were examined using machine learning techniques for detection and classification. The aim of this study is to raise awareness among organizations regarding their internal network security and to encourage a more proactive approach to securing internal systems.Our study utilized a real-time dataset consisting of Vulnerability, Antivirus, and Spyware threat logs obtained from a Next Generation Firewall system. This dataset was evaluated using machine learning algorithms such as Random Forest, Logistic Regression, XGBoost (Extreme Gradient Boosting), MLP (Multi-Layer Perceptron), and SVC (Support Vector Classification) based on accuracy, F1 score, precision, and recall metrics. As a result, the XGBoost algorithm demonstrated the best overall performance, achieving the highest values in terms of accuracy 99.17%, F1 score 97.89%,

precision 97.11%, and recall 98.69%. The MLP algorithm achieved the highest precision rate 98.84% in terms of minimizing false positives.

KEYWORDS: Corporate Networks, Intrusion Detection System, IPSec, Machine Learning, VPN, Wireless Network

1. GİRİŞ

Kurumsal siber güvenlik, özel ve kamu kurumların verilerini, sistemlerini ve operasyonel sürekliliklerini koruyabilmeleri için son derece önemli bir gereklilik haline gelmiştir. Kurumlar, hassas verileri işlediği için, güvenlik stratejileri oluşturmak ve bu stratejileri sürekli geliştirmek büyük önem taşır.

Kurumlar, siber güvenlik alanında çeşitli ve karmaşık tehditlerle karşı karşıya kalmaktadır. Fidye yazılımları, kötü amaçlı yazılımlar ve oltalama saldırıları gibi tehditler hem veri güvenliğini tehlikeye atmakta hem de iş süreçlerinin kesintiye uğramasına yol açabilmektedir. Bunun yanı sıra, kurumsal ağa uzaktan bağlantı sırasında meydana gelebilecek güvenlik ihlalleri ve insan kaynaklı hata faktörleri de önemli risk unsurları arasında yer almaktadır.

2008 yılında Ernst & Young denetim firması tarafından Türkiye'nin de dahil olduğu 50 ülkede 1,400 kuruluşun katılımıyla gerçekleştirilen küresel bilgi güvenliği anketi, bilgi güvenliği önlemlerinin kurumsal itibar üzerinde belirleyici rol oynadığını göstermektedir. Araştırmaya katılanların %85'i, bilgi güvenliği ihlallerinin kurumsal marka imajına ve itibarına zarar verebileceğini belirtirken %72'si bu tür ihlallerin doğrudan gelir kaybına neden olduğunu vurgulamıştır [1]. Kurumsal ağlarda meydana gelebilecek tehditlere yönelik literatür taraması incelendiğinde; Karakaya, çalışmasında kurumların karşılaştığı siber saldırılar ve bu saldırılara ait senaryoları detaylı bir şekilde ele almıştır. Çalışmasında laboratuvar ortamlarında kullanılan çeşitli yardımcı araçlar aracılığıyla, OWASP (Open Web Application Security Project) tarafından 2021 yılında yayımlanan en sık karşılaşılan 10 zafiyet listesine dayalı olarak, örnek atak senaryoları üzerinde testler gerçekleştirmiştir [2].

Bakhareva ve diğerleri, çalışmasında ağ trafiğinin analizine dayalı olarak kurumsal ağlardaki saldırıları tespit etmek için algoritmalar önermiştir. CICIDS2017 veri seti, ikili sınıflandırma (saldırı veya normal trafik) için makine öğrenmesi yöntemlerini karşılaştırmanın yanı sıra DoS (Denial of Service), Port Tarama, Brute Force, Web Attack, Bot ve Sızma gibi tipik saldırıların sınıflarını tanımlamak için çok sınıflı sınıflandırma için kullanmışlardır. Deneyin bir sonucu olarak, CatBoost ve LightGBM algoritmalarının, kötü amaçlı trafiğin çeşitli saldırı gruplarına hem ikili sınıflandırması hem de çok sınıflı sınıflandırması için iyi çalıştığını tespit etmişlerdir [3].

Chikkoppa ve diğerleri, çalışmasında kötü amaçlı yazılımları tanımlamak ve tespit etmek için makine öğrenmesi tekniklerini kullanarak ağın korunmasına yönelik çalışma yapmıştır. CICIDS-2017 veri setini kullanarak, ilgili özellikleri seçerek ve veri setini ağırlık gibi özelliklerine bağlı olarak farklı sınıflara ayırarak kötü amaçlı yazılım algılama tekniklerini göstermektedir. Ayrıca Naif Bayes modelleri, destek vektör algoritmaları, rastgele ormanlar ve karar ağaçları gibi sınıflandırma teknikleri kullanmışlardır. Bu sistemlerin doğruluğu sırasıyla %72,96, %96, %99,67 ve %99,59 olarak belirlediler. Araştırma sonucunda, çeşitli makine öğrenmesi teknikleri arasında rastgele ormanların en yüksek doğruluğu elde ettiği sonucunu elde etmişlerdir [4].

VLL Thing, çalışmasında IEEE 802.11 ağını hedef alan tehditleri ve saldırıları analiz etmiş ve ayrıca özellikle saldırıların yeni olduğu ve daha önce tespit ve sınıflandırma sistemi tarafından hiç karşılaşılmamış olduğu durumlarda doğru tehdit ve saldırı sınıflandırmasına ulaşmanın zorluklarını belirlemiştir. Derin öğrenme yaklaşımı kullanarak anormallik tespiti ve sınıflandırmasına dayalı bir çözüm önermiştir. Sınıflandırmayı çok sınıflı bir problem (yani meşru trafik, flood tipi saldırılar, enjeksiyon tipi saldırılar ve kimliğe bürünme tipi saldırılar) olarak ele almış ve önerilen çözümle saldırıları sınıflandırmada genel olarak %98,6688'lik bir doğruluk elde etmiştir [5].

