KRİTİK ALTYAPILARA YÖNELİK DERİN ÖĞRENME TABANLI SALDIRI TESPİT SİSTEMİ TASARIMI

2023
DOKTORA TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ

Hakan Can ALTUNAY

Tez Danışmanı
Doç. Dr. Zafer ALBAYRAK

KRİTİK ALTYAPILARA YÖNELİK DERİN ÖĞRENME TABANLI SALDIRI TESPİT SİSTEMİ TASARIMI

Hakan Can ALTUNAY

Tez Danışmanı
Doç. Dr. Zafer ALBAYRAK

T.C.
Karabük Üniversitesi
Lisansüstü Eğitim Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalında
Doktora Tezi
Olarak Hazırlanmıştır

KARABÜK
Temmuz 2023

Hakan Can ALTUNAY tarafından hazırlanan "KRİTİK ALTYAPILARA YÖNELİK DERİN ÖĞRENME TABANLI SALDIRI TESPİT SİSTEMİ TASARIMI" başlıklı bu tezin Doktora Tezi olarak uygun olduğunu onaylarım.

Doç. Dr. Zafer ALBAYRAK .....
Tez Danışmanı, Bilgisayar Mühendisliği Anabilim Dalı

Bu çalışma, jürimiz tarafından oy birliği ile Bilgisayar Mühendisliği Anabilim Dalında Doktora tezi olarak kabul edilmiştir. 10/07/2023

Ünvanı, Adı SOYADI (Kurumu)	İmzası
Başkan : Doç. Dr. Zafer ALBAYRAK (SUBÜ)	.....
Üye : Prof. Dr. Necmi Serkan TEZEL (KBÜ)	.....
Üye : Doç. Dr. Salih GÖRGÜNOĞLU (KÜ)	……..
Üye : Dr. Öğr. Üyesi İsa AVCI (KBÜ)	.....
Üye : Dr. Öğr. Üyesi Muhammet ÇAKMAK (SÜ)	…….

KBÜ Lisansüstü Eğitim Enstitüsü Yönetim Kurulu, bu tez ile, Doktora derecesini onamıştır.

Prof. Dr. Müslüm KUZU .....
Lisansüstü Eğitim Enstitüsü Müdürü

"Bu tezdeki tüm bilgilerin akademik kurallara ve etik ilkelere uygun olarak elde edildiğini ve sunulduğunu; ayrıca bu kuralların ve ilkelerin gerektirdiği şekilde, bu çalışmadan kaynaklanmayan bütün atıfları yaptığımı beyan ederim."

Hakan Can ALTUNAY

ÖZET

Doktora Tezi

KRİTİK ALTYAPILARA YÖNELİK DERİN ÖĞRENME TABANLI SALDIRI TESPİT SİSTEMİ TASARIMI

Hakan Can ALTUNAY

Karabük Üniversitesi
Lisansüstü Eğitim Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Tez Danışmanı:
Doç. Dr. Zafer ALBAYRAK
Temmuz 2023, 66 sayfa

Enerji, ulaşım, üretim tesisleri gibi kritik altyapıya sahip sistemlerin siber saldırılara karşı korunması ulusal güvenlik için kritik öneme sahiptir. Gelişen teknoloji ile birlikte kritik altyapılarda internete bağlı cihazların sayısı artmıştır. Bu artışla birlikte kritik altyapılara yönelik gerçekleşen siber saldırıların sayısı ve çeşidi de artmıştır. Kiritk altyapıya sahip sistemler sahada bulunan pek çok cihaz ile birlikte fiziksel bir süreci yönetir. Bu cihazlar eskiden endüstriyel kontrol sistemlerine özeldi. Ancak günümüzde bu görevi genel amaçlı bilgi işlem teknolojileri özellikle de Nesnelerin İnterneti (IoT) gerçekleştirmektedir. Gün geçtikçe endüstride Nesnelerin interneti büyük ölçekli bir ağ haline gelmiştir. Endüstriyel IoT (IIoT) ağları maliyet ve esneklik açısından faydalı olsada, daha fazla siber saldırı çeşidi ile karşı karşıya kalmaktadır.

Kritik altyapılarda bulunan güvenlik ve gizlilik endişeleri, araştırmacıların farklı savunma mekanizmaları geliştirmesine sebep olmuştur. Bu mekanizmaların başında saldırı tespit sistemi (IDS) gelmektedir. Makine öğrenmesi yöntemleri saldırı tespit sistemlerinde tercih edilse de günümüzde artan veri miktarına bağlı olarak derin öğrenme yöntemleri sıklıkla kullanılmaktadır. Bununla birlikte bir sisteme gerçekleştirilen siber saldırılar sonsuza kadar engellenemez. Fakat sistemlerin korunumu için siber saldırıların gerçek zamanlı tespiti gerekmektedir. Kirik altyapıya sahip IIoT ağlarına yönelik izinsiz giriş tespit sistemleri konusunda sınırlı sayıda araştırma bulunmaktadır.

Bu tez çalışmasında, IIoT ağlarındaki güvenlik anormalliklerini tespit etmek için derin öğrenme algoritmalarını hibrit olarak kullanan yeni bir saldırı tespit sistemi modeli öneriyoruz. Hibrit modelde CNN ve LSTM algoritmaları kullanılmıştır. Bu önerilen modelin saldırı algılamadaki doğruluğu güncel ve karmaşık veri setleri olan X-IIoTID ve UNSW-NB15 ile test edilmiştir. Önerilen modelde veri setleri içerisindeki atak paketleri hem ikili hem de çok sınıflı sınıflandırmaya tabi tutulmuştur. Deneysel çalışma sonunda elde edilen sonuçlara göre, önerilen saldırı tespit sisteminin izinsiz girişleri yüksek başarımla etkili bir şekilde tespit edebildiği doğrulanmaktadır.

Anahtar Sözcükler : Kritik altyapı güvenliği, Endüstriyel IoT, derin öğrenme, saldırı tespit sistemi, siber güvenlik

Bilim Kodu **:** 92403

ABSTRACT

Ph.D. Thesis

DEEP LEARNING BASED-INTRUSION DETECTION SYSTEM DESIGN FOR CRITICAL INFRASTRUCTURE

Hakan Can ALTUNAY

Karabük University
Institute of Graduate Programs
Department of Computer Engineering

Thesis Advisor:
Assoc. Prof. Dr. Zafer ALBAYRAK
July 2023, 66 pages

Protecting critical infrastructure systems such as energy, transportation, and production facilities against cyber-attacks is of critical importance in terms of national security. The number of internet-connected devices within critical infrastructures has increased based on advancing technology. Correspondingly, the number and types of cyber-attacks performed on critical infrastructures have also increased. Critical infrastructure systems conduct a physical process accompanied by several devices situated in the field. In previous years, these were the devices specific to the industrial control systems. However, this task is performed today by general-purpose information technologies, particularly the Internet of Things (IoT). The Internet of Things has become a large-scale network within the industry with each passing day. The Industrial Internet of Things (IIoT) networks come under more types of cyber-attacks although being advantageous in terms of cost and flexibility.

Security and privacy concerns arising in critical infrastructures have led researchers to develop various defense mechanisms. The intrusion detection system (IDS) is the leading one among these mechanisms. Machine learning methods are preferred in intrusion detection systems, whereas deep learning methods are frequently used based on today's increasing amount of data. Nevertheless, cyber-attacks performed on a system cannot be prevented forever. All the same, real-time detection of cyber-attacks is a requirement for the protection of the systems. There are limited numbers of studies conducted on intrusion detection systems for IIoT networks with critical infrastructure.

In this thesis study, we proposed a new model of an intrusion detection system using hybrid deep learning algorithms to detect security anomalies in IIoT networks. The CNN and LSTM algorithms were used in the hybrid model. The intrusion detection accuracy of the proposed model was tested through X-IIoTID and UNSW-NB15, which are up-to-date and complex datasets. In the proposed model, the attack packages within the datasets were subjected to both binary and multi-class classification. The results obtained at the end of the experimental study confirm that the proposed intrusion detection system has the ability to efficiently detect intrusions with high performance.

Key Word **:** Critical infrastructure security, Industrial IoT, deep learning, intrusion detection system, cyber security

Science Code : 92403

TEŞEKKÜR

Öncelikle, tez konusunu seçerken isteklerimi göz önünde bulundurup bana yardımcı olan ve tezin üretim sürecinde rehberliği, sabrı, motivasyonu ve desteğini esirgemeyen danışmanım Doç. Dr. Zafer Albayrak'a şükranlarımı sunarım.

Tez çalışmamdaki önemli katkıları sebebiyle tez izleme komitesi ve tez savunma jürisindeki hocalarıma teşekkürü bir borç bilmekteyim. Tez çalışması süresince yardımlarından dolayı Öğr. Gör. Serkan Varan'a teşekkür ederim.

Tez sürecinde benden desteğini bir an için bile esirgemeyen sevgili eşime ve oğluma, eğitim hayatım boyunca sürekli yanımda olan anneme, babama ve kardeşime, çalışmalarımı sabırla destekleyen geniş aileme, en içten teşekkürlerimi sunarım.

İÇİNDEKİLER

Sayfa

ŞEKİLLER DİZİNİ

Sayfa

ÇİZELGELER DİZİNİ

Sayfa

SİMGELER VE KISALTMALAR DİZİNİ

SİMGELER

FP : Yanlış pozitif

FN : Yanlış negatif

TP : Gerçek pozitif

TN : Gerçek negatif

h : Çıktı

x : Girdi

x_i^a : a'nın i. öznitelik haritası

ϕ : Etkinleştirme

k_i : Giriş özniteliği

w_{ji}^a : a'nın i. niteliği ile (a-1) katmanının j. niteliği arasındaki bağlantı ağırlığı

b_j^a : İlgili katmandaki sapma

c : Alt örnekleme

β : Ağırlıklandırma matrisi

y^m : Tam bağlantı katmanının çıktısı

x^{m-1} : Tam bağlantı katmanının girdisi

w^m : Ağırlık katsayısı

b^m : Sapma

X^t : t zaman adımındaki giriş vektörü

h_{t-1} : Zaman adımındaki (t-1) durumu

C_{t-1} : Zaman adımındaki (t-1) bellek hücresi durumu

$f_i^{(t)}$: Unutma kapısı

b^f : Sapma

Z^f : Girdi ağırlığı

D^f : Tekrarlayan ağırlık

$n_i^{(t)}$: i hücresinin durumundaki güncelleme

$p_i^{(t)}$: i hücresinin giriş kapısı

$h_i^{(t)}$: i. gizli durum

$s_i^{(t)}$: i. çıkış kapısı

b : Sapma

Z : Girdi ağırlığı

D : Tekrarlayan ağırlık

b^0 : Sapma

Z^0 : Girdi ağırlığı

D^0 : Tekrarlayan ağırlık

KISALTMALAR

AE	: Oto Kodlayıcı
CNN	: Evrişimli Sinir Ağı
LSTM	: Uzun Kısa Süreli Bellek
RNN	: Tekrarlayan Sinir Ağı
ICS	: Endüstriyel Kontrol Sistemi
IDS	: Saldırı Tespit Sistemi
SCADA	: Denetleyici kontrol ve Veri Toplama
TCP	: Transmission Control Protocol
IP	: Internet Protocol
TCP / IP	: İnternet
IoT	: Nesnelerin İnterneti
IIoT	: Endüstriyel nesnelerin İnterneti
MQTT	: Mesaj Sıraya Alma ve Telemetri Aktarma
LoRaWAN	: Long Range Wide Area Network
PLC	: Programlanabilir Mantıksal Denetleyici
DoS	: Servis Dışı Bırakma
DDoS	: Dağıtık Servis Dışı Bırakma
Backdoor	: Arka Kapı
MiTM	: Ortadaki Adam
Phishing	: Oltalama
SQL	: Structured Query Language
SQLI	: Structured Query Language Injection

WAF	: Güvenlik Duvarı
XSS	: Cross Site Scripting
Malware	: Zararlı Yazılım
Worms	: Solucan
Trojan	: Truva Atı
Ransomware	: Fidye Yazılımı
Spyware	: Casus Yazılım
PSO	: Parçacık Sürü Optimizasyonu
GA	: Genetik Algoritma
SVM	: Destek Vektör Makinesi
RF	: Rasgele Orman
DNN	: Derin Sinir Ağı
XGBoosT	: Extreme Gradient Boosting
DARPA	: Defence Advanced Research Projects Agency
AIDS	: Uygulama Tabanlı IDS
HIDS	: Host Tabanlı IDS
NIDS	: Ağ Tabanlı IDS
GPU	: Grafik İşlemci Birimi
YSA	: Yapay Sinir Ağı
Pooling	: Havuzlama
Flatten	: Düzleştirme
ANN	: Yapay Sinir Ağı
GBM	: Gradient Boosting
DL	: Derin Öğrenme
ELM	: Aşırı Öğrenme Makinesi
GWO	: Gri Kurt Optimizasyon Algoritması
FFA	: Ateş Böceği Optimizasyon Algoritması
TS	: Tabu Arama
DT	: Karar Ağacı
HMI	: İnsan Makine Arayüzü

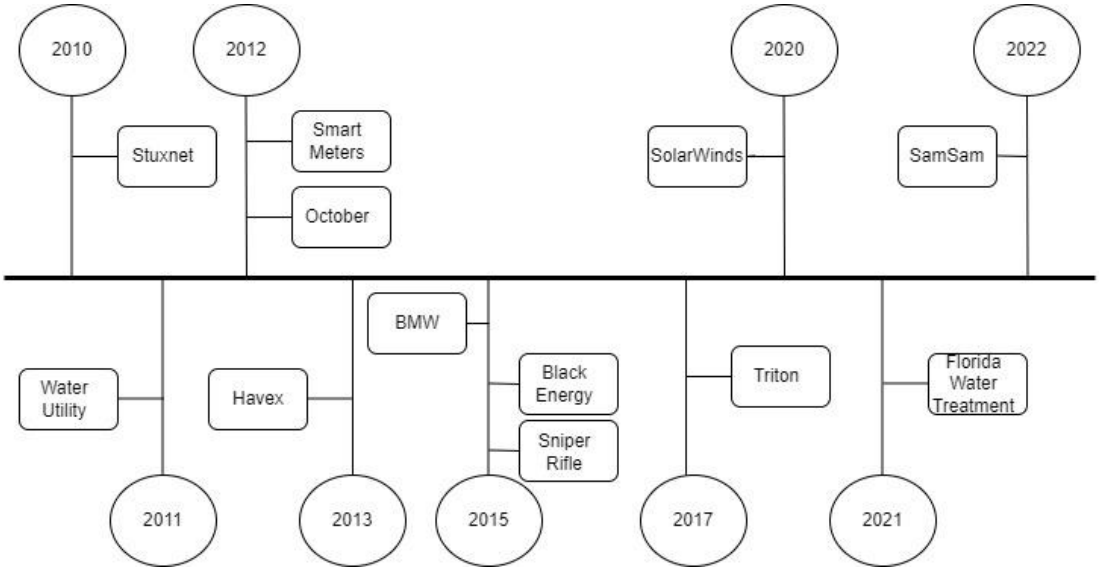
BÖLÜM 1

GİRİŞ

Modern toplumlar, günlük faaliyetlerini yürütmek için gelişmiş siber ve fiziksel altyapılara bağlıdır [1]. Bu altyapılar, hizmetleri yalnızca fiziksel değil, dijital dünyada da korumak için kritik varlıklar olarak da adlandırılır. Günümüzde haberleşme, ulaşım, enerji gibi farklı alanları kapsayan bu altyapıların korunması, ulusal güvenlik endişesi haline gelmiştir [2]. Kritik altyapıların sağladığı hizmetlerin sürekliliği, kontrolü ve güvenliğinin sağlanması maliyetli ve zor bir süreçtir [3].

Denetleyici Kontrol ve Veri Toplama (Supervisory Control And Data Acquisition - SCADA) olarak da adlandırılan kontrol sistemi ile kritik altyapıların işletilmesi sağlanmıştır [4]. SCADA sistemi ile geniş bir alanda kurulan ağlar ve altyapılar tek merkezden yönetilebilmektedir. SCADA sistemi, endüstriyel sistemleri izler ve kontrol eder [5]. SCADA sistemleri, ticari donanım ve yazılımın yanı sıra Ethernet ve TCP/IP gibi açık protokolleri kullanır. Günümüzde SCADA sistemlerinin internete ve kurumsal ağlara bağlanabilmesi siber saldırılarla karşılaşmasına neden olmaktadır [6]. Nükleer santraller, elektrik şebekeleri, su arıtma tesisleri gibi çeşitli alanlarda SCADA sistemlerine yönelik siber saldırılar düzenlenmektedir [7]. Şekil 1.1'de 2010 yılından günümüze tüm dünyada endüstriyel kontrol sistemlerine gerçekleştirilen önemli saldırılar gösterilmektedir.

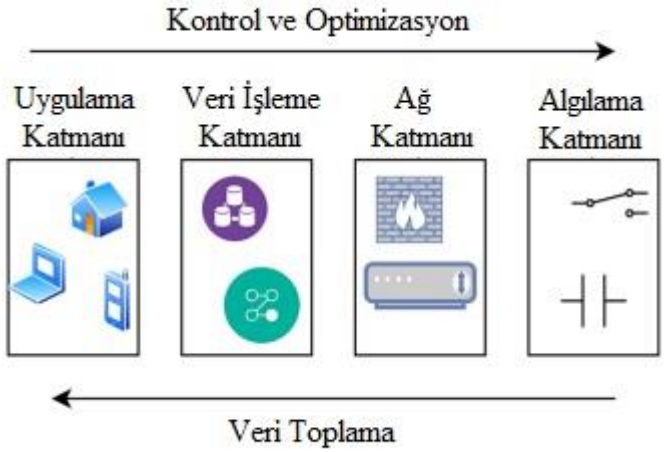
Günlük hayatımızda internet kullanımının, bilgi paylaşımı, kişilerarası iletişim ve etkileşimin artmasını sağladığı gerçektir [8]. Birbirlerini algılayarak iletişime geçen nesnelerin akıllı bağlantısı olarak tanımlanan Nesnelerin İnterneti (IoT – Internet of Things) kavramı günümüzde sıklıkla kullanılmaktadır. IoT ile kablosuz ağ bağlantısına sahip algılayıcılar kullanılarak ağdaki cihazlar üzerindeki veri iletişimi izlenebilmekte ve bu veriler toplanabilmektedir [9].



Şekil 1.1. ICS'ye gerçekleştirilen önemli saldırılar

IoT teknolojisinin endüstride kullanımın başlamasıyla birlikte Endüstriyel IoT (IIoT) kavramı ortaya çıkmıştır [10]. Endüstriyel Nesnelerin İnterneti (IIoT), Nesnelerin İnterneti (IoT) teknolojilerinin, bir üretim ve endüstriyel ortamda siber-fiziksel sistemlerde belirtilen akıllı cihazlarla birlikte kullanılmasını ifade eden terimdir [11]. Nesnelerin İnterneti (IoT), benzersiz şekilde tanımlanabilir "nesneleri" internete bağlayan, algılama, çalıştırma ve potansiyel programlanabilirlik yeteneklerine sahip olan ve "nesneler" hakkında her yerden, her zaman, durum değişikliği de dahil olmak üzere bilgi toplayan bir ağdır [12]. IIoT'de, imalat ve hizmet endüstrileri tarafından cihazları izlemek ve kontrol etmek için kullanılan cihazlara ve ekipmanlara atıfta bulunulur. Şu anda, modern teknoloji şirketleri, imalat ve hizmet endüstrileri için IoT tabanlı bağlantı cihazları ve web uygulamaları tasarlamakta ve geliştirmektedir [13]. Telekomünikasyon, inşaat, madencilik, askeriye, enerji ve sağlık hizmetleri gibi kritik altyapıya sahip sektörler, cihazları ve ekipmanı uzaktan bağlamak, izlemek ve yönetmek için üçüncü taraf IIoT ile ilgili web uygulamalarını ve bağlantı teknolojilerini kullanmaktadır [14].