Yaşar ve Çakır, çalışmasında artan ve giderek karmaşıklaşan siber tehditlere karşı kurumların alabileceği önlemleri kapsamlı bir çerçeve içerisinde sunmuşlardır. Ülkemizdeki kurum ve kuruluşların siber tehditlere karşı güvenliklerini artırmalarına yardımcı olmayı amaçladıkları çalışmada, görüşme ve doküman analizi gibi veri toplama tekniklerini kullanmışlardır. Çalışma sonucunda, kurumların siber güvenlik konusunda yeterli farkındalığa sahip olmadıkları ve tehditlerin etki ve zararlarının giderek arttığı tespit edilmiştir. Ayrıca, verinin gizliliği, bütünlüğü ve erişilebilirliğinin korunmasını hedefleyen kurumsal siber güvenliğin, alınacak temel önlemlerle sağlamlaştırılması ve bu sürecin dinamik bir şekilde sürekli olarak güncellenmesi gerektiği sonucuna varmışlardır [6].

Şentürk ve Irmak, çalışmalarında Microsoft'un geliştirdiği ve organizasyonlar için kritik öneme sahip olan Windows Aktif Dizin Etki Alanı Servisi'ni ele almışlardır. Çalışmada, temsili bir kurumsal ağda Microsoft ürünü olan Powershell kullanılarak gerçekleştirilen bir etki alanı sızma testi detaylandırılmıştır. Bu testte, yetkisiz bir saldırganın, yetkisiz bir son kullanıcı bilgisayarında oturum açarak, etki alanı denetleyicisine erişim sağlama adımları izlenmiştir. Ayrıca örnek bir kurumsal ağ ortamında Windows Aktif Dizin Etki Alanı Servisi üzerinden, yetkisiz bir kullanıcının kötü niyetli eylemlerini gerçekleştirebileceği saldırılar uygulamalı olarak incelenmiştir. Elde edilen bulgular, kurum içindeki personelin kabuk katmanına erişebilmesinin ciddi güvenlik riskleri oluşturduğunu ortaya koymuştur [7].

Atay ve arkadaşları, CSE-CIC-IDS2018 veri kümesi üzerinde siber saldırı tespiti amacıyla yaptıkları çalışmada tek ve iki seviyeli iki ana hibrit yöntem belirlemişlerdir. Çalışmada, çeşitli makine öğrenmesi teknikleri kullanılarak veri kümesi üzerinde analiz yapılmıştır. Bunlar arasında Evrişimsel Sinir Ağı, Rastgele Orman, Hafif Gradyan Artırma ve bunların

birleşimlerinden oluşan hibrit modeller yer almaktadır. Özellikle, Evrişimsel Sinir Ağı ve Rastgele Orman hibrit modelinin %98 doğruluk oranı ve 0,86 macro F1 skoru ile en başarılı sonuçları verdiği ortaya konmuştur. Ayrıca, çalışmalarında GridSearch yöntemiyle hiper parametre optimizasyonu gerçekleştirilmiş ve Sentetik Azınlık Aşırı Örnekleme Tekniği ile yüksek korelasyona sahip özniteliklerin tespit üzerindeki etkileri araştırılmıştır [8].

Hacıbeyoğlu ve arkadaşları, çalışmasında akıllı bir saldırı tespit sistemi tasarlamak amacıyla makine öğrenmesi ve derin öğrenme yöntemlerini incelemişlerdir. Bu tasarım için literatürde yaygın olarak kullanılan iki veri setini, NSL-KDD ve Kyoto 2006+ kullanılmıştır. Makine öğrenmesi yöntemleri, WEKA veri madenciliği aracındaki sınıflandırma algoritmaları kullanılarak uygulanmış ve elde edilen sınıflandırma sonuçları, tasarlanan derin öğrenme modeli ile karşılaştırılmıştır. Bu şekilde, akıllı bir saldırı tespit sistemi için makine öğrenmesi ve derin öğrenme yöntemleri, her iki veri seti üzerinde kapsamlı bir şekilde analiz edilmiş ve gelecekteki çalışmalar için önerilerde bulunulmuştur [9].

Kılınç ve Can, çalışmalarında honeypot teknolojisi ve farklı honeypot türlerini açıklamış ve bir durum çalışması sunmuşlardır. Bu durum çalışmasında, bir kurumsal ağa yerleştirilmiş olan Tpot servisi üzerinden elde edilen çıktılar detaylı bir şekilde incelenmiştir. Çalışma ayrıca, bu çıktılar ile güvenlik araçlarının nasıl entegre edilebileceğini ve güvenlik sağlamak için nasıl bir kaynak oluşturabileceğini de ele almaktadır [10].

Ulus ve Demirci, çalışmalarında kurum içi saldırıların temel özelliklerini ve geçmişteki saldırı örneklerini incelemişlerdir. En yaygın kullanılan saldırı yöntemlerine dair izlenen şüpheli hareketleri tespit etmek amacıyla bir Sunucu İzleme ve Takip Sistemi geliştirmişlerdir. Bu sistemden elde edilen bulgular ve istatistiksel veriler, kurumların risk düzeyleri hakkında bilgi sahibi olmalarını sağlamanın yanı sıra, saldırılara karşı hazırlıklarını güçlendirerek potansiyel saldırganları tespit etme yeteneklerini artıracağını vurgulamışlardır [11].