IIoT, bir aktüatör, sensör, kontrol sistemi, iletişim kanalları, entegrasyon arayüzü, gelişmiş güvenlik sistemleri, araç ağ sistemleri ve akıllı ev ürünleri dahil olmak üzere birçok farklı yapıyı içerir [15]. IIoT içindeki cihazlar internet üzerinden kontrol edilebilir. Endüstriyel nesnelerin internetinin çeşitli endüstrilerde kullanılması, cihaz kalite sistemleri, ürün güvenliği ve yönetim sistemleri, lojistik yönetim sistemleri gibi çeşitli sektörlerin kapasitelerini artırmanın yanı sıra işgücü verimliliğini de önemli ölçüde artırmıştır [16]. Ayrıca IIoT teknolojisi, çeşitli uygulamalara, ara bağlantı bileşenlerine ve ürün desteğine izin vererek fiziksel bir alanı sanal bir alana bağlamaktadır [17]. Hedef cihaza mesaj göndermek için MQTT'yi mesajlaşma protokolü olarak kullanan İnternet Protokolü (IP), PLC'ler arasındaki iletişimi sağlayan Modbus - TCP protokolü ve LoRaWAN gibi verilerin düşük güç tüketimi ile uzun mesafelere iletilmesini sağlayan iletişim protokolleri IIoT'de sıklıkla kullanılmaktadır [18,19]. Ayrıca IIoT cihazlarının büyük bir çoğunluğu gelen verileri alıp işledikten sonra diğer cihazlara iletebilmektedir [20]. Bu özellikler, cihazları endüstriyel nesnelerin interneti sistemlerini ve ait oldukları uygulamaları riske atma potansiyeline sahip bazı gizlilik ve güvenlik tehditlerine karşı hassaslaştırır [21]. IIoT düğümlerinin en önemli özelliği, veri toplama, işleme ve iletme süreçlerinde sürekli aktif olmalarıdır. IIoT'deki tüm katmanlar, uygulama, veri işleme, ağ ve algılama katmanları olarak adlandırılmaktadır [22]. Kontrol, optimizasyon ve veri akışı bu katmanlar arasında gerçekleştirilir. Ayrıca her katman, çeşitli saldırılara ve izinsiz girişlere eğilimli olduğundan IIoT içindeki sistemleri riske atabilir. IIoT ekosistemindeki en yaygın siber tehditler ve yetkisiz erişim ihlalleri, erişim kontrol saldırıları, verilerin kısmen veya tamamen değiştirilmesi, Servis Dışı Bırakma (DoS) saldırıları ve Dağıtık Servis Dışı Bırakma (DDoS) saldırıları ve arka kapı (backdoor) saldırılarıdır. Bazı kuruluşlar, kötü amaçlı saldırılara direnmek ve IIoT düğümlerinin ve ağlarının güvenliğini korumak için İzinsiz Giriş Tespit Sistemlerini (IDS) kullanır. Genel IIoT mimarisi Şekil 1.2'de gösterilmektedir [23].



Şekil 1.2. Genel IIoT mimarisi.

IDS, ağ üzerinden iletilen verilerin bütünlüğünün, gizliliğinin ve güvenliğinin korunmasını sağlar ve bu nedenle IIoT'de çok önemli bir rol oynar. Bir IIoT ağının saldırılara karşı kısmen veya tamamen önlenmesi veya kötü niyetli etkinliğin önlenmesi, tespit edilmesi, tepki verilmesi ve raporlanması görevleri IDS tarafından gerçekleştirilir [24]. Makine öğrenimi algoritmaları kullanmayan ve büyük veriye dayalı istatistiksel yöntemlere dayalı saldırı tespit sistemlerini çalıştırmak zordur. İstatistiksel yöntemlerle oluşturulan IDS'ler davranış temellidir. Önceki kurallara dayalı özetler, ağdaki davranışları modellemek amacıyla kullanılır. Ancak bu özetleri ölçmenin zor olması nedeniyle günümüzde bu yöntem tercih edilmemektedir. Makine öğrenimi ise yapay zekanın bir alt dalıdır [25]. İyileştirme yeteneği ve kapasitesi ile çeşitli sistemlerin deneyimlerden öğrenme ve karar verme süreçlerini herhangi bir açık programlama olmaksızın güçlendirebilir. Genel olarak, makine öğrenimi yaklaşımları denetimli ve denetimsiz olmak üzere ikiye ayrılır [26]. Daha ayrıntılı olarak ise, makine öğrenimi algoritmaları, denetimli, denetimsiz, yarı denetimli ve pekiştirme olarak sınıflandırılır [27]. Denetimli makine öğrenimi yöntemleri, gelecekteki tahminleri belirlemek için tanımlı bir veri kümesinden öğrenme işlemini gerçekleştirir. Ve karar verme süreçlerini iyileştirir. Buna karşılık, tanımlanmamış veriler için denetimsiz makine öğrenimi yaklaşımları kullanılır. Yarı denetimli öğrenme makinelerinde tüm veriler kullanılabilir. Ayrıca algoritma başlarken kurallar sağlanmamış olabilir. Pekiştirme makine öğrenimi yaklaşımı ise, belirli bir ortamdaki etkileşimlerine dayalı olarak ödülleri veya hataları hesaplayan yöntemdir [28].

Geleneksel IDS'lerin sınıflandırılması genellikle imza tabanlı, anomali tabanlı ve hibrit IDS'ler olarak sunulur. İmza tabanlı IDS'ler temel olarak davetsiz misafirlerin davranış modellerini çıkarır. Saldırı imzaları adı verilen bu modellerde daha önce karşılaşılan saldırıların karakteristik özellikleri analiz edilir [29]. Oluşturulan veritabanındaki izinsiz giriş imzaları ile kötü amaçlı yazılım davranışı eşleştiğinde, bu davranışlar izinsiz giriş olarak algılanır. İmza tabanlı IDS'ler yanlış pozitif değerleri üretmez; ayrıca veritabanında imzası bulunan her izinsiz girişi de algılarlar. İmza tabanlı izinsiz giriş tespit sistemlerinin sorunu, imzası bilinmeyen izinsiz girişlerin tespit edilememesi ve buna bağlı olarak yüksek oranda yanlış negative değer üretilmesidir [30]. Yanlış negatif oranını azaltmak için izinsiz giriş imzalarıyla oluşturulan veri tabanının güncellenebilir olması gerekir. Anomali tabanlı IDS'ler ise ağdaki anormal olaylar ile normal olayları ayırt etme prensibine dayanmaktadır [31]. Anomali tabanlı IDS'lerde öncelikle sistemdeki kullanıcıların davranış profilleri belirlenir. Normal davranışlardan farklı olan davranışlar, anormal davranışlar olarak tanımlanır. Normal davranış profillerinin doğru tespit oranı ne kadar yüksek olursa, anormal davranışların doğru tespit oranı da o kadar yüksek olur. Anomali tabanlı IDS'lerde normal davranışlar sürekli olarak güncellenir [32]. Bu nedenle daha önce karşılaşılmayan saldırılar bu yöntemle tespit edilebilmektedir. Anomali tabanlı IDS'lerin dezavantajı, çeşitli izinsiz giriş türlerinde yüksek oranda yanlış alarm üretme olasılığıdır. Hibrit IDS'ler, imza ve anormallik tabanlı IDS'lerin olumsuz yönlerini azaltmak için tasarlanmıştır [33].

Yüksek yanlış pozitif oranları ve düşük tespit doğruluğu, geleneksel IDS'lerde izinsiz giriş tespiti sonuçlarının doğruluğunu azaltır. Ayrıca Slowloris DoS saldırıları gibi olayları da engelleyemezler [34].

1.1. ÇALIŞMANIN AMACI

Yukarıdaki sorunlar göz önüne alındığında, saldırı çeşitlerinin ve veri miktarının her geçen gün arttığı kritik altyapı sistemlerinde izinsiz girişlerin yüksek başarımla tespit edilerek, atak paketlerinin türlerinin belirlenmesi için derin öğrenme tabanlı saldırı tespit sistemi tasarımı, bu tez çalışmasının ana konusunu oluşturmaktadır. Önerilen model, IIoT ağına gerçekleştirilen izinsiz giriş paketlerini, hem ikili sınıflandırma hemde çok sınıflı sınıflandırmaya tabi tutmaktadır. Yukarıda belirtilen iyileştirmeleri

sağlayabilmek için modelde kullanılan derin öğrenme hiperparametreleri uygun biçimde ayarlanmıştır.

1.2. LİTERATÜR İNCELEMESİ

Bu bölümde kritik altyapıya sahip endüstriyel ağlara gerçekleşen siber saldırıların tespit edilmesine yönelik önerilen saldırı tespit sistemleri ile ilgili literatür çalışması açıklanmıştır.

Hawawreh vd. karmaşık IIoT ağları için X-IIoTID adlı bir veri seti oluşturmuşlardır. Bu veri seti farklı marka cihazlardan ve bu cihazların değişik platformlar ile bağlı olduğu diğer cihazlardan gelen verilerin toplanmasıyla elde edilmiştir. Çalışmada sıklıkla kullanılan makine ve derin öğrenme algoritmalarının X-IIoTID veri setindeki atak çeşitlerini algılama oranı belirlenmiştir. Ayrıca çalışmada kullanılan algoritmaların elde ettiği sonuçlar 18 farklı veri seti ile karşılaştırılmıştır. X-IIoTID veri setinin IIoT saldırı veri seti için gereken 20 özelliği elde ettiği ve diğer 18 veri setinden daha üstün olduğu belirlenmiştir [66].

X-IIoTID veri seti kullanılarak gerçekleştirilen başka bir çalışmada veri paylaşımı için güvenli bir model Makkar vd. tarafından önerilmiştir. Önerilen modelde birleşik öğrenme esas alınmıştır. Uç cihazlar fikir birliği hesaplamasına dahil edilmiştir. Önerilen model derin öğrenme algoritmaları ile denenmiştir. Açıklanan sonuçlara göre yüksek verim elde edilmiştir. SecureIIoT adı verilen modelin ikili sınıflandırmada %99.79 doğruluk değerini elde ettiği belirtilmiştir [67].

Diğer bir çalışmada IIoT sistemlerindeki cihazlara yönelik gerçekleştirilen fidye yazılımı saldırılarının önlenmesine yönelik bir model Hawawreh vd. tarafından sunulmuştur. İki kısımdan oluşan modelde, ilk olarak verilerek otomatik kodlayıcı tarafından temizlenerek daha iyi temsil edilmesi sağlanır. Daha sonra ise derin sinir ağı ve toplu normalleştirme kullanılarak saldırıların algılanması ve karar verilmesi sağlanmıştır. Önerilen bu model X-IIoTID, ISOT ve NSL-KDD veri setleri ile ayrı ayrı denenmiştir. Çalışma sonuçlarında önerilen modelin IIoT sistemlerindeki

cihazlara yönelik hedefli fidye yazılımlarını yüksek oranda tespit ettiği belirtilmiştir [68].

UNSW-NB15 ve KDD99 veri setleri kullanılarak yapılan bir çalışmada, beklenti maksimizasyonu, kümeleme algoritması ve yapay sinir ağları yöntemleri olmak üzere çeşitli teknikler ile karşılaştırma işlemi Moustafa vd. tarafından gerçekleştirilmiştir. Bu çalışmada, modellerin değerlendirilmesinde yanlış alarm oranı ve doğruluk değeri kullanılmıştır. KDD99 veri setinde, beklenti maksimizasyonu doğruluk değeri %78.06, yanlış alarm oranı ise %23.79 olarak açıklanmıştır. Diğer veri setinde ise beklenti maksimizasyonu ile %23.79 bir yanlış alarm oranı ve %78.47'lik bir doğruluk değeri elde edildiği belirtilmiştir. Ayrıca, yapay sinir ağları tekniği, UNSW-NB15 veri seti üzerinde test edildiğinde %81.34 doğruluk ve %21.13 yanlış alarm oranı değerlerini elde etmiştir. UNSW-NB15 veri kümesinin diğer veri kümesinin aksine daha dengesiz olduğu belirtilmiştir. [69].

Başka bir çalışmada Kasongo tarafından, IIoT için bir saldırı tespit sistemi önerilmiştir. Bu modelde öznitelik seçimi için Genetik Algoritma (GA) kullanılmıştır. Genetik algoritma uygunluk fonksiyonunda Rastgele Orman (RF) modeli kullanılmıştır. Saldırı tespitinin gerçekleştiği süreçlerde sınıflandırıcı yöntemi olarak Ekstra Ağaçlar, Random Forest, Lineer Regresyon (LR), Naive Bayes (NB), Aşırı Gradyan Artırma ve Karar Ağacı kullanılmıştır. Genetik algoritma - random forest modeli, ikili sınıflandırmada 10, çok sınıflı sınıflandırmada ise 7 özellik vektörü üretmiştir. UNSW-NB15 veri seti kullanılarak gerçekleştirilen uygulamada, doğruluk ikili sınıflandırma işlemi için, %87.61, Auc ise 0.98 olarak elde edilmiştir. Bu sonuçların elde edildiği genetik algoritma-random forest modelinde 16 özellik bulunmaktadır [70].

Liu vd. [71], IoT tabanlı bir saldırı tespit sistemi sunmuşlardır. Bu modelde özellik seçimi için Parçacık Sürü Optimizasyonu (PSO) kullanılmıştır. Saldırıların sınıflandırılmasında ise Destek Vektör Makinesi (SVM) algoritması tercih edilmiştir. Light Gradient Boosting Machine (LightGBM) algoritması bu araştırmada kullanılan parçacık sürü optimizasyonu yöntemine temel oluşturmaktadır. Önerilen modeller UNSW-NB15 veri seti üzerinde test edilmiştir. Performans ölçütü olarak doğruluk

değeri ve yanlış alarm oranı belirlenmiştir. Sonuçlar incelendiğinde, önerilen modelin doğruluk oranı %86.68, yanlış alarm oranı ise %10.62 olarak bulunmuştur. Bu sonuçlar ikili sınıflandırmada önerilen modelin literatürdeki diğer çalışmalara göre yüksek bir yanlış alarm oranına sahip olduğunu göstermektedir. Yazarlar, önerdikleri modeli çok sınıflı sınıflandırma işleminde uygulamamışlardır.

Endüstriyel alanda kullanılan büyük veri sistemleri için, değişken uzun kısa süreli bellek (VLSTM) tabanlı saldırı tespit sistemi, Zhou ve arkadaşları tarafından sunulmuştur. Bu modelde, öznitelikler yeniden oluşturulmuştur. Ardından oluşturulan bu yeni öznitelikler içerisinden seçim işlemi yapılmıştır. Özniteliklerin çıkarılmasında otomatik kodlayıcı kullanılmıştır. Bu sayede karmaşık ve büyük veri kümesinden öznitelikler çıkarılmıştır. Araştırmacılar, bu uygulamada UNSW-NB15 veri setini kullanmışlardır. Önerilen modelin performansını değerlendirirken yanlış alarm oranı, eğri altındaki alan, hassasiyet, geri çağırma ve F1-Skoru ölçütlerini kullanmışlardır. Değişken uzun kısa süreli bellek yönteminin eğri altında kalan alan değeri 0.895, hassasiyet değeri %86, geri çağırma değeri %97.8 ve F1-Skor değeri ise %0.7 olarak elde edilmiştir. Bu sonuçlar literatürdeki bazı çalışmalarda elde edilen değerlerden yüksektir. Fakat kullanılan veri seti içerisindeki saldırı çeşitlerinin düzgün dağılmaması nedeniyle daha fazla uygulama yapılması, daha doğru sonuçların elde edilmesini sağlayacağını yazarlar belirtmişlerdir [72].

Başka bir çalışmada, Gao vd. Tarafından aşırı öğrenme makinesi tabanlı bir saldırı tespit sistemi önerilmiştir. Bu modelde özellik seçimi için uyarlanabilir bir temel bileşen kullanılmıştır. Uyarlanabilir temel bileşen tarafından seçilen özellikler sınıflandırma işlemini gerçekleştirilmesi için aşırı öğrenme makinesine sunulur. Önerilen model NSL-KDD ve UNSW-NB15 veri kümelerinde test edilmiştir. Ayrıca, her iki veri seti için çok sınıflı sınıflandırma işlemi gerçekleştirilmiştir. Test verileri üzerinde elde edilen doğruluk değeri performans metriği olarak kabul edilmiştir. NSL-KDD veri setinde önerilen model %81.22'lik bir doğruluk elde etmiştir. UNSW-NB15 veri setinde ise %70.51 doğruluk oranı elde edilmiştir. Literatürdeki diğer çalışmalara göre daha iyi sonuçlar elde edildiği açıklanmıştır. Ayrıca gerçek hayattaki endüstriyel kontrol sistemlerinde doğruluk oranının yüksek çıkması için daha fazla araştırma yapılması gerektiği de belirtilmiştir [73].

Vinayakumar vd. tarafından başka bir çalışmada, DNN tabanlı bir IDS hazırlanmıştır. Bu araştırmadaki amaçları, yeni saldırı biçimlerini hızlıca tespit edebilen, farklı platformlarda kullanılabilen bir IDS ortaya çıkarmaktır. Araştırmacılar yöntemlerinin başarımını değerlendirmek için 6 farklı veri seti kullanmışlardır. Bunlar, Kyoto, KDD-Cup99, NSL-KDD, WSN-DS, UNSW-NB15 ve CICIDS 2017 veri setleridir. Her bir veri seti için model 1000'in üzerinde epoch yürütülmüştür. Güncel bir veri seti olan UNSW-NB15'i referans kabul eden çalışmada, derin öğrenme ağının ikili sınıflandırmada %76.1 doğruluk, %95.1 hassasiyet, %96.3 geri çağırma ve %79.7 F1-Skoru elde ettiği tespit edilmiştir. Buna karşılık, derin öğrenme ağı çok sınıflı sınıflandırma işlemi için ikili sınıflandırmaya göre daha düşük sonuçlar üretmiştir. Çok sınıflı sınıflandırmada %65.1 doğruluk, %75.6 F1-Skoru, %59.7 hassasiyet ve %65.1 geri çağırma elde etmiştir [74].

Başka bir çalışmada Hanif ve arkadaşları tarafından IoT ağları için yapay sinir ağı tabanlı bir saldırı tespit sistemi sunulmuştur. Bu model, IoT ağlarında büyük bir endişe kaynağı olan güvenlik sorununun üstesinden gelmek için uygulanmıştır. Yazarlar tarafından IoT cihazlarının güvenliğinin sağlanmasında genellikle karmaşık hesaplama yapma kapasitesinden yoksun olduğu gerçeği göz önüne alınmıştır. Yazarlar ilk olarak makine öğrenimi algoritmalarını kullanarak saldırı tespit sistemi oluşturmuşlardır. Saldırı sınıflandırma başarımı UNSW-NB15 veri seti üzerinde test edilmiştir. Yapay sinir ağları tabanlı saldırı tespit sistemi ikili sınıflandırma işlemi için %84.00'lük bir doğruluk değeri elde etmiştir. Çalışmada yapay sinir ağının hiper parametrelerinin sonuca varmak için nasıl ayarlandığı yazarlar tarafından net bir şekilde açıklanmamıştır. Ayrıca, çalışmada herhangi bir özellik seçim yöntemi kullanılmamıştır [75].

Ketzaki, modern iletişim sistemlerinin güvenliğini sağlamak için yapay sinir ağı kullanarak bir saldırı tespit sistemi önermiştir. Bu çalışmada önerilen model özellik çıkarma ve sınıflandırma adımlarından oluşmuştur. Özellik çıkarımı için istatistiksel analiz yöntemleri kullanılmıştır. Sınıflandırma işlemi ise yapay sinir ağları ile yapılmaktadır. UNSW-NB15 veri setini kullanılarak ikili sınıflandırma çalışması değerlendirilmiştir. Test verileri ile elde edilen doğruluk oranı, hazırlanan yapay sinir ağı modellerinin başarımını değerlendirmek için kullanılmıştır. Elde edilen sonuçlara

göre %83.9 doğruluk oranının en yüksek başarı değeri olduğunu göstermektedir. Yazar, gelecekte uygulamanın etkinliğini artırmak için farklı modeller ile çalışmayı amaçladığını belirtmiştir [76].

Başka bir çalışmada, Almomani tarafından J48 sınıflandırıcı algoritması ve destek vektör makinesi kullanan bir saldırı tespit modeli sunulmuştur. Çalışmada evrimsel algoritmaların farklı kullanımları ile özellik seçimi gerçekleştirildi. UNSW-NB15 veri seti çalışmada kullanılmıştır. Genetik algoritma-J48, gri kurt iyileştirici-J48 ve J48 modelleri tarafından elde edilen doğruluk değerleri sırasıyla %86.874, %85.676 ve %86.037 olarak elde edilmiştir. Buna ilaveten, Genetik algoritma-destek vektör makinesi, gri kurt iyileştirici-destek vektör makinesi ve destek vektör makinesi tarafından elde edilen doğruluk değerleri sırasıyla %86.387, %84.485 ve %85.429 olarak tespit edilmiştir. Özellik seçiminde J48 ve destek vektör makinesi yöntemleri etkili sonuçlar elde etse de, büyük ve karmaşık veri setlerinde gelecekte yapılacak çalışmaların derin öğrenme yöntemleri gibi diğer yaklaşımlar kullanılarak yapılması önerilmiştir [77].

Nazir vd. farklı bir çalışmada Tabu Arama algoritması ile rastgele orman algoritmasını birlikte kullanarak yeni bir özellik seçim yöntemi geliştirmişlerdir. Bu yöntemde öznitelikler tabu arama algoritması ile aranır. Daha sonra rastgele orman metoduyla öğrenilen özellikler çıkarılır. UNSW-NB15 veri seti bu çalışmada yöntemin değerlendirilmesinde kullanılmıştır. Çalışmada doğruluk değeri ve yanlış pozitif oranı modellerin performansının başarımında kullanılmıştır. Tabu arama - rastgele orman metodu, sınıflandırma işleminde rastgele orman kullanıldığında %83.12 doğruluk ve %3.7 yanlış pozitif oranı elde etmiştir. Araştırmacılar bu çalışmada UNSW-NB15 veri setinde bulunan atak çeşitlerindeki dengesizlik sorununu dikkate almadıklarını belirtmişlerdir [78].