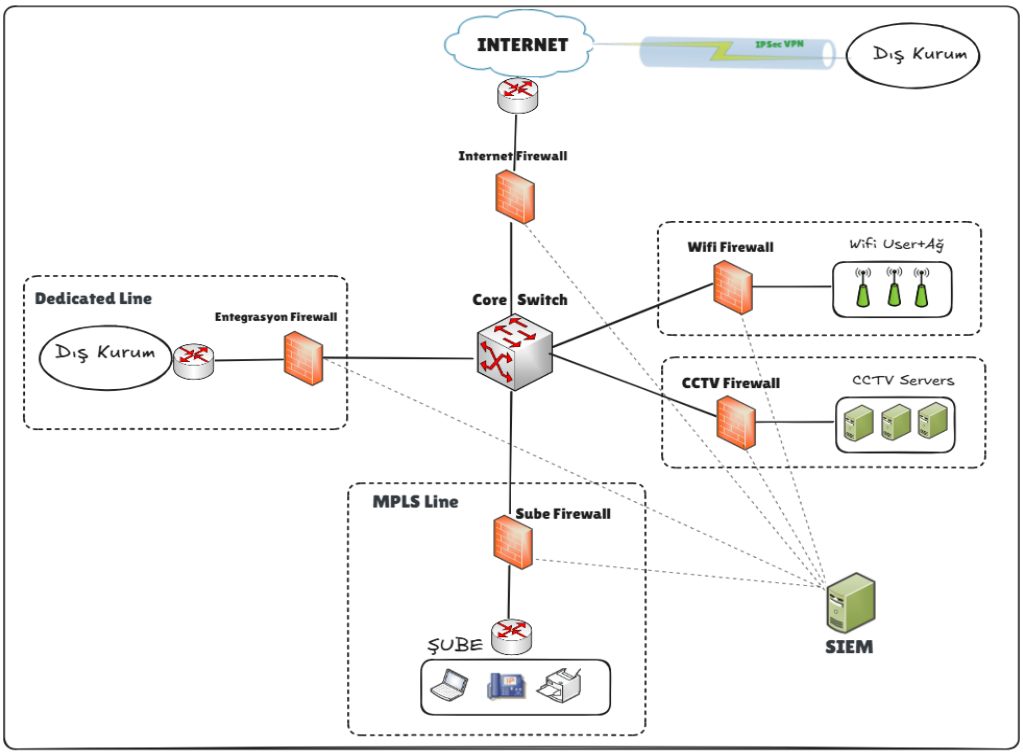
2. MATERYAL VE METOT

Çalışmamızın bu bölümünde, temsili kurumsal ağ topolojisine, protokollere ve modelimizde kullanılacak veri setine, veri önişleme ve model eğitimine yer verilmiştir.

2.1. Kurumsal Ağ Topolojisi

Firewall, ağ trafiğini izleyen ve kontrol eden bir sistemdir. Kullanıcıların ve sunucuların internet ve diğer ağlara erişimini, belirli kurallar ve politikalar çerçevesinde yönlendirir. Bu kurallar hangi trafiğin izinli olduğunu hangi kaynaklara erişim sağlanabileceğini ve hangi zaman dilimlerinde geçerli olduğunu belirler. Firewall, ağ güvenliğini artırmak amacıyla zararlı yazılımların ve yetkisiz erişimleri engeller. Kullanıcıların ve sunucuların yalnızca doğrulanmış ve güvenli kaynaklara erişimini sağlayarak ağ bütünlüğü ve veri güvenliğini sağlar.

Çalışma kapsamında kullandığımız temsili bir kurumsal ağın topolojisi Şekil-1 de verilmiştir. Topolojiyi incelediğimizde, farklı ağ katmanlarını ayırmak ve birbirleri arasındaki iletişimi kontrol altına almak üzere 5 adet yeni nesil firewall cihazı konumlandırılmıştır. Kurumun veri alışverişinde bulunduğu diğer kurumlar ile verinin hassasiyet derecesine göre IPSec VPN veya tamamen internete kapalı kuruma özel dedike hatlar üzerinden entegrasyon sağlanmıştır. Kuruma bağlı şube lokasyonları, MPLS (MultiProtocol Label Switching) devre üzerinden firewall kontrolünde yerel ağa bağlanmaktadır. Kurum personeli ve misafir kullanıcılara hizmet veren kablosuz ağ, diğer ağlardan izole edilmiştir ve yerel ağa firewall kontrolü ile bağlanmaktadır. CCTV cihazları diğer ağlardan izole edilmiştir ve yerel ağa firewall kontrolü ile bağlanmaktadır. Firewall cihazlarına gelen system, threat, configuration logları, SIEM (Security Information And Event Management) ürününe iletilmektedir.



Şekil 1. Kurumsal Ağ Topolojisi

2.2 IPSec ve MPLS Protokol

IPSec (Internet Protocol Security); uçtan uca gizli iletişim için kullanılabilen ve IP katmanında çalışan bir güvenlik protokolüdür. IP katmanının karşılaşabileceği güvenlik riskleri göz önüne alındığında, IPSec üç güvenlik hizmeti sağlar: gizlilik, bütünlük ve kimlik doğrulaması.

Gizlilik: Veri paketlerini şifreleyerek üçüncü tarafların bu verileri okumasını engeller. Hassas bilgilerin korunmasına ve izinsiz erişimlerin önlenmesine yardımcı olur.

Bütünlük: Verilerin iletim sırasında değiştirilmediğinden emin olmak için bütünlük kontrol mekanizmaları sağlar. Veri paketlerinin gönderici tarafından gönderildiği şekilde alıcıya ulaşmasını garanti eder.

Kimlik doğrulama: Veri paketlerinin kimliğini doğrular ve yalnızca güvenilir kaynaklardan gelen verilerin kabul edilmesini sağlar. Sahte veri paketlerinin önlenmesine yardımcı olur.

IPsec, üç alt protokolden oluşur. IKE (Internet Key Exchange) protokolü, AH (Authentication Header) protokolü ve ESP (Encapsulation Security Payload) protokolü. IKE protokolü kimlik doğrulama, kriptografik algoritma seçimi ve anahtar oluşturma için kullanılırken, AH ve ESP güvenli iletişim için kullanılır [12].

MPLS, özel geniş alan ağları üzerinden yönlendirmeyi yönetmek için ağ adresleri yerine “etiketlere” dayalı en kısa yolu kullanarak trafiği yönlendiren bir ağ teknolojisidir. Ölçeklenebilir ve protokolden bağımsız bir çözüm olan MPLS, her veri paketine etiketler atayarak paketin izlediği yolu kontrol eder. Kurum ve kuruluşlar, ülke çapında veya dünyanın çeşitli yerlerinde birden fazla uzak şubeleri olduğunda ve kuruluşun genel merkezindeki veya başka bir şube konumundaki veri merkezlerine veya uygulamalara erişmeleri gerektiğinde genellikle bu teknolojiyi kullanırlar. MPLS network altyapısının etkin bir şekilde kullanılabilmesi, ağın bir akışı taşıma sürecinde karşılaşılan servis gereksinimlerinin toplamı olan farklı servis kalitesi ayarının yapılmasıyla mümkündür [13].