Zong ve arkadaşları tarafından, iki aşamalı bir tabu arama algoritması tabanlı IDS önerilmiştir. Bu yöntemde, ilk aşamada azınlık saldırı sınıfları, ikinci aşamada ise çoğunluk saldırı sınıfları tespit edilmiştir. Bu çalışmada sınıflandırma yöntemi olarak Random Forest kullanılmıştır. Özellik çıkarımı için karar ağacı yöntemi kullanılmıştır. Önerilen modelin değerlendirilmesi için veri seti olarak UNSW-NB15 kullanılmıştır.

Bu araştırmada doğruluk ve yanlış alarm oranı ölçütleri performans değerlendirilmesinde kullanılmıştır. Yazarlar bu çalışmada sadece ikili sınıflandırma işlemi yapmışlardır. Deneysel sonuçlar, önerilen modelin %15.64'lük bir yanlış alarm oranı ve %85.78'lik bir doğruluk değeri elde ettiğini gösterdi. Araştırmacılar bundan sonraki çalışmalarında önerdikleri modeli geliştirmeyi hedeflediklerini belirtmişlerdir [79].

Khammassi vd. tarafından, öznitelik seçimi işleminde Lojistik Regresyon ile genetik algoritma yöntemini kullanan bir model önerilmiştir. Modelde İkili sınıflandırma işlemi ise, C4.5 yöntemi kullanılarak gerçekleştirilmiştir. C4.5 algoritması ağaç tabanlı bir algoritmadır. Yazarlar, önerilen yaklaşımı değerlendirmek için farklı performans ölçütlerini kullanmışşlardır. Buna rağmen test verilerinde temel performans ölçütü elde edilen doğruluk değeri olarak belirlenmiştir. Deney sonunda, yöntem %81.42'lik bir doğruluğa ulaşılmıştır. Bu araştırmada, elde edilen doğruluk değeri literatürdeki diğer çalışmalara göre daha düşük sonuç vermiştir [80].

Başka bir çalışmada, Kasongo ve arkadaşları tarafından farklı makine öğrenme yöntemleri kullanarak bir saldırı tespit sistemi önerilmiştir. Bu modelde özellik çıkarma yöntemi olarak XGBoost (aşırı gradyan artırma) kullanılmıştır. Bu algoritma topluluk temelli bir algoritmadır. Ayrıca sınıflandırma işlemi lineer regresyon yöntemi ile gerçekleştirilmiştir. Deneysel sonuçlar, XGBoosT- Lineer Regresyon modelinin ikili ve çok sınıflı sınıflandırma işlemi için sırasıyla %75.51 ve %72.53 doğruluk elde ettiğini göstermiştir. UNSW-NB15 veri setindeki saldırı çeşitlerinin dengesizliği sorununun giderilmesi için yazarlar, aşırı örnekleme tekniklerinin kullanılmasını önermişlerdir [81].

Jing ve arkadaşları, destek vektör makinesi tabanlı bir ağ saldırı tespit sistemi uygulamışlardır. Bu sistem, IoT ağlarının benzersiz yapısını temsil edecek şekilde tasarlanmıştır. Bu modelin değerlendirilmesi UNSW-NB15 veri seti üzerinde gerçekleşmiştir. Yazarlar, doğruluk, tespit oranı ve yanlış pozitif oranı ve ana performans ölçütlerini dikkate almışlardır. Deneyler ikili ve çok sınıflı sınıflandırma için yapılmıştır. Sonuç, ikili sınıflandırma için %85.99'luk bir doğruluk oranına

ulaştığını göstermiştir. Çok sınıflı sınıflandırma işleminde ise model %75.77 doğruluk elde etmiştir [82].

Başka bir çalışmada Kumar vd. tarafından UNSW-NB15 veri seti kullanılarak çevrimiçi izinsiz giriş tespitini gerçekleştirmek için makine öğrenmesi tabanlı bir saldırı tespit sistemi önerilmiştir. Karar ağacı metodu özellik çıkarımı için kullanılmıştır. Bilgi kazancı yöntemi en iyi olarak 13 özniteliği belirlemiştir. Sınıflandırma süreci için ise araştırmacılar, C5, CHAID, CART ve QUEST ağaç tabanlı sınıflandırıcıları entegre biçimde kullanmışlardır. Sonuç olarak önerilen sistem ikili sınıflandırma işlemi için %84.83'lük bir doğruluk elde etmiştir. Bu çalışmada sunulan saldırı tespit modeli, daha önceden karşılaşmadığı saldırıları tespit edememektedir. Yazarlar gelecekte bu sorunu çözmek için modelde iyileştirmeye gidileceğini açıklamışlardır [83].

Aleesa ve arkadaşları, Uzun Kısa Süreli Bellek (LTSM) ve Recurrent Neural Network (RNN) derin öğrenme yöntemleri ile oluşturulan bir saldırı tespit sistemi önermişlerdir. Yaklaşım UNSW-NB15 veri seti üzerinde değerlendirilmiştir. Buna ilaveten, performans ölçütü olarak doğruluk değeri kullanılmıştır. LSTM yöntemi ikili sınıflandırma işlemi için %85.42 doğruluk elde etmiştir. Yazarlar bu çalışmada literatürdeki diğer çalışmalardan daha iyi sonuçlar elde ettiklerini açıklamışlardır [84].

Elijah ve arkadaşları tarafından [85], topluluk ve derin öğrenme tabanlı bir saldırı tespit sistemi önerilmiştir. Derin öğrenme modeli olarak uzun kısa süreli hafıza algoritması kullanılmıştır. Stokastik Gradyan iniş algoritması, uzun kısa süreli hafıza algoritmasına uygulanan optimizasyon algoritmasıdır. LSTM katmanlarında uygulanan ikili sınıflandırma işleminde Doğrulanmış Doğrusal Birim fonksiyonu kullanılmıştır. Çok sınıflı sınıflandırma işlemi için yazarlar Softmax fonksiyonunu kullanmışlardır. Önerilen model UNSW-NB15 veri seti ile değerlendirilmiştir. LSTM tabanlı saldırı tespit sistemi ikili sınıflandırmada %80.72'lik bir doğruluk değeri elde etmiştir. Buna karşılık, LSTM tabanlı saldırı tespit modeli, çok sınıflı sınıflandırmada ikili sınıflandırmaya göre daha düşük bir değer olan %72.26 doğruluk elde etmiştir.

Başka bir çalışmada Wu ve arkadaşları tarafından, derin sinir ağları kullanılarak saldırı tespit sistemi önerilmiştir. Bu modelde, artık bloklar (ResBlk) kullanılmıştır. Bu blokların görevi bir sonraki katmanı beslemektir. Çalışmada iki farklı veri seti kullanılmıştır. Yaklaşımın değerlendirilmesinde doğruluk değeri performans ölçütü olarak belirlenmiştir. Artık bloklar kullanılarak gerçekleştirilen çalışmanın sonuçlarına göre NSL-KDD veri setinde %99.21, UNSW-NB15 veri setinde ise %86.64 doğrulukla saldırılar tespit edilmiştir. Mevcut performans rakamlarını iyileştirmek ve sürekli hale getirebilmek için daha fazla deney yapılması gerektiği yazarlar tarafından belirtilmiştir [86].

Assiri [87], genetik algoritma ve rastgele orman yöntemi kullanarak anomali tabanlı saldırı tespit sistemi önermiştir. Genetik algoritma kullanılarak nitelik ve parametre seçimi yapılmıştır. Sınıflandırma işlemi ise rastgele orman yöntemi ile gerçekleştirilmiştir. Ayrıca, araştırmacılar sadece ikili sınıflandırma işlemi yapmışlardır. UNSW-NB15veri seti önerilen modelin performansını değerlendirmek için kullanılan veri kümelerinden biridir. Doğruluk, geri çağırma ve kesinlik ölçütleri, önerilen modeli değerlendirmek için kullanılan performans ölçütleridir. Deneysel sonuçlar, modelin %87 geri çağırma ve 87% hassasiyet ile %86.70'lik bir sınıflandırma doğruluğuna ulaştığını göstermiştir.

Khammassi ve arkadaşları tarafından genetik algoritma ve lojistik regresyon tabanlı saldırı tespit modeli geliştirilmiştir. Bu sistemde, çok amaçlı bir öznitelik seçim yöntemi kullanılmıştır. Random forest yöntemi, önerilen yöntemin performansını değerlendirmek için kullanılan makine öğrenmesi yöntemlerinden biridir. Çalışmada UNSW-NB15'te dahil olmak üzere farklı veri setleri kullanılmıştır. Modelin başarımı doğruluk değeri ile belirlenmiştir. Deneysel sonuçlar, önerilen modelin çok sınıflı sınıflandırma görevi için %64.23'lük bir doğruluk elde ettiğini göstermektedir [88].

Literatürdeki çalışmaların incelenmesinden sonra hibrit tabanlı derin öğrenme algoritmalarını kullanan bir saldırı tespit sistemi modelinin eksik olduğu sonucuna varılmıştır. Önerilen yöntemler yüksek doğruluk değerine ulaşsa da, asıl amaç büyük ve güncel veri setlerinde hem ikili hem de çok sınıflı sınıflandırma görevlerini yüksek başarımla yerine getirebilecek bir modelin oluşturulabilmesidir.

Gerçekleştirilen literatür çalışmasından elde edilen sonuçlara göre kritik altyapılara gerçekleştirilen siber saldırılar yüksek doğruluk değerlerine sahip olacak şekilde engellenmelidir. Ayrıca veri önişleme işlemi gerçekleştirilerek saldırı paketlerindeki kayıp değeri en aza indirilmelidir. Bu sorunların çözümü için derin öğrenme tabanlı hibrit bir saldırı tespit sistemi öneriyoruz. Veriler üzerinde normalizasyon işlemi uygulandıktan sonra, derin öğrenme mimarilerindeki hiper parametreleri kullanarak ikili ve çok sınıflı sınıflandırma gerçekleştirip, veri setleri içerisindeki atak çeşitlerinin tespit edilme oranları belirlenmiştir. Veri miktarı fazla ve özniteliklerin çok olduğu karmaşık veri setlerini tercih ettiğimiz için derin öğrenme modelleri bu çaışmada tercih edilmiştir.

1.3. TEZE GENEL BAKIŞ

Hazırlanan tez çalışması altı bölümden oluşmaktadır. Birinci bölümde, teze genel bir giriş yapılarak tez çalışmasının amacı açıklanmış veliteratür incelemesi gerçekleştirilmiştir.

İkinci bölümde siber saldırı kavramı tanımlanıp saldırı çeşitleri hakkında bilgi verilmiştir.

Üçüncü bölümde, saldırı tespit sistemlerinin genel yapısı ve saldırı tespit sistemlerinde kullanılan derin öğrenme algoritmaları açıklanmıştır.

Dördüncü bölümde, tez çalışmasında kullanılan veri setleri açıklanmıştır.

Beşinci bölümde, önerilen model detaylıca açıklanmış ve gerçekleştirilen deneysel çalışma gösterilerek kullanılan değerlendirme ölçütlerine göre elde edilen sonuçlar açıklanmıştır.

Altıncı bölümde, tez çalışması sonucunda elde edilen sonuçlar yorumlanarak tartışılmıştır.

BÖLÜM 2

SİBER SALDIRI VE SİBER SALDIRI TÜRLERİ

Büyük verinin, sosyal ağların, çevrimiçi işlemlerin, internet aracılığıyla depolanan veya yönetilen bilgilerin ve bilgi teknolojileri sistemleri kullanılarak gerçekleştirilen otomatik süreçlerin yönlendirildiği bir dünyada, bilgi güvenliği ve veri gizliliği sürekli olarak risklerle karşı karşıyadır. Yeni araç ve tekniklerin geliştirilmesiyle birlikte siber suç, saldırı sayısı ve kurbanlarına verdiği zarar düzeyi açısından sürekli olarak artmaktadır.

Ağlara, programlara ve verilere yetkisiz erişim sağlamak için yeni yollar geliştiren saldırganlar, tek tek bireylerden küçük veya orta ölçekli şirketlere ve hatta kritik altyapılara kadar hedeflerini oluşturarak bilgilerin gizliliğini, bütünlüğünü ve erişilebilirliğini tehlikeye atmayı amaçlıyor. Her geçen gün genel olarak daha fazla sayıda siber saldırı ile karşılaşılıyor. Aynı zamanda büyük şirketlerin ve kritik altyapıların güvenliğini bozan ve böylece bilgi güvenliğini, iş sürekliliğini ve müşterilerin güvenini etkileyen daha fazla sayıda saldırı gerçekleşmektedir.

2.1. SİBER SALDIRI KAVRAMI

Siber saldırı, içinde bulunan kritik verileri çalmak, değiştirmek veya yok etmek amacıyla bilgisayarı veya bilgisayar ağını hedef alan bir saldırı türüdür [35]. Saldırgan, yetkisiz erişim veya kullanım elde eden herhangi bir birey veya süreç olabilir. Siber saldırı, bir kişi veya gruplar tarafından gerçekleştirilebilir [36]. Siber saldırının amacı, bir kişinin veya bir yönetimin bilgi sistemini ele geçirmektir. Siber saldırı, kötü amaçlı kod kullanır ve bu nedenle bilgisayar verilerini, kodunu veya mantığını değiştirir. Bu, yıkıcı etkilere ve verilerin tehlikeye atılmasına neden olur. Ayrıca bilgi ve kimlik hırsızlığı gibi siber suçlara yol açar. Siber saldırılar genel olarak dört farklı aşamada

gerçekleştirilir. İlk aşamada hedefte ilgili bilgiler toplanır. İkinci aşamada hedefteki bir güvenlik zafiyeti tespit edilerek hedefe sızma işlemi gerçekleştirilir. Üçüncü aşamada hedef sistemde yetki yükseltilme işlemi gerçekleştirilir. Son adımda ise saldırgan hedef sistem içerisinde faaliyetlerini gerçekleştirir [37]. Artan veri miktarı ve çeşitliliğine bağlı olarak karşılaşılan siber saldırıların türleri de değişiklik göstermektedir. Fakat siber güvenliğin özünde CIA üçlüsü olarak adlandırılan bilginin gizliliğini (Confidentiality), bütünlüğünü (Integrity) ve erişilebilirliğini (Availability) korumak bulunmaktadır CIA üçlüsü Şekil1.3'te gösterilmiştir.

Şekil 2.1. CIA üçlüsü.

Gizlilik kısaca mahremiyete eşdeğerdir. Gizlilik önlemleri, hassas bilgileri, yetkisiz erişim korumak için tasarlanmıştır. Verilerin, yanlış ellere geçmesi durumunda verilebilecek hasarın miktarına ve türüne göre sınıflandırılması yaygındır.

Bütünlük, verilerin tüm yaşam döngüsü boyunca tutarlılığını, doğruluğunu ve güvenilirliğini korumayı içerir. Veriler aktarım sırasında değiştirilmemeli ve verilerin yetkisiz kişilerce değiştirilemeyeceğinden emin olmak için adımlar atılmalıdır.

Erişilebilirlik, bilgilerin yetkili taraflar için tutarlı ve kolayca erişilebilir olması gerektiği anlamına gelir. Erişilebilirlik kavramı, bilgileri tutan ve görüntüleyen donanım ve teknik altyapı ile sistemlerin uygun şekilde bakımını içerir.

Bu bilgilerin ışığında, siber saldırı, deneyimli ve yetenekli bir saldırgan tarafından gerçekleştirilirse, tekrarlanan birçok yöntemi içerebilir. Bu nedenle, farklı saldırı türlerini ve bunlarla ilgili aşamaları anlayarak kişi kendisini saldırıdan koruyabilir. Saldırılar hedefli ve hedefsiz saldırılar olmak üzere iki gruba ayrılabilir [38].

2.1.1 Hedefli Saldırı

Hedefli saldırıda, saldırganın belirli bir kuruluşla özel bir ilgisi vardır veya bu tür bir kuruluşu hedef alması için başkaları tarafından yönlendirilmiştir. Bu tür bir saldırının hazırlanması, sistemde, açıktan yararlanmanın en iyi yolunu bulmak için uzun zaman alabilir. Hedefli saldırı, hedeflenmemiş saldırıdan daha fazla tehdide neden olmaktadır. Hedefli saldırılar özel olarak gerçekleştirilir. Hedef odaklı kimlik avı, bir botnet dağıtma, tedarik zincirini alt üst etme vb. örnekler içerir [39].

2.1.2 Hedefsiz Saldırı

Hedefsiz saldırılarda saldırgan, geniş çapta olabildiğince çok cihazı veya kullanıcıyı hedefler. Burada saldırgan, internet teknolojisinin açıklarından faydalanabilir. Kimlik avı, fidye yazılımı, gibi saldırılar örnek olarak verilebilir [40].

2.2. SİBER SALDIRI TÜRLERİ

2.2.1 Servis Dışı Bırakma (DoS) ve Dağıtık Servis Dışı Bırakma saldırısı (DDoS)

Bir hizmet reddi saldırısı, hizmet talebini yanıtlayamayacak şekilde sistem kaynaklarının aşılmasıyla ortaya çıkmaktadır [41]. Bir saldırgan tarafından kontrol edilen kötü amaçlı yazılımlardan etkilenen ana makine, DDoS saldırısı başlatır. Bu tür siber saldırılarda, internete bağlı olan hostun hizmeti aksatılarak, makine veya ağ kaynakları hedeflenen kullanıcı tarafından kullanılamaz hale getirilir [42]. TCP SYN flood saldırısı, smurf saldırısı, ping-of-death saldırısı ve botnet'ler DoS ve DDoS saldırılarının farklı türleridir [43].

Aynı port ve protokolü kullandıklarından ve meşru bir trafik talebini kötü niyetli bir trafik talebinden ayırt etmek zor olduğu için, DoS saldırısını önlemek zaman almaktadır [44].

2.2.2 Ortadaki Adam Saldırısı (MiTM)

Bir istemci ile sunucu arasındaki iletişim arasına üçüncü bir taraf girdiğinde bir MitM saldırısı gerçekleşir. Üçüncü taraf, hem istemciyi hem de sunucuyu taklit eder ve aralarındaki bilgilere erişim sağlar [45]. Bu tür bir saldırı, bir tehdit aktörünün başkasına yönelik verileri ele geçirmesine, göndermesine ve almasına neden olur. Bir MITM saldırısı, işlemlerin, iletişimin veya diğer bilgilerin değiş tokuşunun gerçek zamanlı işleyişini kötüye kullanır [46]. Ortadaki adam saldırısının farklı türleri, oturum ele geçirme, IP sahtekarlığı ve onay mesajlarının değiştirilmesini içerir. Ortadaki adam saldırısını önlemek için bir saldırı tespit sistemi kurulabilir. Birisi ağ akışını ele geçirmeye çalışırsa anında uyarı vermeye yardımcı olur. Ortadaki adam saldırısını önlemek için sanal özel ağ da kullanılabilir. Bu, bir şirketin gizli katmanına Wi-Fi aracılığıyla erişirken ek güvenli katmanlar oluşturmaya yardımcı olur [47].

2.2.3 Oltalama Saldırısı (Phishing)

Kimlik avı saldırısı, güvenilir kaynaklardan geliyormuş gibi görünen sahte e-postaların gönderilmesiyle gerçekleşir. Bu tür saldırıların asıl amacı, kişisel bilgileri ele geçirmektir. Kimlik avı saldırısı, bir sosyal mühendislik tekniği biçimidir. Sisteme kötü amaçlı yazılım yükleyen gömülü bağlantılardan oluşan e-postalar biçimindedir. Bazen bu bağlantı, kötü amaçlı yazılım indirmemize veya kişisel bilgilerimizden vazgeçmemize neden olan meşru olmayan bir web sitesine de yönlendirme sağlayabilir. Hassas verileri elde etmek için kimlik avı saldırısı bazı medya araçlarından, mesajlardan, telefondan vb. yollardan gelmektedir [48].

Kimlik avı saldırısı riskini azaltmak için eleştirel düşünme, bağlantıların üzerinde gezinme, e-posta başlıklarını analiz etme ve sandboxing adı verilen yalıtılmış bir test

ortamı kullanılabilir. Ayrıca bireyler kadar kurum çalışanlarını da bilinçlendirerek phishing saldırılarının engellenme oranı arttırılmaktadır [49].

2.2.4 Sürücüden Yükleme saldırısı (Drive – by – download)

Drive-by-download saldırısı, siber suçlular tarafından kötü amaçlı yazılım yaymak ve yetkisiz erişim elde etmek için gerçekleştirilen yaygın bir siber saldırı türüdür. Bu saldırı, bir bilgisayara yalnızca bir web sitesini ziyaret ederek kötü amaçlı bir yazılım bulaştığında gerçekleşir. Kullanıcının virüs kapmak için herhangi bir yeri tıklamasına gerek yoktur, bu nedenle buna "drive-by" indirme saldırısı denir [50]. Burada suçlular genellikle meşru bir web sitesi kullanır ve web sayfalarının içine kötü amaçlı bir nesne yerleştirir. Kullanıcılar bulaşmaları gözlemleyemez. Kötü amaçlı JavaScript kodundan, iFrame'lere, bağlantılara, yönlendirmelere, siteler arası komut dosyası çalıştırmaya ve diğer kötü amaçlı öğelere kadar değişik yöntemleri vardır [51]. Bir kullanıcı virüslü web sayfasını ziyaret ettiğinde, kullanıcının tarayıcısına kötü amaçlı kodlar otomatik olarak yüklenir. Ardından işletim sistemindeki ve diğer uygulamalardaki bilgisayar güvenlik açıkları otomatik olarak taranır.