2.3 Veri Seti

Makine öğrenmesi algoritmalarının performans ve sonucunu etkileyen önemli hususlardan biride veri setinin kullanışlı, sade ve açık bir formatta olmasıdır. Eksik, hatalı veya gürültülü veriler, modelin doğruluğunu ve genelleme yeteneğini olumsuz etkileyebilir. Bu durum hatalı tahminlere veya yanlış sonuçlara yol açabilir. Bu nedenle veri setinde, veri ön işleme, veri temizleme, veri dönüştürme adımları son derece kritiktir.

Veri ön işleme: Ham verilerin analiz veya modelleme için uygun hale getirilmesi sürecidir. Bu süreç, eksik verilerin doldurulması, aykırı değerlerin tespiti ve giderilmesi, veri normalizasyonu ve ölçeklendirme gibi adımları içerir. Veri madenciliği ve makine öğrenmesi uygulamalarında önemli bir aşamadır [14].

Veri temizleme: Veri temizleme, veri setlerinde bulunan eksik, hatalı, gereksiz, tutarsız verilerin belirlenip düzeltilmesi sürecidir. Örneğin, eksik değerlerin doldurulması veya aykırı değerlerin düzeltilmesi bu sürece dahildir. Veri temizliği, model performansını artırmak için gereklidir [15].

Veri dönüştürme: Verilerin analiz veya modelleme için daha uygun hale getirilmesi sürecidir. Bu süreç, değişken dönüşümü, özellik mühendisliği, veri boyutunun azaltılması ve kodlama işlemlerini içerir. Özellikle büyük veri ve makine öğrenmesi modellerinde önemli bir aşamadır [16]. Bu çalışmada makine öğrenmesi modellerinin geliştirilmesi ve test edilmesi için veri seti olarak, laboratuvar ortamında Next Generation Firewall sisteminden alınmış gerçek zamanlı tehdit logları kullanılmıştır. Bu loglar, sistemin güvenlik performansını ve tehdit algılama yeteneklerini değerlendirmek amacıyla toplanan Vulnerability, Antivirus, Spyware kategorisindeki verileri içermektedir. Veri setinde, 14 adet kolon ve 15,700 adet satırdan oluşan loglar bulunmaktadır. Veri setinde çeşitli düzenleme ve veri temizleme işlemleri yapılmıştır. Veri seti aşağıda belirtilen kolonlardan oluşmaktadır. İlgili öznitelik açıklamaları aşağıdaki Tablo 1'de yer almaktadır.

Tablo 1. Anomali Tespitinde Kullanılan Veri Seti Öznitelikleri

Type	Üretilen Log türünün TEHDİT olduğunu belirtir.
Threat/Content Type	Tehdit logunun alt tipidir. Aşağıdaki örnek değerleri içerir. Vulnerability: tespit edilen güvenlik açığı istismarı spyware: tespit edilen casus yazılım virus: tespit edilen virüs
Threat/Content Name	Tanınan ve spesifik tehditler için bir tanımlama göstergesidir.
Application	Session ile ilişkilendirilen uygulamayı belirtir.
Source Port	Session tarafından kullanılan kaynak portudur.
Destination Port	Session tarafından kullanılan hedef portudur.
IP Protocol	Session ile ilişkili IP protokolüdür.
Severity	Tehdit'in önem derecesini belirtir. İnformational, low, medium, high, critical vb.
Direction	Tehditin yönünü belirtir. (istemciden sunucuya veya sunucudan istemciye)
thr_category	Farklı tehdit imza tiplerini kategorilendirmek için kullanılır.
Subcategory of app	Uygulamaya ait alt kategorisidir
Category of app	Uygulama kategorisi olarak tanımlanır. (saas, networking, media vb)
Risk of app	Uygulamayla bağlantılı risk düzeyini belirtir. (en düşük:1, en yüksek:5).
Action	Session için gerçekleştirilen eylemdir. (Allow, deny, drop vb)

Aşağıda belirtilen Tablo 2'de, Tablo 1'de yer alan özniteliklerin küçük bir örneklemi bulunmaktadır.

Tablo 2. Çalışmada Kullanılan Veri Seti Örneği

Type	Threat/ Content Type	Threat/Content Name	Application	IP Proto	Severity	thr_category	Risk of app	Action
THREAT	vulnerability	HTTP Unauthorized Error	web-browsing	tcp	informational	brute-force	4	alert
THREAT	vulnerability	HTTP WWW-Auth.Fail	web-browsing	tcp	low	brute-force	4	alert
THREAT	vulnerability	Microsoft Windows NTLMSSP	ms-ds-smbv3	tcp	informational	info-leak	3	alert
THREAT	scan	SCAN: Host Sweep(8002)	not-applicable	tcp	Critical	unknown	1	block-ip
THREAT	vulnerability	Microsoft Windows NTLMSSP	ms-ds-smbv3	tcp	informational	info-leak	3	alert
THREAT	spyware	Suspicious TLS Evasion Found	ssl	tcp	informational	spyware	4	alert
THREAT	vulnerability	HTTP WWW-Auth. Fail	web-browsing	tcp	critical	brute-force	5	drop
THREAT	vulnerability	HTTP Unauthorized Error	web-browsing	tcp	informational	brute-force	4	alert
THREAT	virus	Virus/Win32.WGeneric.dxsdhp	ms-ds-smbv3	tcp	medium	portable exe	3	drop
THREAT	virus	Virus/Win32.WGeneric.dxsdhp	ms-ds-smbv3	tcp	medium	portable exe	3	drop
THREAT	spyware	Suspicious HTTP Evasion	git-base	udp	high	spyware	4	drop
THREAT	scan	SCAN: TCP Port Scan	not-applicable	tcp	client-to-server	unknown	1	block-ip
THREAT	spyware	Suspicious HTTP Evasion	web-browsing	tcp	informational	spyware	4	alert
THREAT	spyware	Suspicious TLS Evasion Found	ssl	tcp	informational	spyware	4	alert
THREAT	flood	Session Limit Event	not-applicable	tcp	high	ddos	1	drop
THREAT	flood	Session Limit Event	not-applicable	tcp	high	ddos	1	drop
THREAT	flood	Session Limit Event	not-applicable	tcp	high	ddos	1	drop
THREAT	scan	SCAN: Host Sweep	not-applicable	tcp	high	ddos	1	block-ip

2.4. Veri Önişleme ve Model Eğitimi

Veri setimiz makine öğrenmesi modellerinin eğitimi ve değerlendirilmesi için eğitim ve test verisi olarak iki ayrı bölüme ayrılmıştır. Bu ayrım, modelin performansını değerlendirmek ve genelleme yeteneğini ölçmek için yaygın bir uygulamadır. Veri setinin %80'i eğitim verisi olarak kullanılırken, %20'si test verisi olarak kullanılmıştır. Veri setinde severity değeri informational, low olan loglar zararsız, medium, high ve critical seviyesine sahip olan loglar zararlı olarak sınıflandırılmıştır.

3. BULGULAR VE TARTIŞMA

Bu bölümde, çalışmamız kapsamında kullanılan makine öğrenmesi modellerinin sonuçlarına dair bulgular yer almaktadır. Bu çalışmamızın öncelikli amacı, kurumsal ağlarda kurum için güvenilir kabul edilen ağ bağlantıları üzerinden bilinçli veya bilinçsiz bir şekilde gerçekleştirilen atak ve tehditlerin makine öğrenmesi teknikleri kullanılarak tespiti ve sınıflandırılmasıdır. Çalışma kapsamında kullanılan makine öğrenmesi algoritmaları, kurumsal ağlarda iç ağdan gelen tehdit ve atakların etkin bir biçimde sınıflandırılması amacıyla seçilmiştir. Seçilen algoritmaların her biri, farklı matematiksel temellere dayanmakta olup ağ güvenliği loglarının çok boyutlu ve heterojen yapısına karşı farklı perspektiflerden yaklaşarak genel başarımı artırmayı hedeflemektedir.

Çalışmamızda Next Generation Firewall sisteminden alınmış gerçek zamanlı Vulnerability, Antivirus, Spyware threat loglarına dair veri seti kullanılmıştır. Bu veri seti, Random Forest, Logistic Regression, XGBoost, MLP (Multi layer Preception), SVC (Support Vector Classification) makine öğrenmesi algoritmaları kullanarak doğruluk, kesinlik, duyarlılık ve F1 skoru kriterleri ile değerlendirilmiştir.

Algoritma sonuçları, Python kod parçalarını satır satır çalıştırmaya ve her kod parçacığının sonuçlarını almamıza olanak sağlayan Jupyter Notebook ile çalıştırılmıştır. Çalışmada kullanılan bilgisayar, Windows 10 işletim sistemi, 2.20 GHz, 32 GB RAM özelliklerine sahiptir. Veri seti üzerinde 5 adet algoritma kullanılmıştır. Doğruluk oranına göre en yüksek başarımı gösteren algoritma XGBoost olmuştur. Çalışmamızda kullanılan algoritmaların sonuçları, %85 ile %99 arasında doğruluk oranına sahiptir. Machine Learning Algoritma sonuçları Tablo 3'te bulunmaktadır.

Tablo 3. Makine Öğrenmesi Algoritma Sonuçları

Makine Öğrenmesi Algoritması	Doğruluk Oranı	F1 Skoru	Kesinlik	Duyarlılık
Random Forest	0,85891	0,90648	0,97149	0,84962
LogisticRegression	0,89904	0,93902	0,91354	0,96596
XGBoost	0,99171	0,97896	0.97110	0,98694
MLP (Multi-Layer Perceptron)	0,97261	0,98288	0,98839	0,97744
SVC (Support Vector Machine)	0,86560	0,91334	0,94921	0,88009

3.1. Random Forest Algoritma Sonucu

Random Forest, birden fazla karar ağacı kullanarak verilerin genellemesi ve daha doğru tahminler yapması nedeniyle tercih edilir. Özellikle karmaşık ve büyük veri setlerinde dengesiz sınıfları iyi bir şekilde yönetir. Random Forest algoritmasının siber güvenlikteki etkinliği, saldırı tespitinde hassasiyet ve doğruluk sağladığı birçok çalışmada gösterilmiştir. Örneğin, Zhang ve arkadaşları çalışmasında siber saldırıların tespitinde Random Forest' in üstün performansını vurgulamıştır [17]. Random Forest algoritmasının sonuçlarına göre, doğruluk oranı %85'dir. F1 skoru değeri %90 olarak ölçülmüştür. F1 skoru, yanlış pozitif ve yanlış negatiflerin dengeli bir biçimde değerlendirildiği durumlarda, modelin performansını ölçmek için önemli bir metriktir. Kesinlik değeri %97, duyarlılık değeri %84 olarak ölçümlenmiştir. Şekil 2'de Random Forest algoritması kullanılarak veri kümemiz sınıflandırılmış ve modelin performansı ve sonuçları bulunmaktadır.

```
# Veriyi Eğitim ve test setlerine ayirma
X_train, X_test, y_train, y_test = train_test_split(features, target, test_size=0.2, random_state=42)

# Random Forest modelini oluşturma ve eğitme
rf = RandomForestClassifier(n_estimators=100, random_state=42)
rf.fit(X_train_res, y_train_res)

# Test seti üzerinde tahmin yapma
y_pred = rf.predict(X_test_scaled)

# Performans metriklerini hesaplama
accuracy = accuracy_score(y_test, y_pred)
f1 = f1_score(y_test, y_pred, pos_label='zararlı')
precision = precision_score(y_test, y_pred, pos_label='zararlı')
recall = recall_score(y_test, y_pred, pos_label='zararlı')

print("✅ Random Forest Sonuclari:")
print(f"Doğruluk: {accuracy:}")
print(f"F1 Skoru: {f1:}")
print(f"Precision: {precision:}")
print(f"Recall: {recall:}")
```

```
✅ Random Forest Sonuclari:
Doğruluk: 0.8589171974522293
F1 Skoru: 0.906480895081275
Precision: 0.9714932126696832
Recall: 0.849624060150376
```