Yazılımın hızlı ve düzenli bir şekilde güncellenmesi, istenmeyen yazılım uygulamalarının ve tarayıcı eklentilerinin güvenlik duvarı ve web filtreleme yazılımları kullanılarak kaldırılması, sürücüden yükleme saldırılarının önlenmesi için kullanılabilir [52]. Ayrıca, ayrıcalıklı bir hesap kullanarak internette gezinirken, her türlü kötü niyetli yazılım, herhangi bir açık izin olmaksızın bir sisteme kendi kendine girebilir. İki ayrı hesap tutularak sisteme bu tür girişler engellenebilir. Biri günlük aktiviteler için, diğeri ise yazılım yüklemek için yönetici hesabı için kullanılabilir [53].

2.2.5 Parola Saldırısı (Password Attack)

Kullanıcının kimliğini doğrulamanın en yaygın yöntemi parola kullanmaktır ve bu parolaları izinsiz elde etmek, etkili bir saldırı yaklaşımıdır. Parola saldırısı, kullanıcının parolasının meşru olmayan yollarla elde edildiği veya parolasının çözüldüğü tekniktir [54]. Kullanıcı masasına bakarak, tahmin ederek, , ağ bağlantısını koklayarak (sniffing), düz metin şifreyi elde ederek gibi farklı yollarla parola atakları

gerçekleştirilir. Parolaları sık sık değiştirerek, tahmin edilmesi zor ve uzun parolalar kullanarak, parola saldırısının etkisinden korunulabilmektedir. Kaba kuvvet (Brute force) ve sözlük saldırısı (Dictionary Attack), parolanın elde edilebileceği iki ana tekniktir [55].

2.2.6 SQL Enjeksiyonu Saldırısı (SQL Injection)

SQL (Structured Query Language), veritabanında bulunan verileri, depolamak, değiştirmek ve almak için kullanılan bir bilgisayar programlama dilidir. SQL dili, gerekli görevi gerçekleştirmek için seç, güncelle, sil gibi komutları kullanır [56]. SQL ayrıca veritabanına karşı sorgular yürütebilir, veritabanına kayıtlar ekleyebilir ve veritabanında yeni tablolar oluşturabilir. SQL Injection saldırısı, veritabanını manipüle ederek bilgilere erişmek için kötü amaçlı koddan yararlanır. Bu bilgiler, hassas kuruluş bilgilerini, müşteri veya kullanıcının özel verilerini içerebilir. Bu durum, kullanıcı verilerinin yasa dışı olarak görüntülenmesine, tablo verilerinin silinmesine ve yetkisiz veritabanı saldırılarına neden olabilir [57].

SQL enjeksiyonu gerçekleştirmek isteyen bir saldırgan, veritabanındaki doğrulanmamış güvenlik açıklarından yararlanmak için standart bir SQL sorgusunu manipüle eder. Saldırganlar, SQL komutlarını değiştirmek için yanlış filtrelenmiş karakterler de kullanabilir [58]. Oluşmaları durumunda SQLI saldırılarını önlemenin ve bunlara karşı koruma sağlamanın birkaç etkili yolu vardır. Yasa dışı kullanıcı girişlerini belirlemek için giriş doğrulaması yapılabilir. Ancak bu yöntem, yasal ve yasa dışı tüm girdilerin eşleştirilmesi mümkün olmadığı için pek uygun değildir. Bu nedenle, SQLI'yi kaldırmak için genellikle bir güvenlik duvarı (WAF) kullanılır. İmza tabanlı saldırı tespit sistemi de SQL enjeksiyonlarını tanımlamak ve engellemek için kullanılabilir [59].

2.2.7 Cross Site Scripting (XSS) Saldırısı

Cross site scripting, güvenilir bir web sitesine veya hassas bir web uygulamasına kötü amaçlı kod ekleyen yaygın bir enjeksiyon saldırısı türüdür [60]. Başka bir deyişle, saldırgan web sitesinin veritabanına kötü amaçlı bir kod veya JavaScript enjekte

ettiğinde XSS zafiyeti oluşur. Davetsiz misafir, son kullanıcının web sayfasına kötü amaçlı JavaScript kodu enjekte eder ve web sayfasını indirmesini sağlar [61]. Kurbanın tarayıcısı, yanıt içindeki kötü amaçlı komut dosyasını çalıştırarak kurbanın çerez bilgilerini saldırganın sunucusuna gönderir. Üç ana XSS saldırısı türü vardır Bunlar, Persistent XSS, Reflected XSS ve DOM tabanlı XSS'di. Persistent XSS'de, kötü amaçlı kod web sitesinin veritabanından kaynaklanırken, Reflected XSS türünde, kötü amaçlı kod kurbanın isteğinden kaynaklanır. DOM tabanlı XSS ise, yukarıda belirtilen yöntemlerin bir alternatifidir. Burada güvenlik açığı sunucu tarafında değil istemci tarafında bulunmaktadır. Cross site scripting çalıştırma, kodlama veya doğrulama yoluyla önlenebilir [62].

2.2.8 Malicious Software (Malware) Saldırısı

Malware saldırısı, kötü amaçlı yazılımların izinsiz bir şekilde, kullanıcının bilgisayarına yüklendiği bir siber saldırı türüdür [63]. Bu siber saldırı türünde, virüs, casus yazılım veya fidye yazılımı gibi kötü amaçlı yazılımlar ağa erişebilir, belirli bilgi işlem süreçlerini kesintiye uğratabilir, hassas bilgileri veya diğer kullanıcı verilerini çalabilir. Ayrıca bu yöntemle hedeften yasadışı olarak para kazanabilir. Günümüzde kötü amaçlı yazılımlar, herhangi bir kimlik bilgisinden çok ticari veya finansal bilgileri hedeflemektedir [64].

En yaygın kötü amaçlı yazılım türü şunlardır:

a-)Virüs: Herhangi bir bilgisayar programına eklenen, çalıştırıldığında kodları çoğaltan ve değiştiren kötü amaçlı yazılımlardır. Bir dosya indirerek veya herhangi bir programı çalıştırarak yayılabilir.

b-) Solucanlar (Worms): E-posta ekleri yoluyla bilgisayarlara veya ağlara yayılır. Solucanlar, hizmet reddi saldırılarına neden olabilir.

c-) Truva atları (Trojan): En tehlikeli kötü amaçlı yazılımlardan biridir. Yararlı bir programda gizlenir ve virüsler gibi çoğalmaz.

d-) Fidye yazılımı (Ransomware): Kullanıcıyı kilitleyen bir tür kötü amaçlı yazılımdır. Saldırgan, istediği fidye ödenmediği takdirde kullanıcıyı tehdit eder ve eriştiği verileri erişime açmaz.

e-) Casus yazılım (Spyware): Kullanıcı etkinliğini kullanıcı onayı olmadan denetleyen ve saldırgana bildiren bir tür kötü amaçlı yazılımdır [65].

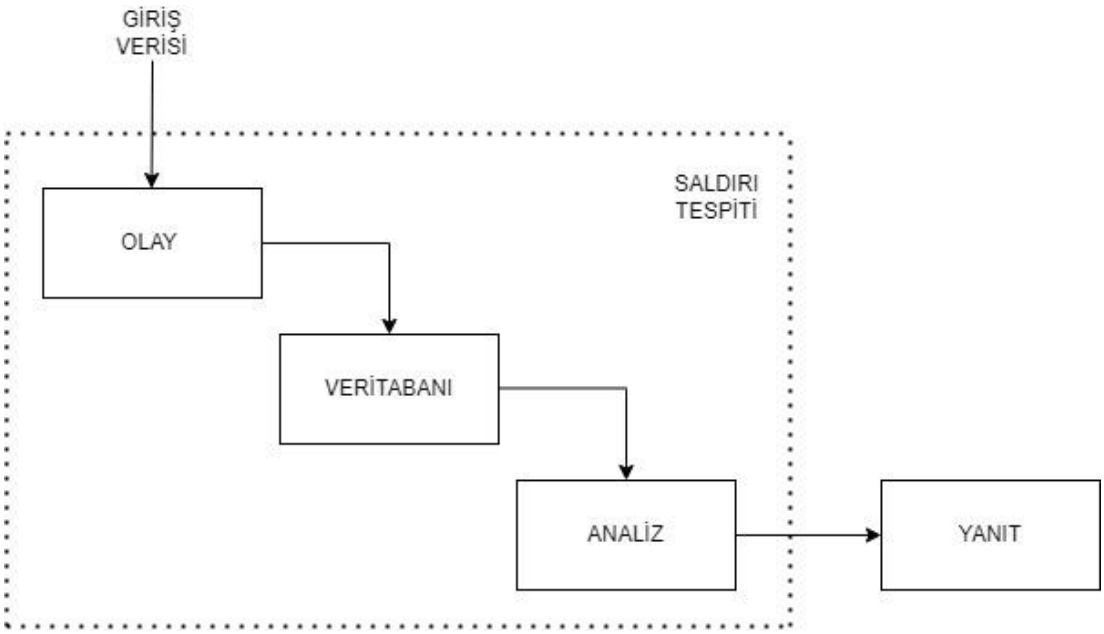
BÖLÜM 3

SALDIRI TESPİT SİSTEMLERİ

İlk otomatik IDS, 1980 yılında James P. Anderson tarafından ABD Hava Kuvvetlerine yönelik bilgisayar güvenlik tehditleri hakkındaki raporunda tanımlanmış ve kısa bir süre sonra ilk kural tabanlı model inşa edilmiştir. 1980'lerin sonunda birçok kuruluş, sürekli büyüyen ağ bağlantılı operasyonlarını korumak için bu sistemi benimsemeye başlamıştır. 1990'larda ağ tehditleri daha çeşitli geldi ve sonuç olarak IDS' de aynısını yaparak anormallik tespit yöntemini yarattı [89]. 1990'ların sonlarında ve 2000'lerin başlarında IDS'ler veri miktarının artmasına bağlı olarak yanlış pozitif değerler üretti ve bu sonuçlar IDS'ler tutarsız hale gelmeye başladı. Bununla birlikte, 2000'lerin ortalarından bu yana, taşınabilir bilgisayarların yaygınlaşması, bulut tabanlı sistemlerin ortaya çıkışı ve her iki yılda bir paylaşılan veri miktarının ikiye katlanması nedeniyle IDS yöntemleri daha sık kullanılmaya başlandı [90]. IDS, tüm ağ sistemine daha yüksek güvenlik sağlarmaktadır. Operasyonun başından sonuna kadar olayların değerlendiririr. Yetkisiz değişiklikler konusunda kullanıcıları uyarır ve bilinen tehditlerle karşılaştırma işlemi yaparken mevcut tehditlerin istatistiksel analizini yapar. Yönetici yardımını en aza indirgeyerek sistem yönetimini sağlar. Ayrıca ağdaki anormal olayların araştırılmasını sağlayarak, kötü amaçlı yazılımları engeller veya kötü amaçlı oluşumlar göründüğünde bildirimler oluşturur [91].

IDS'nin en önemli eksikliklerinden biri, sağlanan sonuçların kesin olmamasıdır. Doğrulukla ilgili iki ölçü kullanılır: İyi niyetli etkinlik yanlışlıkla şüpheli olarak işaretlendiğinde ortaya çıkan Yanlış Pozitif (FP) ve tersi gerçekleştiğinde, yani kötü niyetli etkinlik tespit edilmediğinde ortaya çıkan Yanlış Negatif (FN) [Ref]. Ho vd. anomali tabanlı IDS kullanarak iki binden fazla FP ve FN vakasını toplayıp analiz edip ve şu sonuçlara varmışlardır: ilk olarak, FN'lerden daha fazla FP vakası vardır. İkinci olarak, birçok FP'nin nedeni kural yapılandırması ve özel uygulama davranışının açıkça tanımlanmamış olmasıdır [92].

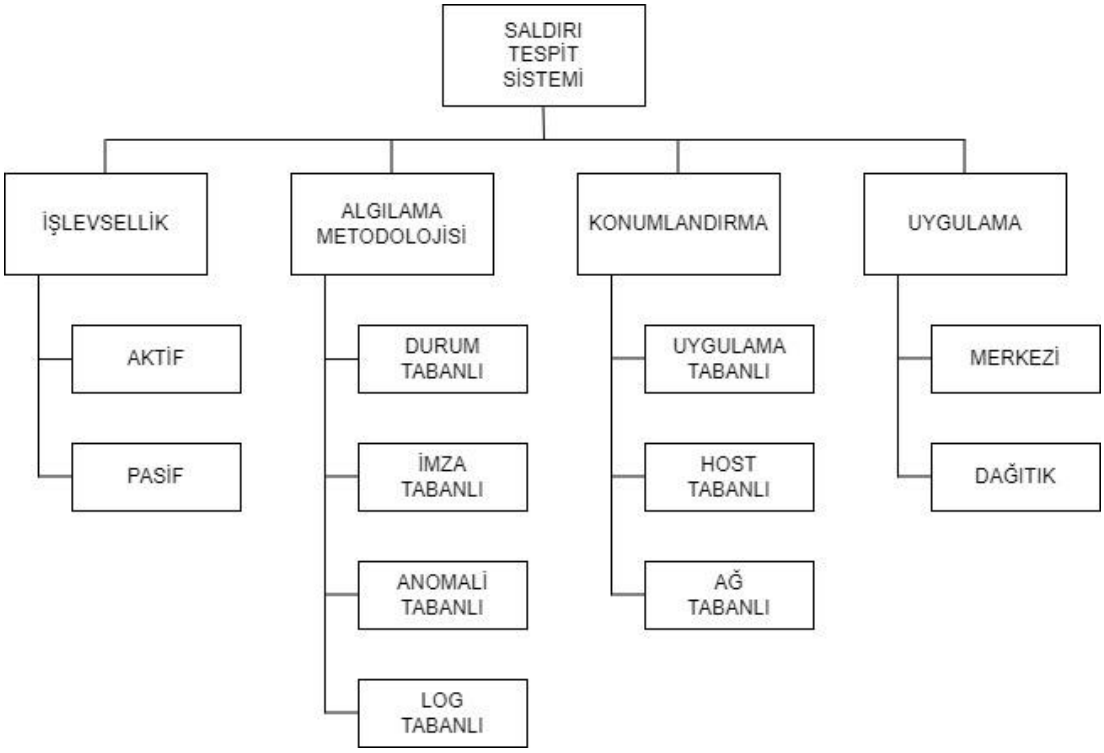
DARPA'nın İzinsiz Giriş Tespiti Çalışma Grubu (IDWG), genelleştirilmiş bir IDS Mimarisi tanımlamıştır. Bu mimari dört modülden oluşur: İlk modül olan olay kutuları, bilgi giriş noktalarıdır ve sistemin tüm sensörlerini içerir. İkinci modül olan veritabanı Kutuları, olayların saklandığı yerdir. Üçüncü modül olan Analiz Kutuları'nda olası saldırılar tespit edilir. Son modül olan Yanıt Kutuları ise, sistem tarafından gerçekleştirilen eylemlerdir [93].



Şekil 3.1. IDWG Genel IDS Mimarisi

3.1. SALDIRI TESPİT SİSTEMLERİNİN SINIFLANDIRILMASI

IDS'ler farklı şekilde sınıflandırılabilir. Bu tez çalışmasında Şekil 3.2'de gösterildiği ve aşağıda açıklandığı şekilde dört gruba ayrılmıştır.



Şekil 3.2. Saldırı tespit sistemlerinin sınıflandırılması.

3.1.1 İşlevselliğine göre Saldırı Tespit Sistemleri

Bu sınıflandırma, olası bir tehdit belirlendiğinde IDS tarafından gerçekleştirilen eyleme bakılmasını esas alır. Kendi içerisinde iki farklı gruba ayrılır.

a) Pasif IDS: Yöneticinin harekete geçmesini bekleyen, olası saldırıları tespit edip rapor edecek şekilde yapılandırılmış saldırı tespit sistemleridir. Yönetici tehdidi zamanında tanıyamazsa, bu yöntem güvenlik riskini azaltmak yerine artırır.

b) Aktif IDS: Bu IDS'ler yalnızca tehdidi bildirmekle kalmaz, aynı zamanda kötü niyetli etkinliği engellemek için önceden belirlenmiş eylemleri de gerçekleştirir [94].

3.1.2 Konumlandırılmasına göre Saldırı Tespit Sistemleri

Bu sınıflandırmada saldırı tespit sisteminin ağ ortamının neresinde uygulandığına atıfta bulunulur. Kendi içerisinde üç gruba ayrılır.

a-)Uygulama Tabanlı (AIDS): Bir ana bilgisayarda belirli bir uygulamayı korumak için özelleştirilmiş saldırı tespit sistemleridir. AIDS genellikle web sunucuları ve posta sunucuları gibi doğrudan internete maruz kalan uygulamalar için uygulanır.

b) Host Tabanlı (HIDS): Şifresi çözüldükten sonra şifrelenmiş trafiği analiz etmeye izin veren ve yalnızca bir bilgisayarda çalışan IDS türüdür. Dezavantajı, tüm ağa değil, yalnızca bir bilgisayara bakması ve büyük ağlar için yönetilmesinin imkansız hale gelmesidir.

c) Ağ Tabanlı (NIDS): Ağ tabanlı saldırı tespit sistemleri tüm ağın veya ağın bir bölümünün trafiğini izler, ancak şifreli trafiği göremez [95].

3.1.3 Uygulamaya göre Saldırı Tespit Sistemleri

IDS'nin network içerisinde uygulanma şekline göre sınıflandırılmasıdır.

a) Tek Ana Bilgisayar (Merkezi): Tüm sistemin bilgilerini toplayan ve işleyen tek bir merkezi ana bilgisayar vardır ve yalnızca algılamadan sorumludur.

b) Çoklu Ana Bilgisayar (Dağıtılmış): Çeşitli konumlarda (Uygulamalar, Ana Bilgisayarlar ve Ağ) kurulu birkaç bağımsız izleme sisteminden veri toplar ve birden fazla giriş noktasına sahip dağınık veya işbirlikçi saldırıları tespit etmenin yanı sıra tahsis edilen her bölümü koruma amacıyla veri toplar [96].

3.1.4 Algılama Metodolojisine göre Saldırı Tespit Sistemleri

Atak paketlerinin belirlenmesi için yapılan analiz türünün kullanıldığı sınıflandırma çeşididir.

a) Durum Tabanlı: Satıcı tarafından sağlanan protokol profillerini alır ve alınan komutları kötü amaçlı etkinlik arayarak analiz eder. Fakat, bu algılama işlemi doğru oluşturulmuş komutları yok saymaktadır.

b) İmza Tabanlı: Potansiyel olarak zararlı eylemleri izole ve tahmin etmek için bilinen zararlı davranış kalıplarını veya imzalarını kullanır. Kötüye kullanım tespiti olarak da bilinmektedir.

c) Anomali Tabanlı: Normal kullanım için bir referans noktası oluşturmaya ve ardından olası izinsiz girişleri ortaya çıkarmak için bu referans noktasından sapan kalıpları bulmaya çalışır.

d) Log Tabanlı: Bu metodoloji, saldırıları tespit etmek için kaynak olarak belirlenen bir ortamın uygulama ve cihaz kayıtlarını analiz eder [97].

3.2. SALDIRI TESPİT SİSTEMLERİNDE DERİN ÖĞRENME MİMARİLERİ

Derin öğrenme algoritmaları, veri miktarındaki artışa bağlı olarak son zamanlarda saldırı tespit sistemlerinde sıklıkla kullanılmaktadır. Derin öğrenme, Şekil 3.3'de gösterildiği gibi yapay zekanın bir alt dalı olan makine öğrenmesinin özel bir kullanımıdır. Derin sinir ağındaki birçok soyutlama katmanı, giriş verilerini çıkış katmanına iletir. Özellik seçme mühendisliğinin öğrenme süreci üzerinde hiçbir etkisi yoktur. Çeşitli kriterlere göre hata olasılığına dayalı olarak bir sınıflandırmanın doğru olup olmadığını belirlemek için çeşitli istatistiksel yöntemler kullanılabilir. Derin öğrenmenin odak noktası, çok katmanlı denetimsiz öğrenmenin hiyerarşik ağlarında sınıflandırma işlemini gerçekleştirmektir. Derin öğrenme ağları, kullanılan teknolojiye, derin öğrenme yöntemine ve izinsiz giriş tespit yöntemine göre değişiklik gösterebilir. Günümüzde evrişimli sinir ağları (CNN), tekrarlayan sinir ağları (RNN), uzun kısa süreli hafıza (LSTM), oto kodlayıcılar (AE) gibi derin öğrenme yöntemleri saldırı tespit sistemlerinde tercih edilmektedir. Bu tez çalışmasında CNN ve LSTM derin öğrenme algoritmaları kullanılmıştır [98].

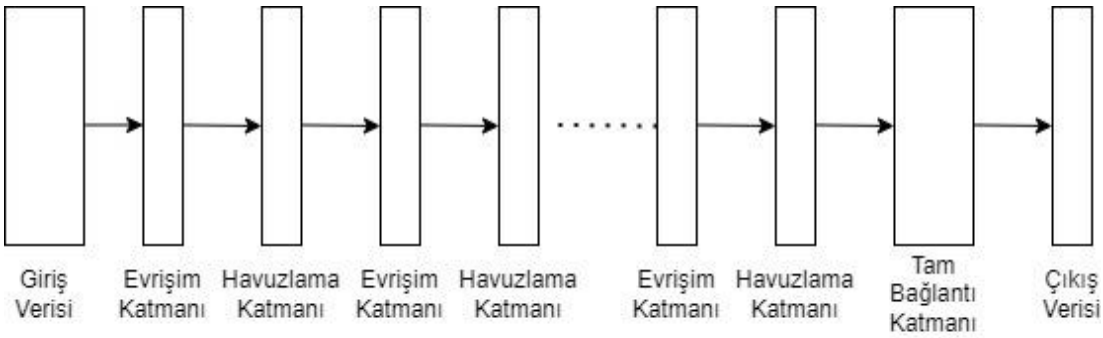
Şekil 3.3. Yapay Zeka İçeriği

3.2.1 Evrişimli Sinir Ağı (CNN)

1959'da, iki nörofizyolog David Hubel ve Torsten Wiesel yaptıkları çalışmaya katmanlı yapıyı kullanmaya başladılar. Bu katmanlar, önce yerel özellikleri ayıklamakta, ardından çıkarılan özellikleri daha yüksek düzeyde temsil için birleştirerek görsel kalıpları tanıma işlevini yerine getirmektedirler. Daha sonra bu kavram, Derin Öğrenmenin temel ilkelerinden biri haline gelmiştir [99]. 1980 yılında Hubel ve Wiesel'in çalışmasından ilham alan Kunihiko Fukushima, birden çok katman içeren kendi kendini organize eden bir Sinir Ağı olan Neocognitron'u önermiştir. Önerilen bu mimari, görsel kalıpları öğrenme yoluyla hiyerarşik olarak tanıyabilen CNN'nin, ilk teorik modeli olmuştur. Neocognitron mimarisi üzerinde LeCun ve arkadaşları tarafından 1989'da, MNIST el yazısı rakam veri setini başarıyla tanıyan LeNet-5 adlı modern bir CNN çerçevesi geliştirerek büyük bir gelişme yaşandı. LeNet-5, hata geri yayılım algoritması kullanılarak eğitildi ve görsel modelleri, herhangi bir ayrılmış özellik mühendisliği mekanizması kullanmadan doğrudan giriş görüntülerinden tanıyabilmektedir. LeNet-5'i keşfettikten sonra, büyük eğitim verilerinin olmaması gibi çeşitli sınırlamalar nedeniyle, algoritmadaki yenilikler ve yetersiz işlem gücü nedeniyle, CNN çeşitli karmaşık problemlerde iyi performans gösterememiştir. Ancak günümüzde, Büyük Veri çağında, büyük etiketli veri kümeleri, daha yenilikçi algoritmalar ve özellikle güçlü GPU makineleri bulunmaktadır. Bu tür yükseltmelerle, 2012'de Krizhevsky ve arkadaşları AlexNet'i

tasarladı. AlexNet, CNN modelinin bilgisayarlı görü ve doğal dil işleme alanlarında uygulanmasının yolunu açmıştır. Bununla birlikte SCADA, IoT, IIoT gibi farklı sistem ve ağlarda CNN mimarileri saldırı tespiti için kullanılmaktadır.

Evrişimli Sinir Ağı (CNN), derin ileri beslemeli mimariye sahip ve diğer ağlara kıyasla şaşırtıcı genelleme yeteneğine sahip bir Yapay Sinir Ağı (YSA) türüdür. Soyutlanmış özellikleri öğrenme oranı yüksektir. Derin bir CNN modeli, birden çok soyutlama düzeyiyle giriş verilerinin çeşitli özelliklerini öğrenebilen sonlu bir işleme katmanları kümesinden oluşur. Başlangıç katmanları, yüksek seviyeli özellikleri (düşük soyutlama ile) öğrenir ve çıkarır. Daha derin katmanlar, düşük seviyeli özellikleri (yüksek soyutlama ile) öğrenir ve çıkarır [100]. CNN'nin temel kavramsal modeli Şekil 3.4'de gösterilmiştir.



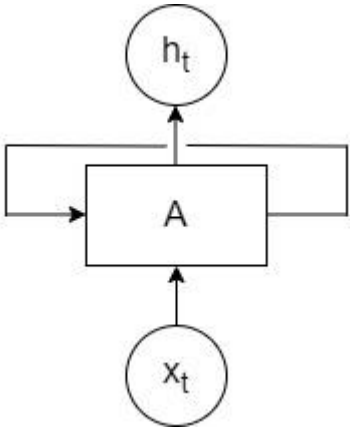
Şekil 3.4. CNN algoritması temel kavramsal modeli.

3.2.2 Tekrarlayan Sinir Ağları (RNN)

RNN'ler güçlü ve sağlam bir sinir ağı türüdür. Dahili belleğe sahip tek algoritma olduğu için uygulamada umut verici sonuçlar üreten algoritmaları bulunmaktadır.

Diğer birçok derin öğrenme algoritması gibi, tekrarlayan sinir ağları da nispeten eskidir. Başlangıçta 1980'lerde kullanılmaya başlansa da, gerçek potansiyelleri son yıllarda ortaya çıkmıştır. Artık çalışmak zorunda olduğumuz devasa miktarda veriyle birlikte hesaplama gücündeki artış ve 1990'larda uzun kısa süreli belleğin (LSTM) icadı, RNN'leri ön plana çıkarmıştır.

RNN'ler, dahili bellekleri sayesinde aldıkları girdilerle ilgili önemli özellikleri hatırlayabilirler, bu da onların bir sonraki adımı tahmin etmede çok kesin değerler üretmelerini sağlar. Bu nedenle zaman serileri, konuşma, metin, finansal veriler, ses, video, hava durumu ve çok daha fazlası gibi sıralı veriler için tercih edilen algoritmalardır. Diğer derin öğrenme yöntemleri giriş verisi olarak sabit boyutlu bir vektörü kabul eder. Çıkış verisi de yine sabit boyutlu bir vektördür. RNN'ler kendi giriş hafızalarını ileri beslemeli sinir ağlarının aksine, giriş verilerinin rastgele dizilişlerini işlemek için kullanabilirler. Tekrarlayan sinir ağları, diğer algoritmalara kıyasla bir dizi ve bağlamı hakkında çok daha derin bir anlayış oluşturabilir. Genel bir RNN mimarisi Şekil 3.5'te gösterilmiştir. Tekrarlayan sinir ağlarında h çıktıyı, x ise girdiyi ifade eder. Her katmanda çıkan sonuç bir sonraki katman için girdi olarak kullanılır [101].



Şekil 3.5. Genel RNN mimarisi.

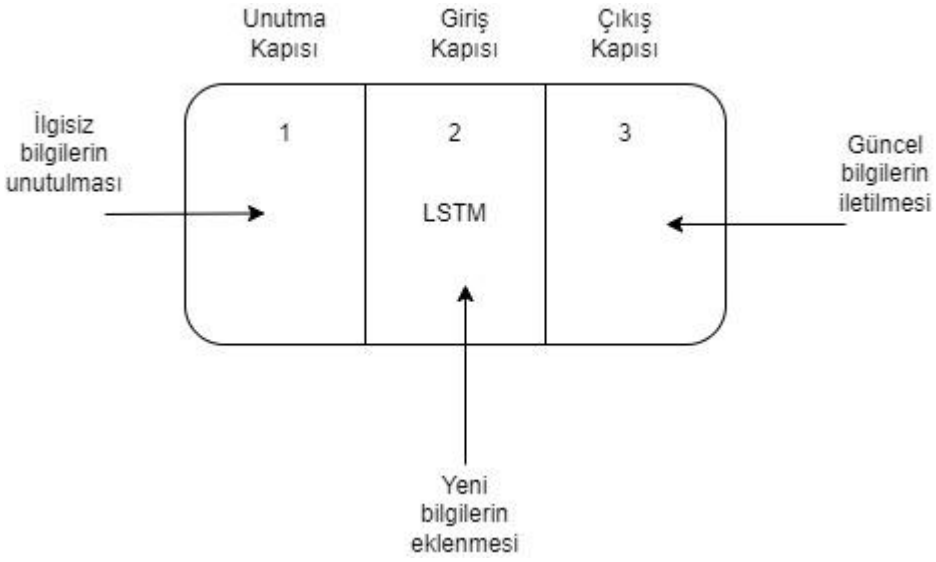
3.2.3 Uzun Kısa Süreli Hafıza (LSTM)

Uzun Kısa Süreli Bellek Ağları, bilginin kalıcı olmasını sağlayan derin öğrenme tabanlı, sıralı bir sinir ağıdır. RNN'nin karşılaştığı yok olan gradyan problemini çözebilen, özel bir Tekrarlayan Sinir Ağı türüdür. Yüksek düzeyde, LSTM bir RNN hücresi gibi çalışır. LSTM ağ mimarisi, Şekil 3.6'da gösterildiği gibi üç bölümden oluşur ve her parça ayrı bir işlevi yerine getirir.

İlk kısım, bir önceki zaman damgasından gelen bilginin hatırlanıp hatırlanmayacağına veya ilgisiz olup unutulmayacağına karar verir. İkinci kısımda ise hücre, bu hücreye

gelen girdiden yeni bilgiler öğrenmeye çalışır. Son olarak, üçüncü bölümde, hücre güncellenen bilgileri mevcut zaman damgasından bir sonraki zaman damgasına aktarır. LSTM'nin bu bir döngüsü, tek seferlik bir adım olarak kabul edilir.

Bir LSTM biriminin bu üç parçası, kapılar olarak bilinir. Bellek hücresine veya LSTM hücresine giren ve çıkan bilgi akışını kontrol ederler. Birinci kapıya Unutma kapısı, ikinci kapıya Giriş kapısı ve sonuncusuna da Çıkış kapısı denir. Bu üç kapıdan ve bir bellek hücresinden veya LSTM hücresinden oluşan bir LSTM birimi, geleneksel ileri beslemeli sinir ağlarında, her bir nöronun bir gizli katmanı ve bir geçerli durumu olduğu bir nöron katmanı olarak düşünülebilir [102].

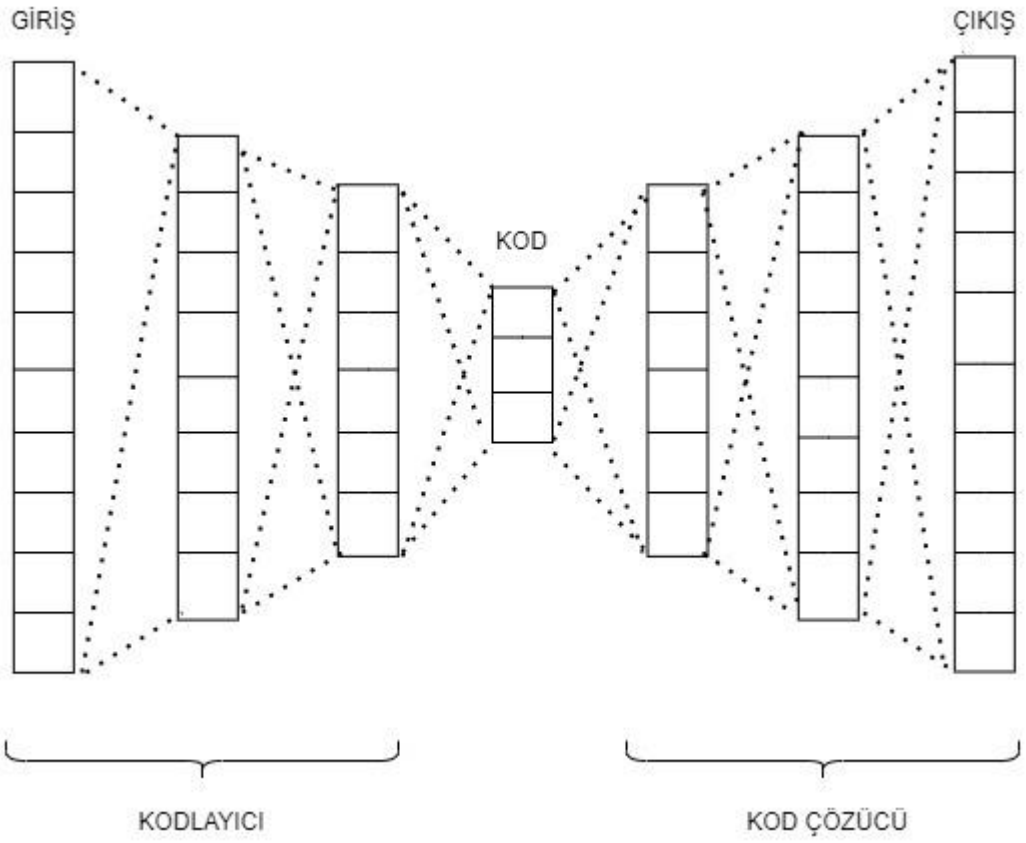


Şekil 3.6. Genel LSTM mimarisi.

3.2.4 Derin Oto Kodlayıcılar (AE)

Yapay sinir ağının özel bir uygulaması olan derin oto kodlayıcılar, denetimsiz öğrenme gerçekleştirmek için kullanılır. Otomatik kodlayıcı, verilerin verimli bir şekilde nasıl sıkıştırılacağını ve kodlanacağını öğrenen, ardından verilerin azaltılmış kodlanmış temsilden orijinal girdiye mümkün olduğunca yakın bir temsile nasıl yeniden yapılandırılacağını öğrenen denetimsiz bir yapay sinir ağıdır.

Otomatik kodlayıcı, tasarımı gereği, verilerdeki gürültünün nasıl göz ardı edileceğini öğrenerek veri boyutlarını azaltır. Bir otomatik kodlayıcı 3 bileşenden oluşur. Bunlar; kodlayıcı, kod ve kod çözücüdür. Kodlayıcı girişi sıkıştırır ve kodu üretir, kod çözücü daha sonra girişi yalnızca bu kodu kullanarak yeniden oluşturur. Diğer bir ifade ile bir otomatik kodlayıcı oluşturmak için 3 yapıya ihtiyaç duyulur. Bunlar, bir kodlama yöntemi, kod çözme yöntemi ve çıktıyı hedefle karşılaştırmak için bir kayıp fonksiyonudur [103]. Şekil 3.7'de genel bir derin otokodlayıcı mimarisi gösterilmektedir.



Şekil 3.7. Genel AE mimarisi.

BÖLÜM 4

DENEYSEL ÇALIŞMADA KULLANILAN VERİ SETLERİ

Bu çalışmada önerilen modellerin performansları güncel bir veri seti olan X-IIoTID ve literatürde sıklıkla tercih edilen bir veri seti olan UNSW-NB15 kullanılarak belirlenmiştir.

4.1. UNSW-NB15 VERİ SETİ

UNSW-NB15 [104], IDS çalışmaları tarafından yaygın olarak kullanılan gelişmiş bir veri setidir. UNSW NB15 veri setinde Çizelge 4.1'de listelenen kırk iki özellik vardır. Çizelge 4.1'deki özellikler listesinde sunulduğu gibi, otuz dokuz özellik sayısaldır ve üç özellik kategoriktir.

UNSW-NB15 izinsiz giriş tespit veri seti, Reconnaissance, DoS, Analysis, Fuzzers, Exploits, Shellcode, Generic, Worms, Backdoor ve Benign [105] saldırılarını içerir. UNSW-NB15 veri setinde yer alan saldırı tiplerinin değer dağılımı Çizelge 4.2'de sunulmaktadır. Çizelge 4.2'de görüldüğü gibi veri setindeki saldırı tipi sayıları dengeli bir dağılım göstermemektedir. Daha önceki çalışmalarda bu şekilde oluşturulan veri setlerinde sayısı az olan saldırı türlerinde ağın ezberleme sorunu yaşadığı bilinmektedir [106]. Bu sorunun önüne geçmek için literatürde sentetik veri üretme yöntemi kullanılmaktadır. Bu yöntem sayesinde veri seti içerisinde sayısı az olan saldırı türleri birbiriyle dengelenmekte ve bu sayede hem dengeli bir veri seti elde edilmekte hem de toplam veri miktarı artırılmaktadır [107].

Çizelge 4.1. UNSW-NB15 veri setindeki özellikler

Özellik	**Değer**	**Bölüm**	**Özellik**	**Değer**	**Bölüm**
service	nominal	primary	stcpb	int	content
state	nominal	primary	dwin	int	content
sinpkt	float	time	djit	float	time
dpkts	int	primary	dmean	int	content
dbytes	int	primary	trans_depth	int	content
synack	float	time	ct_state_ttl	int	general
dttl	int	primary	İs_ftp_login	binary	general
dload	float	primary	sttl	int	primary
ct_dst_src_ltm	int	connection	ct_ftp_cmd	int	general
sload	float	primary	is_ftp_login	int	general
ct_src_dport_ltm	int	connection	ct_flw_http_mthd	int	content
dloss	int	primary	dtcpb	int	content
sloss	int	primary	is_sm_ips_ports	binary	general
dur	float	primary	ct_srv_src	int	connection
proto	nominal	flow	smean	int	content
tcprtt	float	time	ct_dst_ltm	int	connection
synacks	float	time	ct_dst_sport_ltm	int	connection
dinpkt	float	time	spkts	int	primary
sjit	float	time	ct_drc_ltm	int	connection
sbytes	int	primary	ackdat	float	time
swin	int	content	ct_srv_dst	int	connection
response_body_len	int	content			

Çizelge 4.2. UNSW-NB15 veri setindeki atak çeşitleri ve sayıları.

Atak Çeşitleri	**Atak Sayıları**
Fuzzers	18184
Backdoor	1746
Analysis	2000
General	40000
Shellcode	1133
Reconnaissance	10491
DoS	12264
Worms	130
Exploits	33393
Benign	56000

4.2. X-IIoTID VERİ SETİ

Çalışmada kullanılan bir diğer veri seti ise X-IIoTID'dir. Güncel bir veri seti olarak X-IIoTID, birbiriyle uyumlu çalışan ve karmaşık yapıya sahip IIoT cihazları için oluşturulmuştur. Bütüncül bir yaklaşımla oluşturulan bu veri seti, IIoT ağındaki cihaz ve protokollerden elde edilen trafik bilgileri ve değişikliklerden oluşmaktadır. Veri kümesi ayrıca çeşitli aygıtlardan ve bağlantılardan günlükler, kaynaklar ve uyarı özellikleri içerir. X-IIoTID veri setinde 820.834 saldırı türü ve 68 özellik bulunmaktadır. X-IIoTID veri kümesindeki saldırı türlerinin sayıları Çizelge 4.3'te [66] sunulmuştur.

Çizelge 4.3. X-IIoTID veri setindeki atak çeşitleri ve sayıları.

Atak Çeşitleri	**Atak Sayıları**
Generic Scanning	50277
Scanning Vulnerability	52852
Discovering Resources	23148
Fuzzing	1313
Brute Force	47241
Dictionary	2572
Insider Malicious	17447
Reverse Shell	1016
Mitm	117
Modbus Register	5953
MQTT	23524
TCP Relay	2119
Command Control	2863
Exfiltration	22164
Ransomware	458
RDoS	141261
Fake Notification	28
False Data Injection	5094
Normal	4211417

Bu çalışmada veri setinin %70'i modellerin eğitim aşamasında, %15'i doğrulama aşamasında ve %15'i test aşamasında kullanılmıştır. Eğitim sürecinde en iyi sonuçların alınıp alınmadığını kontrol etmek için doğrulama süreci yapılmıştır.

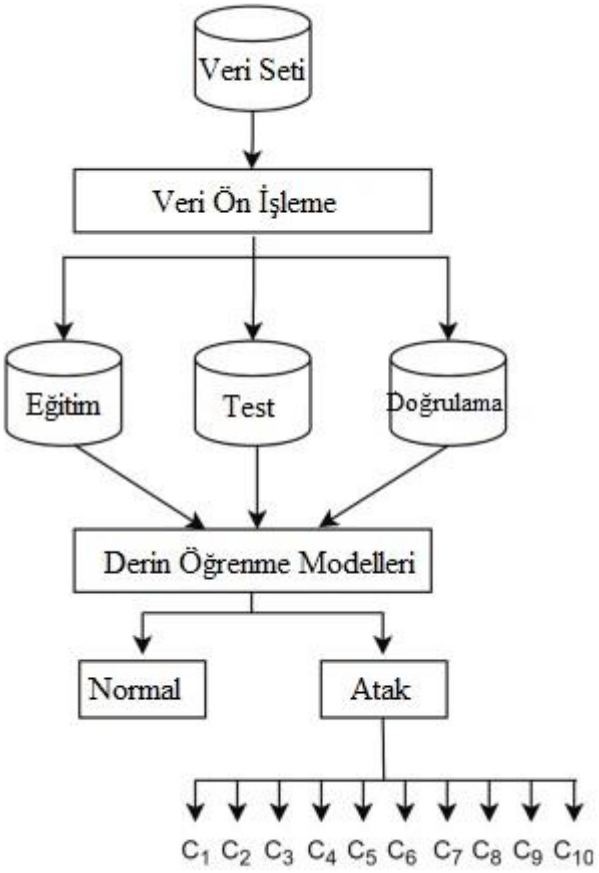
BÖLÜM 5

ÖNERİLEN MODEL

CNN'ler ve LSTM'ler bu çalışmada kullanılan derin öğrenme modelleridir. Bu modellerin çıktılarını hesaplamak için her bir derin öğrenme yönteminde doğruluk ve kayıp metrikleri kullanılmıştır. Ayrıca performans değerlendirme ölçütleri olarak doğruluk, hatırlama, F1 puanı ve kesinlik kullanılmıştır. RNN'ler ve LSTM ağlarındaki geçici bileşen özelliği, bu yöntemleri diğer sinir ağlarından ayırır. Bu iki model, zamansal bileşen aracılığıyla zaman tüketim özelliğini ve sıralama değerlendirmesini içerir. RNN ve LSTM'yi diğer sinir ağlarından ayıran temel fark, zaman tüketmeleri ve diziyi zamansal bileşenlerine göre dikkate almalarıdır. Araştırmalar, saldırı tespit sistemlerinin tasarımında sinir ağının en verimli yöntem olduğunu göstermektedir. Yine de son zamanlarda, bellek ağları, dönüştürücüler ve dil görevleri, performans oranları açısından sinir ağlarına göre öncelikli hale gelmiştir. RNN'ler yamalara bölünebilir ve aynı zamanda parçalar halinde yaklaşılabilen görüntülere genişletilebilir. Belirli bir hafıza biçimini sağlayan tekrarlayan ağlar ve hafızanın insan deneyiminin bir parçası haline gelmesi, öğrenme sürecindeki etkileyici unsurlardır [108].