Şekil 2. Random Forest Algoritma Modeli Sonuçları

3.2. Logistic Regression Algoritma Sonucu

Logistic Regression, ayrık değişkenler arasındaki ilişkiyi kurmak için kullanılır. Geçmiş verilerden öğrenilen ilişkilere dayanarak bir olayın gerçekleşme olasılığını tahmin etmeye yönelik istatiksel bir model sağlar [18]. Bu çalışmamızda Logistic Regression algoritması, Antivirüs, Vulnerability, Antispyware güvenlik sistemlerinden gelen loglarda zararlı ve zararsız davranışların istatiksel olarak ayırt edilebilmesi amacıyla kullanılmıştır. Logistic Regression algoritmasının sonuçlarına göre, doğruluk oranı %89'dur. F1 skoru değeri %93 olarak ölçülmüştür. Recall (Duyarlılık) değeri %96'dır. Precision (Kesinlik) değeri %91'dir. Kesinlik değeri, false positive durumların sınırlanması gereken durumlarda önemlidir. Şekil 3'te Logistic Regression algoritması kullanılarak veri kümemiz sınıflandırılmış ve modelin performansı ve sonuçları sunulmuştur.

```python
# Eğitim/test ayrımı
X_train, X_test, y_train, y_test = train_test_split(features_scaled, target, test_size=0.2, random_state=42)

# Logistic Regression modelini oluştur ve eğit
model = LogisticRegression(random_state=42, max_iter=1000, C=1.0)
model.fit(X_train_balanced, y_train_balanced)

# Tahmin
y_pred = model.predict(X_test)

# Performans metrikleri
accuracy = accuracy_score(y_test, y_pred)
f1 = f1_score(y_test, y_pred, pos_label='zararlı')
precision = precision_score(y_test, y_pred, pos_label='zararlı')
recall = recall_score(y_test, y_pred, pos_label='zararlı')

print("✅ Logistic Regression Sonuclari:")
print(f"Doğruluk: {accuracy:}")
print(f"F1 Skoru: {f1:}")
print(f"Precision: {precision:}")
print(f"Recall: {recall:}")
```

```
✅ Logistic Regression Sonuclari:
Doğruluk: 0.8990445859872611
F1 Skoru: 0.939026735910752
Precision: 0.9135479041916168
Recall: 0.9659675504550851
```

Şekil 3. Logistic Regression Algoritma Modeli Sonuçları

3.3. XGBoost Algoritma Sonucu

XGBoost, yüksek doğruluk ve verimlilik sağlayan bir gradient boosting algoritmasıdır. Büyük veri setlerinde hızlı çalışması ve düzenleme teknikleri ile model doğruluğunu artırması nedeniyle tercih edilir. XGBoost, hiperparametre ayarlamaları ile farklı veri setlerine kolayca adapte olabilir. İç ağdan gelen saldırıların ve normal aktivitelerin çeşitli özelliklerini modellemek için özelleştirilebilir [19]. XGBoost algoritmasının sonuçlarına göre, doğruluk oranı %99'dur. F1 skoru değeri %97 olarak ölçülmüştür. Precision (Kesinlik) değeri %97'dir. Recall (Duyarlılık) değeri %98'dir. Şekil 4'te XGBoost algoritması kullanılarak veri kümemiz sınıflandırılmış ve modelin performansı ve sonuçları sunulmuştur.

```python
# Eğitim/test ayrımı
X_train, X_test, y_train, y_test = train_test_split(features_scaled, target_encoded, test_size=0.2, random_state=42)

# Modeli oluşturma ve eğitme
model = xgb.XGBClassifier(random_state=42)
model.fit(X_train_balanced, y_train_balanced)

# Tahmin
y_pred = model.predict(X_test)

# Performans metrikleri
accuracy = accuracy_score(y_test, y_pred)
f1 = f1_score(y_test, y_pred, pos_label=1)  # 'zararlı' sayısal karşılığı 1
precision = precision_score(y_test, y_pred, pos_label=1)
recall = recall_score(y_test, y_pred, pos_label=1)

# Sonuçlar
print("✅ XGBoost Sonucu:")
print(f"Doğruluk : {accuracy}")
print(f"F1 Skoru: {f1}")
print(f"Precision: {precision}")
print(f"Recall: {recall}")
```

```
✅ XGBoost Sonucu:
Doğruluk : 0.9917197452229299
F1 Skoru: 0.9789644012944984
Precision: 0.971107544141252
Recall: 0.9869494290375204
```

Şekil 4. XGBoost Algoritma Modeli Sonuçları

3.4. MLP (Multi-Layer Perceptron) Algoritma Sonucu

MLP, yapay sinir ağları yapısına sahip bir algoritmadır ve veri setindeki karmaşık, doğrusal olmayan ilişkileri öğrenme kapasitesine sahiptir. Bu özellik, zararlı ve zararsız aktivitelerin belirgin şekilde ayrılmadığı durumlarda avantaj sağlamaktadır. MLP, karmaşık veri örüntülerini öğrenme kapasitesi ve doğrusal olmayan anomalilerin tespiti üzerinde etkilidir [20]. Bu çalışmamızda MLP algoritması, Antivirüs, Vulnerability ve Antispyware logları içinde doğrudan belirgin olmayan tehdit davranışlarını tespit etmek amacıyla kullanılmıştır.