İkili ve çok sınıflı sınıflandırma süreçleri için önerilen derin öğrenme modellerinin mimarisi Şekil 5.1'de sunulmuştur. İlk adımda, veri ön işleme katmanında UNSW-NB15 ve X-IIoTID veri kümelerine veri temizleme uygulandı. Daha sonra veri setleri min-maks yöntemi kullanılarak normalleştirildi. Daha sonra veri kümeleri eğitim, doğrulama ve test kümelerine ayrılmıştır. Derin öğrenme sınıflandırıcı adımında, normalize edilmiş veri seti üzerinden CNN ve LSTM sınıflandırıcıları üretilmiştir. Derin öğrenme modelleri, ikili ve çok sınıflı sınıflandırıcılar olarak eğitildi. Değerlendirme aşamasında, son adım olarak, derin öğrenme yöntemlerine dayalı olarak önerilen IDS, oluşturulan her modele ait hem ikili hem de çok sınıflı

sınıflandırmalar için doğruluk, hatırlama, kesinlik, kayıp ve F1-Skor kullanılarak değerlendirilmiştir.



Şekil 5.1. İkili ve çok sınıflı sınıflandırma model mimarisi.

Evrişimli sinir ağları, saldırı tespitinde yaygın olarak kullanılan çok katmanlı bir yapay sinir ağı yöntemidir. Evrişimli sinir ağları, evrişim adı verilen matematiksel bir işlem yürütür. Evrişim, özel bir doğrusal işlem türüdür. Evrişimli ağlar, genel matris çarpımı yerine katmanlardan en az birinde evrişimi kullanan sinir ağlarıdır [109]. CNN'ler genellikle bir evrişim katmanı, bir havuzlama katmanı ve tamamen bağlı bir katmandan oluşur. Evrişimli sinir ağları, karmaşık öznitelikleri otomatik olarak çıkarmayı öğrenebilir. Evrişim katmanında [110] gelişmiş bir öznitelik gösterimi elde edilir. Konvolüsyon işlemi, Denklem 5.1'de gösterildiği gibi sunulur. Burada gösterilen x_i^a, evrişim katmanı a'nın i. özniteliği haritasıdır. $\emptyset$, etkinleştirme işlevini temsil eder. ki, katmanın (a-1) giriş özniteliği kümesidir. w_{ji}^a, evrişim katmanı a'nın i.

niteliği ile (a-1) katmanının j. niteliği arasındaki bağlantı ağırlığıdır. b_j^a, ilgili katmandaki sapmadır.

$$x_i^a = \emptyset \left[\sum_{i \in ki} x_j^{a-1} * w_{ji}^a + b_j^a \right] \tag{5.1}$$

Havuzlama katmanı, evrişim katmanını takip eder. Havuzlama katmanının amacı, öznitelik haritasının boyutunu küçültmektir. Bu işlem, önemli özniteliklerin tanımlanmasını sağlar, veri karmaşıklığını azaltır ve ağın çevresel değişikliklere karşı toleransını artırır. Havuzlama katmanı, Denklem 5.2'de gösterildiği gibi sunulabilir.
Burada gösterilen c, alt örnekleme işlevini, β ise ağırlıklandırma matrisini gösterir. Sınıflandırma işlemi, evrişim katmanı ve havuzlama katmanını takip eden tam bağlantılı katman aracılığıyla gerçekleştirilir. Şekil 5.2, tamamen bağlı katmanın çıktı işlevini göstermektedir.

$$x_i^a = \emptyset \left[\beta_i^a C(x_i^{a-1} + b_i^a) \right] \tag{5.2}$$

Denklem 5.3'te gösterilen m, katman indeksini gösterirken, y^m tamamen bağlı katmanın çıktısını, x^{m-1} tam bağlı katmanın girişini, w^m ağırlık katsayısını ve b^m sapmasını gösterir [110].

$$y^m = \emptyset \left[w^m x^{m-1} + b^m \right] \tag{5.3}$$

LSTM hücreleri, tekrarlayan ağların gizli birimlerinin yerini alır ve tekrarlayan bağlantılara sahiptir [109]. LSTM bloğunda, x^t, t zaman adımındaki giriş vektörünü, h_{t-1}, zaman adımındaki (t-1) gizli durumu ve c_{t-1}, zaman adımındaki (t-1) bellek hücresi durumunu gösterir. Bunlar bloğun girişlerini oluşturur. LSTM giriş, unutma ve çıkış kapılarını içerir. LSTM'nin hücre durumu, unutma, giriş ve çıkış kapılarına ilişkin hesaplamalar aşağıdaki denklemlerde belirtilmiştir. Buna göre, t zaman adımındaki i hücresi için unutma kapısı $f_i^{(t)}$, Denklem 5.4'de gösterilmiştir. b^f, Z^f ve D^f sırasıyla

unutma kapıları için sapmayı, girdi ağırlığını ve tekrarlayan ağırlıkları gösterir. Denklem 5.5'te $n_i^{(t)}$ i hücresinin durumundaki güncellemeyi gösterirken b, Z ve D sırasıyla sapmayı, giriş ağırlığını ve LSTM hücresine giren tekrarlayan ağırlıkları gösterir. $p_i^{(t)}$, i hücresi için giriş kapısının hesaplanması anlamına gelir ve Denklem 5.6'da gösterilir ve bu hesaplama unutma kapısına benzer şekilde yapılır. Denklem 5.7'deki $h_i^{(t)}$ i. gizli durumu, $s_i^{(t)}$ ise i. çıkış kapısını gösterir. Çıkış kapısı denklemi Denklem 5.8'de gösterilmiştir. b^0, Z^0 ve D^0 sırasıyla sapmayı, giriş ağırlığını ve tekrarlayan ağırlıkları gösterir.

$$f_i^t = \sigma\left(b_i^f + \sum j\, Z_{ij}^f\, x_j^{(t)} + \sum j\, D_{ij}^f\, h_j^{(t-1)} \right) \tag{5.4}$$

$$n_i^t = f_i^{(t)} n_{ji}^{(t-1)} + p_i^{(t)} \sigma\left(b_i + \sum j Z_{ij}\, x_j^{(t)} + \sum j D_{ij}\, h_j^{(t-1)} \right. \tag{5.5}$$

$$p_i^{(t)} = \sigma\left(b_i^p + \sum j Z_{ij}^p\, x_j^t + \sum j D_{ij}^p\, h_j^{(t-1)} \right. \tag{5.6}$$

$$h_i^t = \tan h\,(n_i^{(t)}) s_i^{(t)} \tag{5.7}$$

$$s_i^{(t)} = \sigma\left(b_i^0 + \sum j Z_{ij}^0\, x_j^t + \sum j D_{ij}^0\, h_j^{(t-1)} \right. \tag{5.8}$$

Önerilen modelin her aşamasında gerçekleştirilen işlemler aşağıda ayrıntılı olarak açıklanmıştır:

Aşama 1: Birinci aşama olan ön işleme kendi içinde ikiye ayrılmıştır. Bu, UNSW-NB15 veri kümesinin bazı özellik sütunlarındaki çok sayıda eksik değerden kaynaklanmıştır. Bu eksik değerler, öğrenme modellerine dahil edilemez. Bu aşamada, özellik sütunundaki tüm boş hücreler, eksik değeri temsil eden "0" değeri ile doldurulduu. Ardından, hafızada saklanan hücreler metinlere dönüştürüldü; her bir kategorik değer belirli bir sayısal değer olarak temsil edilmiş ve dönüşüm işlemi gerçekleştirilmiştir. Verilerin normalleştirme süreci, çeşitli seviyelerde temsil edilen

sistemler için özellikle yardımcı oldu. Yapay sinir ağlarının daha tutarlı bir şekilde üretilmesine yardımcı olmak için min-maks normalizasyon işlemi gerçekleştirilmiştir. Bu yöntem, tüm veri bağlantılarını doğru bir şekilde yürütme avantajına sahiptir. Artan fonksiyon, sınıflandırma işlemine min-maks eklendiğinden gerçek değer aralığının altına düşer. Bununla birlikte, özellik değerleri mevcut aralıkta kalabilir [111].

$$X_{new} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{5.9}$$

Aşama 2: Veri setleri bu aşamada eğitim, doğrulama ve test olmak üzere üç gruba ayrıldı. Veri setlerinin %70'ini eğitim, %15'ini doğrulama ve %15'ini test seti olarak ayırdık.

Aşama 3: CNN, LSTM ve CNN + LSTM derin öğrenme tekniklerine dayalı üç tür derin öğrenme modeli aracılığıyla bir saldırı tespit sistemi önerdik. CNN yöntemi için uygulanan hiperparametreler Çizelge 5.1'de ve LSTM yöntemi için uygulananlar Çizelge 5.2'de sunulmaktadır. Önerilen CNN, LSTM ve CNN + LSTM modelleri Şekil 5.2, 5.3 ve 5.4'de gösterilmektedir. CNN + LSTM modelinin sözde kodu Çizelge 5.3'de gösterilmektedir.

Çizelge 5.1. CNN modelinde uygulanan hiperparametreler.

Hiperparametreler	**Değer**
Optimizer	Gradient Descent
Dropout	0.4
Layers	2
Feature Layer	Fully Connected
Classify Layer	Fully Connected
Learning Rate	0.001, 0.0001
Epoch	100

Çizelge 5.2. LSTM modelinde uygulanan hiperparametreler.

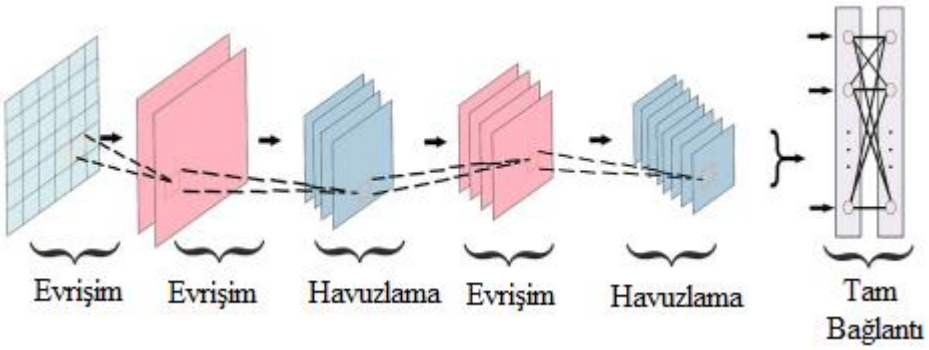
Hiperparametreler	**Değer**
Activation Function in Hidden Layer	ReLu
Epoch	100
Batch Size	120
Optimizer	Gradient Descent

Layers	3
Number of Neurons	256, 128, 64
Activation function in output layer	Softmax

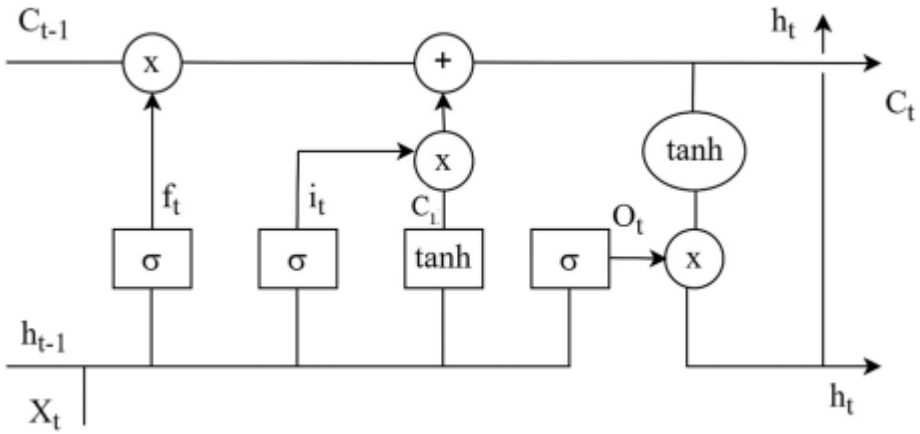
Çizelge 5.3. CNN+LSTM hibrit model sözde kodu.

Sözde Kod
*1: **Input:** Train_X, Train_Y*
2: Hyper-Parameters: optimizer, rate ,feature_layers, poolsize, batchsize
3: Initialize( )
4: Normalization(Train_X , Train_Y)
5:Convolution_1=Sequential((Convolution2D(optimizer,dropout,name=”Conv2D _1), MaxPooling2D(poolsize), dropout(rate))
6: Convolution_1.compile(Train_X, Train_Y, epochs, batchsize)
7: Convolution_1.fit(Train_X, Train_Y, epochs, batchsize)
8:Convolution_1_feature=Model(inputs,convolution_1(“Convolution2D”).output)
9: Convolution_1_feature.predict(Train_X)
10:Lstmmodel=Sequential(Lstm(units, activation, recurrent_activation), flatten (units,activation)
11: Lstmmodel.compile(lossfunction, optimizer)
12: Lstmmodel.fit(Convolution_1_feature, Train_Y, batchsize, epochs)

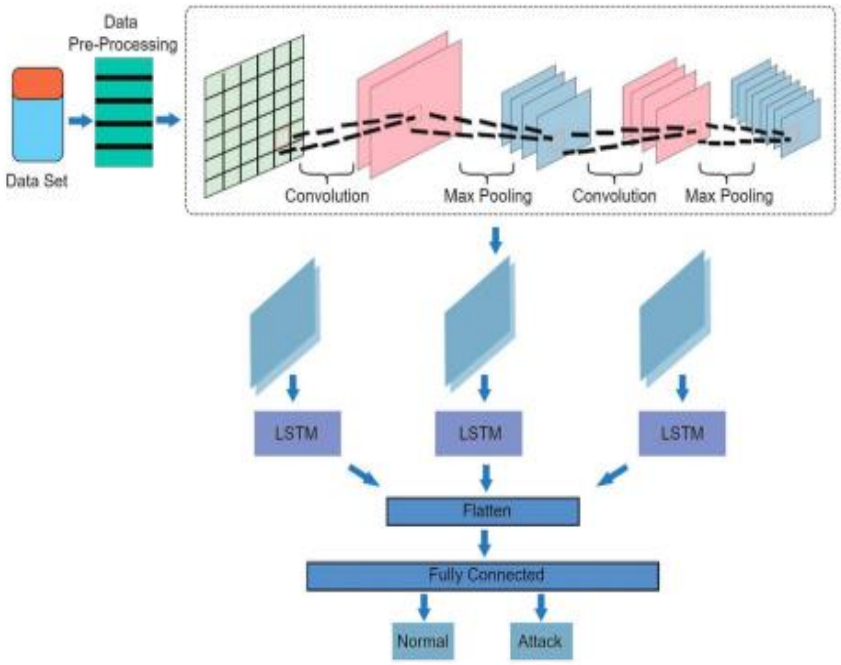
Aşama 4: Bu aşamada önerilen derin öğrenme yaklaşımları saldırı tespit yöntemleri olarak değerlendirilmiş ve belirlenen metriklere göre performansları ölçülmüştür. Her modelin performansını değerlendirmek için belirlenen metrikler, doğruluk, kesinlik, geri çağırma, F1 Puanı ve kayıp değeridir. Ayrıca veri kümeleri içerisindeki saldırı türlerinin tespitinde doğruluk oranı dikkate alınmıştır. Eğitim sürecinde dönem başına doğrulama kaybını izlemek için bir geri çağırma görevi vardı. Doğrulama kaybının otuz epoch boyunca iyileştirilememesi durumunda eğitim süreci durdurulur. Buna göre, eğitim süreci sona erdiğinde değerlendirme süreci de durdurulacaktır.



Şekil 5.2. Tam bağlantı katmanının çıktı işlevi.



Şekil 5.3. LSTM mimarisi.



Şekil 5.4. Önerilen CNN+LSTM mimarisi.

5.1. MODELİN DEĞERLENDİRİLMESİNDE KULLANILAN ÖLÇÜTLER

Önerilen yaklaşımın performansını değerlendirmek için doğruluk, F1 Puanı, hatırlama ve kesinlik metrikleri kullanılmıştır. Denklemlerde geçen TP, saldırı sayısını, TN ise doğru şekilde sınıflandırılan normal trafik sayısını gösterir. FP, aslında normal veri olan ancak yanlış sınıflandırılan ve saldırı olarak kabul edilen trafik sayısını gösterir. FN, normal trafik olarak yanlış sınıflandırılan saldırıların sayısını gösterir [112].

$$Doğruluk = \frac{TP + TN}{TP + TN + FP + FN} \tag{5.10}$$

$$Geri\ Çağırma = \frac{TP}{FN + TP} \tag{5.11}$$

$$Kesinlik = \frac{TP}{TP + FP} \tag{5.12}$$

$$F1 - Score = \frac{2 * Kesinlik * Geri\ Çağırma}{Kesinlik + Geri\ Çağırma} \tag{5.13}$$

5.2. KULLANILAN VERİ SETLERİNDE ELDE EDİLEN SONUÇLAR

Bu bölümde CNN, LSTM ve CNN + LSTM yöntemleri kullanılarak oluşturulan IDS ile elde edilen ikili ve çok sınıflı sınıflandırmaların sonuçları açıklanmıştır. Ayrıca bu çalışmada UNSW-NB15 ve X-IIoTID veri setlerinde elde edilen doğruluk oranları önceki çalışmalarda elde edilenlerle karşılaştırılmıştır. Son olarak, bu çalışmada kullanılan veri kümelerindeki saldırı türlerinin doğru tespit edilme oranları sunulmuştur.

UNSW-NB15 veri setinde LSTM kullanılarak gerçekleştirilen hem ikili hem de çok sınıflı sınıflandırma işlemleri için elde edilen sonuçlar Çizelge 5.4'de sunulmaktadır. İkili ve çok sınıflı etiketlemede performans metriklerinden optimum sonuçlar almak için tekrarlayan sinir ağlarının doğru konumlandırılması gerekir. Bu işlem ağın eğitim süresini kısalttığı gibi kayıp değerini de düşürür. Bu bağlamda, uzun kısa süreli bellek yönteminin ikili sınıflandırmadaki performansının, çok sınıflı sınıflandırmadaki performansına göre üstün olduğu belirlenmiştir. Test setinin doğruluk oranı ikili sınıflandırmada %91.14, çok sınıflı sınıflandırmada ise %91.10 bulunmuştur. Hem ikili hem de çok sınıflı sınıflandırmalarda test kaybı değeri %6.41'dir. LSTM modelinde eğitim için doğrulama doğruluk değeri ikili sınıflandırmada %91.05 ve çok sınıflı sınıflandırmada %91.08 olarak belirlenirken bu oran ikili sınıflandırmada %91.12 ve çok sınıflı sınıflandırmada %91.10 olarak belirlenmiştir. Eğitim kaybı değeri her iki sınıflandırma işlemi için %11,83 olurken, doğrulama için bu oran ikili sınıflandırmada %11,72 ve çok sınıflı sınıflandırmada %11,70 olmuştur. Eğitim sürecinde 100 epoch sonucu alınmıştır.

Çizelge 5.4. UNSW-NB15 veri setinde LSTM modeli ile elde edilen sonuçlar.