MLP algoritmasının sonuçlarına göre, doğruluk oranı %97'dir. F1 skoru değeri %98 olarak ölçülmüştür. Precision (Kesinlik) değeri %98'dir. Recall (Duyarlılık) değeri %97'dir. Şekil 5'te MLP algoritması kullanılarak veri kümemiz sınıflandırılmış ve modelin performansı ve sonuçları sunulmuştur.

```python
# Eğitim/test ayrımı
X_train, X_test, y_train, y_test = train_test_split(features_scaled, target, test_size=0.2, random_state=42)

# MLP modelini oluştur ve eğitme
mlp = MLPClassifier(hidden_layer_sizes=(100,50), max_iter=500, random_state=42)
mlp.fit(X_train_balanced, y_train_balanced)

# Tahmin
y_pred = mlp.predict(X_test)

# Performans metrikleri
accuracy = accuracy_score(y_test, y_pred)
f1 = f1_score(y_test, y_pred, pos_label='zararlı')
precision = precision_score(y_test, y_pred, pos_label='zararlı')
recall = recall_score(y_test, y_pred, pos_label='zararlı')

print("✅ MLP Sonuclari:")
print(f"Doğruluk: {accuracy:}")
print(f"F1 Skoru: {f1:}")
print(f"Precision: {precision:}")
print(f"Recall: {recall:}")
```

```
✅ MLP Sonuclari:
Doğruluk: 0.9726114649681529
F1 Skoru: 0.9828889773179467
Precision: 0.9883953581432573
Recall: 0.9774436090225563
```

Şekil 5. MLP Algoritma Modeli Sonuçları

3.5. SVC (Support Vector Classification) Algoritma Sonucu

SVC, sınıflandırma, regresyon, yenilik algılama görevleri ve özellik azaltma için yaygın olarak kullanılmaktadır. Kurumsal güvenlik cihazlarından elde edilen Antivirüs, Vulnerability, AntiSpyware logları gibi veriler yüksek boyutlu özellikler içermektedir. SVC, ağ verilerindeki normal ve anormal davranışları ayırt etme yeteneğine sahiptir [21]. Bu çalışmamızda SVC algoritması, zararlı ve zararsız trafik örnekleri arasındaki sınırların net şekilde ayrılması amacıyla kullanılmıştır. SVC algoritmasının sonuçlarına göre, doğruluk oranı %86'dır. F1 skoru değeri %91 olarak ölçülmüştür. Precision (Kesinlik) değeri %94'dir. Recall (Duyarlılık) değeri %88'dir. Şekil 6'da SVC algoritması kullanılarak veri kümemiz sınıflandırılmış ve modelin performansı ve sonuçları sunulmuştur.

```python
# Etiketler sayısal hale getirilir
label_encoder = LabelEncoder()
y_train_encoded = label_encoder.fit_transform(y_train_resampled)
y_test_encoded = label_encoder.transform(y_test)

# SVC modelini oluşturma ve eğitme
svc = SVC(kernel='rbf', probability=True, random_state=42)
svc.fit(X_train_resampled, y_train_encoded)

# Tahmin yapma
y_pred_encoded = svc.predict(X_test)
y_pred = label_encoder.inverse_transform(y_pred_encoded)

# Metrik hesabı
accuracy = accuracy_score(y_test, y_pred)
f1 = f1_score(y_test, y_pred,pos_label='zararlı')
precision = precision_score(y_test, y_pred,pos_label='zararlı')
recall = recall_score(y_test, y_pred,pos_label='zararlı')

print("✅ SVC Sonuc:")
print(f"Doğruluk: {accuracy:}")
print(f"F1 Skoru: {f1:}")
print(f"Precision: {precision:}")
print(f"Recall: {recall:}")
```

```
✅ SVC Sonuc:
Doğruluk: 0.8656050955414013
F1 Skoru: 0.913347022587269
Precision: 0.9492104139991464
Recall: 0.8800949742777998
```

Şekil 6. SVC Algoritma Modeli Sonuçları

4. SONUÇ VE ÖNERİLER

Siber saldırıların hızla çoğalması ve karmaşıklaşması nedeniyle, sürekli güncellenen ve gelişen dinamik siber güvenlik stratejilerine olan ihtiyaç giderek artmaktadır. Saldırıların tespit edilmesi ve önlenmesi için çeşitli yöntemler mevcut olsa da sürekli değişen saldırı türlerine ve yoğunluklarına karşı hazır olunması gerekmektedir.

Kurumlar her geçen gün daha fazla sayıda ağ bağlantısına ve cihazlara sahip olmakta, bu durum beraberinde çeşitli güvenlik tehditlerini getirmektedir. Kurumsal ağlarda çoğu zaman odak noktası dış kaynaklı siber saldırılar ve tehditler olurken, kurum içinden gelebilecek saldırılar göz ardı edilebilmektedir. Oysaki, kurum iç ağında meydana gelebilecek bir güvenlik ihlali, en az kurum dış ağından gelen bir saldırı kadar yıkıcı etkilere, maddi kayıp ve itibar zedelenmesine neden olabilir.

Çalışmamızda kullanılan veri seti Next Generation Firewall sisteminden alınmış gerçek zamanlı Vulnerability, Antivirus, Antispyware threat loglarına dair veri seti kullanılmıştır. Bu veri seti için 5 farklı makine öğrenmesi algoritması

çalıştırılmıştır. Elde edilen sonuçlar incelendiğinde, XGBoost algoritmasının %99,17 doğruluk oranı, %97,89 F1 skoru, %97,11 kesinlik ve %98,69 duyarlılık değerleriyle genel olarak en yüksek performansı sergilediğini ortaya koymaktadır. Bu sonuçlar, XGBoost'un zararlı örnekleri doğru bir şekilde tanımlama ve sınıflandırma yeteneğini kanıtlamaktadır.

MLP algoritması, %98,84 kesinlik oranı ile yanlış pozitif sınıflandırmaları minimize etme konusunda en yüksek başarıyı göstermiştir. %97,26 doğruluk oranı ve %98,28 F1 skoru ile de güvenilir bir performans sergilemiştir.

Logistic Regression algoritması, %89,90 doğruluk oranı ve %93,90 F1 skoru ile özellikle duyarlılık metriğinde %96,59 ile zararlı olayların kaçırılma riskini minimize etmiştir.

Random Forest algoritması, %85,89 doğruluk oranı ve %90,64 F1 skoru ile diğer algoritmalara göre orta düzeyde bir performans sunarken, duyarlılık metriği %84,96 ile nispeten daha düşük kalmış olup bu durum zararlı olayların tespitinde bazı eksiklikler olduğunu göstermektedir.

SVC algoritması, %86,56 doğruluk oranı ve %91,33 F1 skoru performansı sergilemiştir. %94,92 kesinlik oranı ile zararsız olayları yanlış bir şekilde zararlı olarak sınıflandırma riskini düşük tutmayı başarmış, duyarlılık metriğinde %88,01 değerinde bir performans göstermiştir.

İlerleyen dönemlerde, siber güvenlik çözümlerinin hibrit yapıda ve yapay zeka entegreli olarak gelişmesi öngörülmektedir. Makine öğrenmesi algoritmalarının performansını artırmak için daha geniş ve çeşitlilik içeren veri setleriyle modellerin eğitilmesi ve test edilmesi, geliştirilen modellerin gerçek zamanlı saldırı tespit sistemlerine entegre edilmesi, farklı algoritmaların güçlü yönlerini birleştirerek hibrit yaklaşımlar geliştirilmesi kurumsal ağ güvenliğinin dinamik bir şekilde güçlendirilmesine katkıda bulunabilir.

ÇIKAR ÇATIŞMASI BEYANI

Bu araştırmada çıkar çatışması bulunmamaktadır.

VERİ KULLANILABİLİRLİĞİ BEYANI

Araştırma sırasında üretilen veya kullanılan tüm veriler, modeller ve kodlar gönderilen makalede yer almaktadır.

KAYNAKLAR

[1] H. Yılmaz, "TS ISO/IEC 27001 Bilgi Güvenliği Yönetimi Standardı kapsamında bilgi güvenliği yönetim sisteminin kurulması ve bilgi güvenliği risk analizi," *Kamu İç Denetçileri Derneği (KİDDER) Denetişim Dergisi*, 2014–2015.

[2] M. Karakaya, *Kurumsal güvenlik için siber tehditlerin incelenmesi ve saldırı senaryoları (Research of cyber threats and attack scenarios for corporate security)*, 2022.

[3] N. Bakhareva, A. Shukhman, A. Matveev, P. Polezhaev, Y. Ushakov, and L. Legashev, "Attack detection in enterprise networks by machine learning methods," in *Proc. 2019 Int. Russian Automation Conf. (RusAutoCon)*, Sep. 2019, pp. 1–6.

[4] B. Chikkoppa, J. Hanumanthappa, V. Pati, S. Allagi, L. S. Rodriguez-Baca, and C. F. Cruzado, "A comparative study of malware detection in enterprise networks," in *Proc. 2024 2nd World Conf. on Communication & Computing (WCONF)*, Jul. 2024, pp. 1–5.

[5] V. L. Thing, "IEEE 802.11 network anomaly detection and attack classification: A deep learning approach," in *Proc. IEEE Wireless Communications and Networking Conf. (WCNC)*, Mar. 2017, pp. 1–6.

[6] H. Çakır and H. Yaşar, "Kurumsal siber güvenliğe yönelik tehditler ve önlemleri," *Düzce Üniversitesi Bilim ve Teknoloji Dergisi*, vol. 3, no. 2, pp. 488–507, 2015.

[7] Z. Senturk and E. Irmak, "Windows aktif dizin etki alanı servisi ve kurumsal ağ güvenliği: PowerShell erişiminin analizi ve önlemler," *Gazi Univ. J. Sci. Part C: Design and Technology*, vol. 12, no. 3, pp. 475–487, 2024.

[8] M. K. Pehlivanoğlu, R. Atay, and D. E. Odabaş, "İki seviyeli hibrit makine öğrenmesi yöntemi ile saldırı tespiti," *Gazi Mühendislik Bilimleri Dergisi*, vol. 5, no. 3, pp. 258–272, 2019.

[9] M. Hacıbeyoğlu, F. N. Arıcı, and M. Karaaltun, "Intrusion detection system application with machine learning," *Afyon Kocatepe Univ. Fen ve Mühendislik Bilimleri Dergisi*, vol. 24, no. 5, pp. 1166–1180, 2024.

[10] Ç. Kılınç and Ö. Can, "Siber güvenlikte T-Pot honeypot uygulanması: Kurumsal ağ üzerinde örnek durum çalışması," *Computer Science*, IDAP-2023, pp. 24–30, 2023.

[11] H. Ulus and M. Demirci, "Kurum içi saldırıların tespiti ve önlenmesi için sunucu izleme uygulaması," *Gazi Univ. J. Sci. Part C: Design and Technology*, vol. 6, no. 3, pp. 507–523, 2018.

[12] J. Guo, C. Gu, X. Chen, and F. Wei, "Model learning and model checking of IPsec implementations for Internet of Things," 2019.

[13] P. Kırcı, Ç. Toka, and F. Erdem, "IP/MPLS ağlar üzerinde sanal-yerel servisler ve yönlendirici konfigürasyonları," *Uludağ Univ. Mühendislik Fakültesi Dergisi*, vol. 20, no. 2, pp. 155–165, 2015.

[14] S. García, J. Luengo, and F. Herrera, *Data Preprocessing in Data Mining*. Cham, Switzerland: Springer, 2015.

[15] E. Rahm and H. H. Do, “Data cleaning: Problems and current approaches,” *IEEE Data Engineering Bulletin*, vol. 23, no. 4, pp. 3–13, 2000.

[16] D. Pyle, *Data Preparation for Data Mining*. San Francisco, CA, USA: Morgan Kaufmann, 1999.

[17] J. Zhang, M. Zulkernine, and A. Haque, “Random-forests-based network intrusion detection systems,” *IEEE Trans. Syst., Man, Cybern., Part C (Appl. Rev.)*, vol. 38, no. 5, pp. 649–659, 2008.

[18] M. Bhavsar, K. Roy, J. Kelly, and O. Olusola, “Anomaly-based intrusion detection system for IoT application,” *Discover Internet of Things*, vol. 3, no. 1, p. 5, 2023.

[19] X. Wang and X. Lu, “A host-based anomaly detection framework using XGBoost and LSTM for IoT devices,” *Wireless Communications and Mobile Computing*, vol. 2020, no. 1, 2020.

[20] R. Chalapathy and S. Chawla, “Deep learning for anomaly detection: A survey,” 2019.

[21] K. Ghanem, F. J. Aparicio-Navarro, K. G. Kyriakopoulos, S. Lambotharan, and J. A. Chambers, “Support vector machine for network intrusion and cyber-attack detection,” in *Proc. Sensor Signal Processing for Defence Conf. (SSPD)*, Dec. 2017, pp. 1–5.