Ölçütler	**İkili Sınıflandırma**	**Çok Sınıflı Sınıflandırma**
Eğitim Doğruluğu	91.12%	91.10%
Değerlendirme Doğruluğu	91.05%	91.08%
Test Doğruluğu	91.14%	91.10%
Eğitimde Kayıp	11.83%	11.83%
Değerlendirmede Kayıp	11.72%	11.70%
Testte Kayıp	6.41%	6.41%

UNSW-NB15 veri setinde evrişimli sinir ağı (CNN) kullanılarak gerçekleştirilen hem ikili hem de çok sınıflı sınıflandırma işlemleri için elde edilen sonuçlar Çizelge 5.5'de sunulmuştur. İkili ve çok sınıflı etiketleme için oluşturulan veri kümesindeki performans ölçümlerinden elde edilen sonuçlar bu adımda değerlendirilmektedir. Bu bağlamda evrişimli sinir ağının ikili sınıflandırmadaki performansı, çok sınıflı sınıflandırmadaki performansına eşit olarak belirlenmiştir. Test setinin doğruluk oranı hem ikili hem de çok sınıflı sınıflandırmalar için %90,09 olarak belirlenmiştir. İkili sınıflandırmada test kaybı değeri %8,30 iken, çok sınıflı sınıflandırmada bu oran %8,34 olmuştur. Hem ikili hem de çok sınıflı sınıflandırmalar için doğrulama kaybı değeri %12,70 ve eğitim kaybı değeri %12,75 olarak belirlenmiştir. Doğrulama setinde

doğruluk oranı ikili sınıflandırmada %90,05 iken, çok sınıflı sınıflandırmada bu oran %90,08 olmuştur. Eğitim doğruluk oranı ise hem ikili hem de çok sınıflı sınıflandırmalar için %90,43 olarak belirlenmiştir.

Çizelge 5.5. UNSW-NB15 veri setinde CNN modeli ile elde edilen sonuçlar.

Ölçütler	**İkili Sınıflandırma**	**Çok Sınıflı Sınıflandırma**
Eğitim Doğruluğu	90.43%	90.43%
Değerlendirme Doğruluğu	90.05%	90.08%
Test Doğruluğu	90.09%	90.09%
Eğitimde Kayıp	12.75%	12.75%
Değerlendirmede Kayıp	12.70%	12.70%
Testte Kayıp	8.30%	8.34%

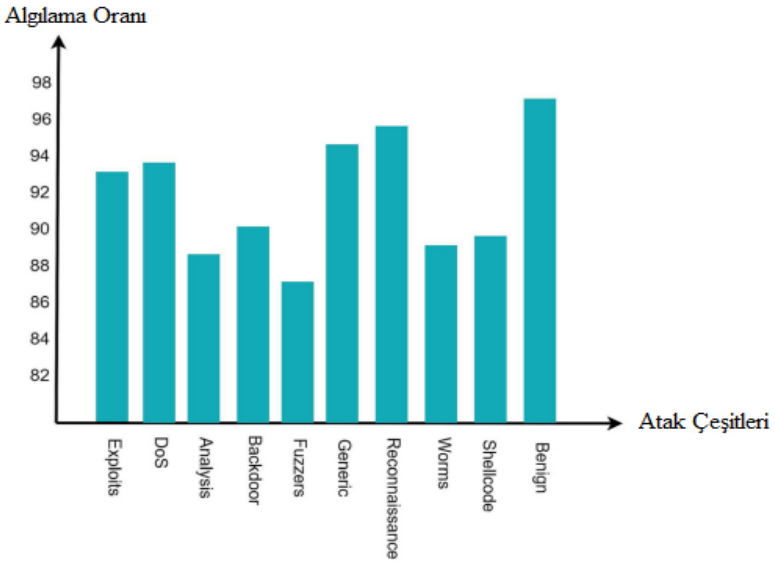
UNSW-NB15 veri seti üzerinde CNN + LSTM modeli kullanılarak yapılan ikili ve çok sınıflı sınıflandırma işlemlerinin sonuçları Çizelge 5.6'da sunulmaktadır. CNN + LSTM modeli ikili sınıflandırma için test setinde %93,21, çoklu sınıflandırma için %92,9 doğruluk elde etmiştir. Test setindeki kayıp değer ikili ve çok sınıflı sınıflandırmalar için sırasıyla %6,21 ve %6,28 olarak belirlenmiştir. Doğrulama setinde hem ikili hem de çok sınıflı sınıflandırma için doğruluk değeri %93,11 olarak belirlendi. Doğrulama setindeki kayıp değer, ikili ve çok sınıflı sınıflandırmalar için sırasıyla %5,89 ve %5,98 olarak belirlenmiştir. Eğitim setindeki doğruluk değeri ikili ve çok sınıflı sınıflandırmalar için sırasıyla %93.84 ve %93.26 olarak belirlenmiştir. Eğitim setindeki eksik değer ikili ve çok sınıflı sınıflandırmalar için sırasıyla %5,19 ve %6,07 olarak belirlenmiştir.

Çizelge 5.6. UNSW-NB15 veri setinde CNN+LSTM modeli ile bulunan sonuçlar.

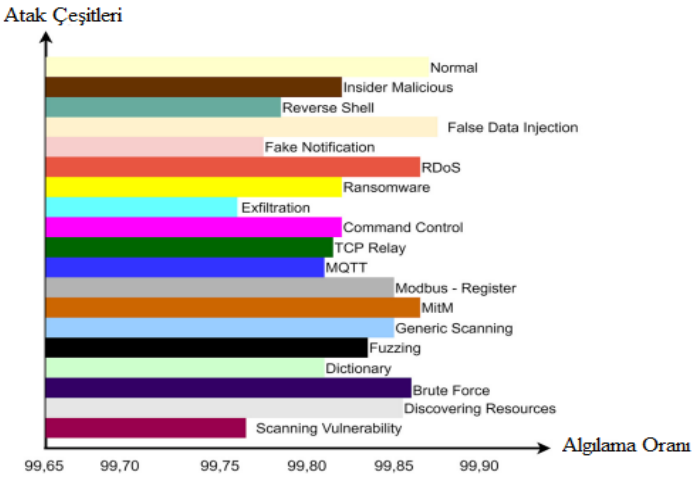
Ölçütler	**İkili Sınıflandırma**	**Çok Sınıflı Sınıflandırma**
Eğitim Doğruluğu	93.84%	93.26%
Değerlendirme Doğruluğu	93.11%	93.11%
Test Doğruluğu	93.21%	92.90%
Eğitimde Kayıp	5.19%	6.07%
Değerlendirmede Kayıp	5.89%	5.98%
Testte Kayıp	6.21%	6.28%

Kullanılan veri setlerinde saldırı türlerinin sınıflandırılmasında elde edilen doğruluk oranları Şekil 5.5 ve Şekil 5.6'da sunulmuştur. Şekil 4.5'de görüldüğü gibi UNSW-NB15 veri setinde optimum doğruluk oranı ile tespit edilen ilk üç saldırı türü Benign,

Keşif ve Genel saldırı türleridir. Şekil 5.5 ve Çizelge 4.2 birlikte dikkatlice incelendiğinde, sayısı az olan saldırı türlerinin isabetli tespit oranlarının, sayısı yüksek olan saldırı türlerine göre daha düşük olduğu görülmektedir. Derin öğrenme yöntemleri kullanılarak geliştirilen saldırı tespit sistemlerinde daha iyi sonuç alınabilmesi için kullanılan veri miktarının fazla olması ve veri seti içerisindeki saldırı türlerinin dengeli dağılması önemlidir.



Şekil 5.5. UNSW-NB15 veri setindeki atak paketlerinin doğru algılanma oranları.



Şekil 5.6. X-IIoTID veri setindeki atak paketlerinin doğru algılanma oranları.

X-IIoTID veri setinde CNN modeli, test setinde çok sınıflı sınıflandırma için %99,26 ve ikili sınıflandırma için %99,15 doğruluk elde etmiştir. Test setinde hem ikili hem de çok sınıflı sınıflandırma için kayıp değer %0,41 olarak belirlenmiştir. CNN modeli doğrulama setinde ikili sınıflandırma için %99,18 ve çok sınıflı sınıflandırma için %99,09 doğruluk elde etmiştir. Doğrulama setinde kayıp değer ikili sınıflandırma için %0,27 ve çok sınıflı sınıflandırma için %0,32 olarak belirlenmiştir. Ayrıca, CNN yaklaşımı, eğitim setinde hem ikili hem de çok sınıflı sınıflandırma için %99.11'lik bir doğruluk elde etmiştir. Eğitim setinde ikili sınıflandırma için kayıp değer %0,56, çok sınıflı sınıflandırma için %0,59 olarak belirlenmiştir.

X-IIoTID veri setinde, önerilen diğer model olan LSTM, test setinde ikili sınıflandırma için %99,05 ve çok sınıflı sınıflandırma için %98,91 doğruluk elde etmiştir. LSTM modelinin ulaştığı doğrulama setindeki doğruluk değeri, ikili ve çok sınıflı sınıflandırmalar için sırasıyla %98,97 ve %99,02 olarak belirlenmiştir. LSTM modeli, eğitim setinde hem ikili hem de çok sınıflı sınıflandırma için %98.99 doğruluk elde etti. Test setindeki kayıp değer ikili ve çok sınıflı sınıflandırmalar için sırasıyla %0,52 ve %0,92 olarak belirlenmiştir. LSTM modeli, doğrulama setinde ikili sınıflandırma için %0,78 ve çok sınıflı sınıflandırma için %0,72'lik bir kayıp değeri vermiştir. Eğitim setinde ise hem ikili hem de çok sınıflı sınıflandırma için kayıp değer %0,75 olarak belirlenmiştir.

Son olarak tarafımızca önerilen CNN + LSTM modeli X-IIoTID veri seti üzerinde çalıştırılmıştır. Test setinde ikili sınıflandırma için %99,84, çok sınıflı sınıflandırma için %99,80 doğruluk elde etmiştir. Test setinde hem ikili hem de çok sınıflı sınıflandırma için CNN + LSTM modelinin kayıp değeri %0,12 olarak belirlenmiştir. Doğrulama setinde hem ikili hem de çok sınıflı sınıflandırma için %99,78 doğruluk elde edilirken, aynı sette ikili sınıflandırma için kayıp değer %0,14 ve çok sınıflı sınıflandırma için %0,15 olarak belirlendi. Eğitim setinde ise ikili sınıflandırma için %99.81, çok sınıflı sınıflandırma için bu oran %99.79 doğruluk elde edilmiştir. Eğitim setinde ikili sınıflandırma için kayıp değer %0,11, çok sınıflı sınıflandırma için %0,19 olarak belirlenmiştir.

CNN modeli aracılığıyla X-IIoTID veri setindeki ikili ve çok sınıflı sınıflandırma prosedürleri için elde edilen değerler Çizelge 5.7'de sunulmaktadır. LSTM modeli ile elde edilen sonuçlar Çizelge 5.8'de, CNN + LSTM modeli ile elde edilen sonuçlar ise ise Çizelge 5.9'da sunulmuştur.

Çizelge 5.7. X-IIoTID veri setinde CNN modeli ile elde edilen sonuçlar.

Ölçütler	**İkili Sınıflandırma**	**Çok Sınıflı Sınıflandırma**
Eğitim Doğruluğu	99.11%	99.11%
Değerlendirme Doğruluğu	99.18%	99.09%
Test Doğruluğu	99.10%	99.26%
Eğitimde Kayıp	0.56%	0.59%
Değerlendirmede Kayıp	0.27%	0.32%
Testte Kayıp	0.41%	0.41%

Çizelge 5.8. X-IIoTID veri setinde LSTM modeli ile elde edilen sonuçlar.

Ölçütler	**İkili Sınıflandırma**	**Çok Sınıflı Sınıflandırma**
Eğitim Doğruluğu	98.97%	98.99%
Değerlendirme Doğruluğu	98.99%	99.02%
Test Doğruluğu	99.05%	98.91%
Eğitimde Kayıp	0.75%	0.75%
Değerlendirmede Kayıp	0.78%	0.72%
Testte Kayıp	0.52%	0.92%

Çizelge 5.9. X-IIoTID veri setinde CNN+LSTM modeli ile bulunan sonuçlar.

Ölçütler	**İkili Sınıflandırma**	**Çok Sınıflı Sınıflandırma**
Eğitim Doğruluğu	99.81%	99.79%
Değerlendirme Doğruluğu	99.78%	99.78%
Test Doğruluğu	99.84%	99.80%
Eğitimde Kayıp	0.11%	0.19%
Değerlendirmede Kayıp	0.14%	0.15%
Testte Kayıp	0.12%	0.12%

UNSW-NB15 veri setinde, LSTM modeli ikili sınıflandırma için %91,08'lik bir geri çağırma değeri, %90,98'lik bir kesinlik değeri ve %91,02'lik bir F1-Score değeri elde ederken, çok sınıflı sınıflandırmada %91,04'lük bir kesinlik ve geri çağırma değeriyle birlikte %91.04'lik F1-Score değerine ulaşmıştır. XIIoTID veri setinde, LSTM modeli ikili sınıflandırma için %99,01 geri çağırma değeri, %99,03 kesinlik değeri ve %99,01 F1-Skor değeri elde ederken, çok sınıflı sınıflandırmada %98,82 geri çağırma değeri, %98,86 kesinlik değeri elde etmiştir. Ayrıca eğitim setinde çok sınıflı sınıflandırma

için %98,84'lük bir F1 Puanı değeri elde edilmiştir. Çizelge 5.10'da her iki veri setinde ikili ve çok sınıflı sınıflandırma işlemlerinde LSTM modeli ile elde edilen kesinlik, geri çağırma ve F1-score değerleri gösterilmiştir.

Çizelge 5.10. LSTM modeli ile elde edilen ölçüt değerleri.

	İkili Sınıflandırma	Çok Sınıflı Sınıflandırma
UNSW-NB15		
Kesinlik	%90.98	%91.04
Geri Çağırma	%91.08	%91.04
F1-Score	%91.02	%91.04
X-IIoTID		
Kesinlik	%99.01	%98.86
Geri Çağırma	%99.03	%98.82
F1-Score	%99.03	%98.84

UNSW-NB15 veri setinin test adımında, CNN modeli, ikili sınıflandırma için %90,07'lik bir geri çağırma değeri, %90,06'lık bir kesinlik değeri ve %90,06'lık bir F1-Score değeri elde ederken, çok sınıflı sınıflandırmada %90,05'lik bir kesinlik ve geri çağırma değeri ile %90,05'lik F1-Score değerine ulaşmıştır. X-IIoTID veri setinde CNN modeli, ikili sınıflandırma için %99,05 geri çağırma değeri, %99,02 kesinlik değeri ve %99,03 F1-Skor değeri elde ederken, çok sınıflı sınıflandırma işleminde %99,14 hatırlama değeri, %99,18 kesinlik değeri elde etti. Ayrıca eğitim setinde çok sınıflı sınıflandırma için %99,16'lık bir F1 Puanı değeri elde edilmiştir. Çizelge 5.11'de her iki veri setinde ikili ve çok sınıflı sınıflandırma işlemlerinde CNN modeli ile elde edilen kesinlik, geri çağırma ve F1-score değerleri gösterilmiştir.

Çizelge 5.11. CNN modeli ile elde edilen ölçüt değerleri.

	İkili Sınıflandırma	Çok Sınıflı Sınıflandırma
UNSW-NB15		
Kesinlik	%90.06	%90.05
Geri Çağırma	%90.07	%90.05
F1-Score	%90.06	%90.05
X-IIoTID		
Kesinlik	%99.02	%99.18
Geri Çağırma	%99.05	%99.14
F1-Score	%99.03	%99.16

UNSW-NB15 veri setinde, hibrit CNN + LSTM modeli, ikili sınıflandırma için %93,10 geri çağırma değeri, %92,91 kesinlik değeri ve %93,00 F1-Skor değeri elde ederken, çok sınıflı sınıflandırmada %92,45'lik geri çağırma, kesinlik ve F1-Score değeri elde etmiştir. Hibrit CNN + LSTM modeli, X-IIoTID veri setinde ikili sınıflandırma için %99,55 geri çağırma değeri, %99,67 kesinlik değeri ve %99,60 F1-Score değeri elde ederken, çok sınıflı sınıflandırma işleminde %99,72 geri çağırma, %99,38 kesinlik değeri ve %99,54 F1-Score değeri tespit edilmiştir. Çizelge 5.12'de her iki veri setinde ikili ve çok sınıflı sınıflandırma işlemlerinde CNN+LSTM modeli ile elde edilen kesinlik, geri çağırma ve F1-score değerleri gösterilmiştir.

Çizelge 5.12. CNN+LSTM modeli ile elde edilen ölçüt değerleri.

	İkili Sınıflandırma	Çok Sınıflı Sınıflandırma
UNSW-NB15		
Kesinlik	%92.91	%92.45
Geri Çağırma	%93.10	%92.45
F1-Score	%93.00	%90.45
X-IIoTID		
Kesinlik	%99.67	%99.38
Geri Çağırma	%99.55	%99.72
F1-Score	%99.60	%99.54

Elde edilen sonuçlara göre CNN + LSTM modeli, UNSW-NB15 ve X-IIoTID veri setlerinde ikili ve çok sınıflı sınıflandırma işlemleri için hem literatürdeki diğer çalışmalara hem de bu çalışmada kullanılan diğer iki modele göre daha yüksek doğruluk değerine ulaşmıştır. Çalışma, CNN ve LSTM. Çalışmada yapılan testler sonucunda IIoT ağındaki saldırıları tespit etmek için üretilen en son veri setlerinden biri olan X-IIoTID veri setinde elde edilen doğruluk değerlerinin UNSW-NB15 veri setindeki değerlerden daha yüksek olduğu gözlemlendi. Sonuç olarak, X-IIoTID'nin karmaşık IIoT ağları için iyi tanımlanmış bir veri kümesi olduğu sonucuna varıldı.

Hem literatürdeki UNSW-NB15 ve X-IIoTID veri setleri kullanılarak yapılan çalışmalardan hem de önerdiğimiz modellerden elde edilen saldırı algılamadaki doğruluk sonuçları Çizelge 5.13 ve Çizelge 5.14'de sunulmaktadır.

Çizelge 5.13. UNSW-NB15 veri setinde literatürdeki çalışmalar ile tez çalışmasında elde edilen sonuçların karşılaştırılması.

Kaynak Numarası	**Kullanılan Model**	**İkili Sınıflandırma Sonucu**	**Çok Sınıflı Sınıflandırma Sonucu**
[69]	ANN	%81.34	-
[70]	GA – RF	%87.61	-
[70]	GA -ET	-	%77.64
[71]	PSO – GBM	%86.68	-
[72]	VLSTM	-	-
[73]	ELM	-	%70.52
[74]	DL – ANN	%76.1	%65.10
[75]	ANN	%84	-
[76]	ANN	%83.9	-
[77]	GA – SVM	%86.38	-
[77]	GWO – SVM	%84.48	-
[77]	FFA – SVM	%85.42	-
[78]	TS – RF	%83.12	-
[79]	IG – TS	%85.78	-

[80]	GA - DT	%81.42	-
[81]	XGB – LR	%75.51	%72.53
[82]	SVM	%85.99	%75.77
[83]	IG – TREE	%84.83	-
[84]	DL – LSTM	%85.42	-
[85]	DL - LSTM	%80.72	%72.26
[86]	CNN – RNN	%86.64	-
[87]	GA – RF	%86.70	-
[88]	GA – RF	-	%64.23
Önerilen Model	CNN	%90.09	%90.09
Önerilen Model	LSTM	%91.14	%91.10
Önerilen Model	CNN+LSTM	%93.21	%92.90

Çizelge 5.14. X-IIoTID veri setinde literatürdeki çalışmalar ile tez çalışmasında elde edilen sonuçların karşılaştırılması

Kaynak Numarası	**Kullanılan Model**	**İkili Sınıflandırma Sonucu**	**Çok Sınıflı Sınıflandırma Sonucu**
[66]	DL – ML	%99.54	%99.45
[67]	DL	%99.79	-
[68]	CDAE - DNN	%98.33	%97.21
Önerilen Model	CNN	%99.15	%99.26
Önerilen Model	LSTM	%99.05	%98.91
Önerilen Model	CNN+LSTM	%99.84	%99.80

BÖLÜM 6

SONUÇLAR VE TARTIŞMA

Ağ tabanlı denetim kullanan birçok IDS, bir düğümün güvenliğinin ihlal edilip edilmediğini belirlemek için ağ etkinliğini inceler. Bu denetim genel (trafik veya frekans analizi) veya protokole özel (derin paket incelemesi) olabilir. Önerilen saldırı tespit modelleri, IIoT ağlarında güvenliği sürekli olarak değerlendirmekten sorumludur. Kritik altyapıya sahip endüstriyel ağlarda ağ etkinliğinin yalnızca genel denetimi, güvenliği ihlal edilmiş HMI / PLC veya PLC yeniden programlama saldırısı gibi diğer saldırıları algılamada her zaman etkili olmadığından, protokole özgü denetim gerçekleştiren algılama mekanizmaları paralel olarak kullanılmaktadır. Her iki yöntem de derin öğrenme yöntemlerini temel aldığı için yeni tip saldırıları tespit etme kabiliyetine sahiptir. Anomali tabanlı tespit sistemlerinin en önemli avantajı, sıfır gün saldırılarını tespit etme yetenekleridir. Yüksek doğruluk, düşük yanlış alarm oranı, gerçek zamanlı iletişim gereksinimleri ve düşük ek yük arasında denge kurmak için çeşitli tekniklerin bir kombinasyonuna ihtiyaç vardır. Bu nedenle ortam üzerinde yüksek ek yüke neden olmadan gerçek zamanlı durumlarda çalışabilir. Ağ ortamına her zaman yalnızca bir ağa bağlı aygıtın erişim kazanmasına özen gösterilir. Böylece eşzamanlı ve eşzamansız verilerin iletimi karışmaz ve kesin iletişim zamanlaması elde edilebilir.

Siber saldırıların büyük bir yüzdesinin içerideki kullanıcılar tarafından tetiklendiği bilindiği için saldırı tespitinde, daha fazla karmaşıklık ortaya çıkmaktadır. Bu nedenle çevre savunması tek başına sistemi savunamaz. Bu gibi durumlarda, kişinin karşı karşıya kaldığı soru, sistemin kendi dinamiklerinde devam etmekte olan bir saldırının yeterli göstergesinin olup olmadığıdır. Bu faaliyetler yelpazesine rağmen, bunların yarısının özünde insan hatası olduğu kanıtlanmıştır. Bu nedenle, ana akım siber güvenlik uygulamasının yararlanabileceği yolları belirlemek için insan hatasıyla ilgili olayların hacmine dayalı olarak siber güvenliğin insani yönlerine ilişkin araştırmaların

sayısı artırılmalıdır. Güvenlik önlemleri, kalıcı saldırganların çevre koruması ne olursa olsun eninde sonunda erişim elde edeceklerini göz ardı etme eğilimindedir. Modern güvenlik çözümlerinin ana hedeflerinden biri, saldırganların sistem içine erişim sağladıktan sonra faaliyetlerini tespit edip bozabilecek yeni yöntemler geliştirmek olacaktır. İzinsiz giriş tespit/önleme sistemlerinde, veri hırsızlığı saldırılarını tespit edebilen, önleyebilen ve azaltabilen yeni stratejilerin uygulanmasına özel dikkat gösterilmelidir.

Bu tez çalışmasında, önerilen saldırı tespit sistemi, IIoT ağlarında izinsiz girişleri yüksek başarımla tespit etmiştir. Önerilen saldırı tespit sistemlerinde CNN, LSTM ve hibrit CNN+LSTM derin öğrenme yöntemleri kullanılmıştır. Bu araştırmadaki anormal örüntüleri belirlemek için X-IIoTID ve UNSW-NB15 veri setleri kullanılmıştır. Özellik seçim sürecinde herhangi bir ek makine öğrenimi yöntemi kullanılmamıştır; derin öğrenme yöntemleri bu görevi tek başına gerçekleştirmiştir. Önerilen derin öğrenme modellerinin deneysel sonuçları, aynı veri setleri kullanılarak geliştirilen önceki yöntemlerden daha üstündür. Bu sonuçlar, derin öğrenme yöntemlerinin büyük ve karmaşık veri kümelerindeki anormal olayları tespit etmedeki üstünlüğünü göstermektedir.

Gelecek çalışmalar için farklı güncel veri setlerinde önerilen derin öğrenme tabanlı modellerin algılama doğruluğunu test etmek ile gerçek zamanlı siber saldırıların tespitine yönelik performanslarının belirlenmesi kritik altyapıların güvenliğinin artırılmasına katkı sağlayacaktır.

KAYNAKLAR

1. Mohammed, A. S., Anthi, E., Rana, O., Saxena, N., and Burnap, P., "Detection and mitigation of field flooding attacks on oil and gas critical infrastructure communication." *Computers & Security*, 124, 103007, (2022).

2. Mohlabeng, M. R., and Osunmakinde, I. O., "Towards Railway Cable Infrastructure Protection: Turning Cross-Sectional Explorative Analytics to Answers." *In Distributed Computing and Intelligent Technology: 19th International Conference, ICDCIT 2023*, Bhubaneswar, India, January 18–22, 2023, Proceedings (pp. 270-289). Cham: Springer Nature Switzerland, (2023).

3. Khaustova, V., Tirlea, M. R., Dandara, L., Trushkina, N., and Birca, I., "Development of critical infrastructure from the point of view of information security". *Univers strategic*, 53, (1), (2023).

4. Salvador, L. C. R., Dai, N. H. P., and Zoltán, R., (2023, "SCADA Systems: Security Concerns and Countermeasures." *In 2023 IEEE 21st World Symposium on Applied Machine Intelligence and Informatics (SAMI)* (pp. 000251-000254). IEEE, (2023).

5. Babayigit, B., and Abubaker, M., "Industrial Internet of Things: A Review of Improvements Over Traditional SCADA Systems for Industrial Automation." *IEEE Systems Journal*, (2023).

6. Altunay, H. C., Albayrak, Z., Özalp, A. N., and Çakmak, M., "Analysis of anomaly detection approaches performed through deep learning methods in SCADA systems." *In 2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, (pp. 1-6). IEEE, (2021).

7. Zhu, Q., Zhang, G., Luo, X., and Gan, C., "An industrial virus propagation model based on SCADA system." *Information Sciences*, 630, 546-566, (2023).

8. Das, S. K., Benkhelifa, F., Sun, Y., Abumarshoud, H., Abbasi, Q. H., Imran, M. A., and Mohjazi, L., "Comprehensive Review on ML-based RIS-enhanced IoT Systems: Basics, Research Progress and Future Challenges." *Computer Networks*, 109581, (2023).

9. Rekha, S., Thirupathi, L., Renikunta, S., and Gangula, R., "Study of security

issues and solutions in Internet of Things (IoT).” *Materials Today: Proceedings*, 80, 3554-3559, (2023).

10. Centenaro, M., Granelli, F., and Vangelista, L., " A Survey on Technologies, Standards and Open Challenges in Satellite IoT ", IEEE Communications Survey & Tutorials, 23, (2021).

11. Aversano, L., Bernardi, M. L., Cimitile, M., and Pecori, R., " A Systematic Review on Deep Learning Approaches for IoT Security ", *Computer Science Review*, 40, (2021).

12. Yazdinejad, A., Dehghantanha, A., Parizi, R. M., Srivastava, G., and Karimipour, H., "Secure intelligent fuzzy blockchain framework: Effective threat detection in iot networks.", *Computers in Industry*, 144, 103801, (2023).

13. Qi, Q., Xu, Z., and Rani, P., "Big data analytics challenges to implementing the intelligent Industrial Internet of Things (IIoT) systems in sustainable manufacturing operations.", *Technological Forecasting and Social Change*, 190, 122401, (2023).

14. Zhang, F., Wang, H., Zhou, L., Xu, D., and Liu, L., "A blockchain-based security and trust mechanism for AI-enabled IIoT systems.", *Future Generation Computer Systems*, 146, 78-85, (2023).

15. Irshad, A., Mallah, G. A., Bilal, M., Chaudhry, S. A., Shafiq, M., and Song, H., "SUSIC: A Secure User Access Control mechanism for SDN-enabled IIoT and Cyber Physical Systems.", *IEEE Internet of Things Journal*, (2023).

16. Oñate, W., and Sanz, R., "Analysis of architectures implemented for IIoT.", *Heliyon*, e12868, (2023).

17. Chen, H., Jeremiah, S. R., Lee, C., and Park, J. H., "A Digital Twin-Based Heuristic Multi-Cooperation Scheduling Framework for Smart Manufacturing in IIoT Environment.", *Applied Sciences*, 13(3), 1440, (2023).

18. Özalp, A. N., Albayrak, Z., Çakmak, M., & Özdoğan, E., "Layer-based examination of cyber-attacks in IoT", *In 2022 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*,1-10, IEEE, (2022).

19. Özarpa, C., Aydın, M., and Avcı, İ., " International Security Standards for

Critical Oil, Gas, and Electricity Infrastructures in Smart Cities: A Survey Study", In the Proceedings of the third International Conference on Smart City Applications, (2021).

20. Lv, Z., Qiao, L., Singth, A. K., and Wang, Q., "AI-Empowered IoT Security for Smart Cities ", ACM Transactions on Internet Technology, 1–21, (2021).

21. Andrei, C. C., Tudor, G., and Calin, M. A., " Industrial Internet of Things (IIoT) Integration in Power Grids ", 9th International Conference on Modern Power Systems (MPS), IEEE, Romania, (2021).

22. Khraisat, I., Gondal, P., Vamplew, P., and Kamruzzaman, J., "Survey of Intrusion Detection Systems: Techniques, Datasets and Challenges ", Cyber Security, 2, 1-22, (2019).

23. Qin, W., Chen, S., and Peng, M., " Recent Advances in Industrial Internet: Insights and Challenges ", Digial Communications and Networks, Elsevier, 1-13, (2020).

24. Zang, X. D., "Machine Learning", In A Matrix Algebra Approach to Artificial Intelligence, Singapore: Springer, 223-440 (2020).

25. Pengfei, S., Pengju, L., Qi, L., Chenxi, L., Xiangling, L., Ruochen, H., and Jinpeng, C., "Dl-ids: Extracting features using cnn-lstm hybrid network for intrusion detection system", Security and Communication Networks Hindawi, (2020).

26. Aldweesh, A. A., Derhab, A., and Emam, A. Z., " Deep Learning Approaches for Anomaly-Based Intrusion Detection Systems: A Survey, Taxonomy, and Open Issues ", Knowledge-Based Systems, Elsevier, 189, (2020).

27. Öncül, A. B., and Çelik, Y., "A hybrid deep learning model for classification of plant transcription factor proteins, Signal Image and Video Processing ", Springer, 1-7, (2022).

28. Alabadi, M. S., and Çelik Y., " Anomaly Detection for Cyber-Security Based on Convolution Neural Network: A Survey ", International Congress on Human-Computer Interaction, Optimization and Robotic Applications, IEEE, (2020).

29. Suthar, K., and He, Q. P., " Multiclass Moisture Classification in Woodchips Using IIoT Wi-Fi and Machine Learning Techniques ", Computers and

Chemical Engineering, Elsevier, 151, (2021).

30. Shafiq, M., Tian, Z., Bashir, A. K., Du, X., and Guizani, M., " IoT Malicious Trafic Identification using Wrapper-Based Feature Selection Mechanisms ", Computer Security, 94, (2020).

31. Zhiang, X., Yijia, G., Chinmay, C., Qiaozhi, H., Shengbo, C., and Keping, Y., "A simple federated learning-based scheme for security enhancement over internet of medical things,", Journal of Biomedical and Health Informatics, IEEE, 27, (2023).

32. Bhavsar, M., Roy, K., Kelly, J., and Olusola, O., "Anomaly-based intrusion detection system for IoT application.", *Discover Internet of Things*, 3(1), 5, (2023).

33. Xiao, J., Yang, L., Zhong, F., Chen, H., and Li, X., "Robust anomaly-based intrusion detection system for in-vehicle network by graph neural network framework.", *Applied Intelligence*, 53(3), 3183-3206, (2023).

34. Hnamte, V., and Hussain, J., "DCNNBiLSTM: An efficient hybrid deep learning-based intrusion detection system.", *Telematics and Informatics Reports*, 10, 100053, (2023).

35. Zhao, N., Zhao, X., Xu, N., and Zhang, L., "Resilient event-triggered control of connected automated vehicles under cyber attacks.", *IEEE/CAA Journal of Automatica Sinica*, (2023).

36. Duo, W., Zhou, M., and Abusorrah, A., "A survey of cyber attacks on cyber physical systems: Recent advances and challenges.", *IEEE/CAA Journal of Automatica Sinica*, 9(5), 784-800, (2022).

37. Inayat, U., Zia, M. F., Mahmood, S., Khalid, H. M., and Benbouzid, M., "Learning-based methods for cyber attacks detection in IoT systems: a survey on methods, analysis, and future prospects.", *Electronics*, 11(9), 1502, (2022).

38. Lian, Z., Shi, P., Lim, C. C., and Yuan, X., "Fuzzy-model-based lateral control for networked autonomous vehicle systems under hybrid cyber-attacks.", *IEEE Transactions on Cybernetics*, (2022).

39. Abdullahi, M., Baashar, Y., Alhussian, H., Alwadain, A., Aziz, N., Capretz, L. F., and Abdulkadir, S. J,, "Detecting cybersecurity attacks in internet of things using artificial intelligence methods: A systematic literature

review.", *Electronics*, 11(2), 198, (2022).

40. Makkar, A., and Park, J. H., "SecureCPS: Cognitive inspired framework for detection of cyber attacks in cyber–physical systems.", *Information processing & management*, 59(3), 102914, (2022).

41. Yang, C., Liang, J., and Chen, X., "Distributed event-based H∞ consensus filtering for 2-D TS fuzzy systems over sensor networks subject to DoS attacks.", *Information Sciences*, 641, 119079, (2023).

42. Wen, G., Wang, P., Lv, Y., Chen, G., and Zhou, J., "Secure consensus of multi-agent systems under denial-of-service attacks.", *Asian Journal of Control*, 25(2), 695-709, (2023).

43. Ali, T. E., Chong, Y. W., and Manickam, S., "Machine Learning Techniques to Detect a DDoS Attack in SDN: A Systematic Review.", *Applied Sciences*, 13(5), 3183, (2023).

44. Bouke, M. A., Abdullah, A., ALshatebi, S. H., Abdullah, M. T., and El Atigh, H., "An intelligent DDoS attack detection tree-based model using Gini index feature selection method.", *Microprocessors and Microsystems*, 98, 104823, (2023).

45. Al-Juboori, S. A. M., Hazzaa, F., Jabbar, Z. S., Salih, S., and Gheni, H. M., "Man-in-the-middle and denial of service attacks detection using machine learning algorithms.", *Bulletin of Electrical Engineering and Informatics*, 12(1), 418-426, (2023).

46. Al-Shareeda, M. A., and Manickam, S., "Man-in-the-middle attacks in mobile ad hoc networks (MANETs): Analysis and evaluation.", *Symmetry*, 14(8), 1543, (2022).

47. Thankappan, M., Rifà-Pous, H., and Garrigues, C., "Multi-Channel Man-in-the-Middle attacks against protected Wi-Fi networks: A state of the art review.", *Expert Systems with Applications*, 118401, (2023).

48. Ansari, M. F., Sharma, P. K., an Dash, B., "Prevention of phishing attacks using AI-based Cybersecurity Awareness Training.", *Prevention*, (2022).

49. Sharma, P., Dash, B., and Ansari, M. F., "Anti-phishing techniques–a review of Cyber Defense Mechanisms.", *IJARCCE*, 11(7), 153-160, (2022).

50. Iglesias, P., Sicilia, M. A., and García-Barriocanal, E., "Detecting Browser Drive-By Exploits in Images Using Deep Learning.", *Electronics*, 12(3), 473, (2023).

51. Javed, A., Ikwu, R., Burnap, P., Giommoni, L., and Williams, M. L., "Disrupting drive-by download networks on Twitter.", *Social Network Analysis and Mining*, 12(1), 117, (2022).

52. Iglesias, P., Sicilia, M. A., and García-Barriocanal, E., "Detecting Browser Drive-By Exploits in Images Using Deep Learning.", *Electronics*, 12(3), 473, (2023).

53. Mohanty, S., and Acharya, A. A., "MFBFST: Building a stable ensemble learning model using multivariate filter-based feature selection technique for detection of suspicious URL.", *Procedia Computer Science*, 218, 1668-1681, (2023).

54. Aslan, Ö., Aktuğ, S. S., Ozkan-Okay, M., Yilmaz, A. A., and Akin, E., "A comprehensive review of cyber security vulnerabilities, threats, attacks, and solutions.", *Electronics*, 12(6), 1333, (2023).

55. Guembe, B., Azeta, A., Misra, S., Osamor, V. C., Fernandez-Sanz, L., and Pospelova, V., "The emerging threat of ai-driven cyber attacks: A Review.", *Applied Artificial Intelligence*, 36(1), 2037254, (2022).

56. Alghawazi, M., Alghazzawi, D., and Alarifi, S., "Detection of sql injection attack using machine learning techniques: a systematic literature review.", *Journal of Cybersecurity and Privacy*, 2(4), 764-777, (2022).

57. Hallo, M., and Suntaxi, G., "A survey on SQL injection attacks, detection and prevention techniques-a tertiary study.", *International Journal of Security and Networks*, 17(3), 193-202, (2022).

58. Altulaihan, E. A., Alismail, A., and Frikha, M., "A Survey on Web Application Penetration Testing.", *Electronics*, 12(5), 1229, (2023).

59. Marashdeh, Z., Suwais, K., and Alia, M., "A survey on sql injection attack: Detection and challenges.", *In 2021 International Conference on Information Technology (ICIT)*, 957-962, IEEE, (2021).

60. Rodríguez, G. E., Torres, J. G., Flores, P., and Benavides, D. E., "Cross-site scripting (XSS) attacks and mitigation: A survey.", *Computer Networks*, 166,

106960, (2020).

61. Cui, Y., Cui, J., and Hu, J., "A survey on xss attack detection and prevention in web applications.", *In Proceedings of the 2020 12th International Conference on Machine Learning and Computing*, 443-449, (2020).

62. Huang, Y., Li, Y. J., and Cai, Z., "Security and privacy in metaverse: A comprehensive survey.", *Big Data Mining and Analytics*, 6(2), 234-247, (2023).

63. Ling, X., Wu, L., Zhang, J., Qu, Z., Deng, W., Chen, X., and Wu, Y., "Adversarial attacks against Windows PE malware detection: A survey of the state-of-the-art.", *Computers & Security*, 103134, (2023).

64. Gopinath, M., and Sethuraman, S. C., "A comprehensive survey on deep learning based malware detection techniques.", *Computer Science Review*, 47, 100529, (2023).

65. Akter, M. S., Shahriar, H., Ahamed, S. I., Gupta, K. D., Rahman, M., Mohamed, A., and Wu, F., "Case Study-Based Approach of Quantum Machine Learning in Cybersecurity: Quantum Support Vector Machine for Malware Classification and Protection, arXiv:2306.00284v1, (2023).

66. Hawawreh, M. A., Sitnikova, E., and Aboutorab, N., " X-IIoTID: A Connectivity-Agnostic and Device-Agnostic Intrusion Data Set for Industrial Internet of Things ", *IEEE Internet of Things Journal*, 9 (2022).

67. Makkar, A., Kim, T. W., Singh A. K., Kang, J., and Park, J. H., " SecureIIoT Environment: Federated Learning Empowered Approach for Securing IIoT from Data Breach ", *IEEE Transactions on Industrial Informatics*, (2022).

68. Hawawreh, M. A., Sitnikova, E., and Aboutorab, N., " Asynchronous Peer-to-Peer Federated Capability-Based Targeted Ransomware Detection Model for Industrial IoT ", *IEEE Access*, (2021).

69. Moustafa, N., and Slay, J., " The Evaluation of Network Anomaly Detection Systems: Statistical Analysis of the UNSW-NB15 Data Set and the Comparison with the KDD99 Data Set ", *Information Security Journal, Global Perspective*, 18-31 (2016).

70. Kasongo, S. M., ", An advanced intrusion detection system for iiot based on ga and tree based algorithms", *IEEE Access*, 9, (2021).

71. Liu, J., Yang, D., Lian, M., and Li, M., ", Research on intrusion detection based on particle swarm optimization in iot", *IEEE Access*, 9, (2021).

72. Zhou, X., Hu, Y., Liang, W., Ma, J., and Jin, Q., "Variational lstm enhanced anomaly detection for industrial big data", *IEEE Transactions on Industrial Informatics*, 17, (2021).

73. Gao, J., Chai, S., Zhang, B., and Xia, Y., ", Research on network intrusion detection based on incremental extreme learning machine and adaptive principal component analysis", *Energies*, 12 (2019).

74. Viayakumar, R., Alazab, M., Soman, K., Poornachandran, P., Al-Nemrat, A., and Venkatraman, S., "Deep learning approach for intelligent intrusion detection system", *IEEE Access*, 7 (2019).

75. Hanif, S., İlyas, T., and Zeeshan, M., "Intrusion detection in iot using artificial neural networks on unsw-15 dataset", *IEEE 16th International Conference Smart Cities, Improving Quality of Life Using ICT & IoT AI (HONET-ICT)*, 152-156 (2019).

76. Ketzaki E., Drosou, A. Papadopoulos, S., and Tzovaras, D., "A lightweighted ann architecture for the classification of cyber-threats in modern communication networks", *10th International Conference Networks of the Future (NoF)*, 17-24 (2019).

77. Almomani, O., "A feature selection model for network intrusion detection system based on pso, gwo, ffa and ga algorithms", *Symmetry*, 12, (2020).

78. Nazir, A., and Khan, R. A., ", A novel combinatorial optimization based feature selection method for network intrusion detection", *Computers & Security*, Elsevier, 102, (2021).

79. Zong, W., Chow, Y. W., and Susilo, W., "A two-stage classifier approach for network intrusion detection", *International Conference Information Security Practice and Experience Cham, Springer*, 329-340 (2018).

80. Khammassi, C., and Krichen, S., "n, A ga-lr wrapper approach for feature selection in network intrusion detection", *Computer & Security, Elsevier*, 255-270, (2017).

81. Kasongo, S. M., and Sun, Y., "Performance Analysis of Intrusion Detection Systems Using a Feature Selection Method on the UNSW-NB15 Dataset ",

Journal of big Data, Springer, (2020).

82. Jing, D. D., and Chen, H. B., " Svm based network intrusion detection for the unsw-nb15 dataset", *13th International Conference on ASIC (ASICON), IEEE*, (2019).

83. Kumar, V. V., Sinha, D., Das, A. K., Pandey, S. C., and Goswami, R. T., " An integrated rule based intrusion detection system: Analysis on unsw-nb15 data set and the real time online dataset", *Cluster Computing, Springer*, 1397-1418 (2020).

84. Aleesa, A., Younis, M., Mohammed, A. A and sahar, N. M., "Deep intrusion detection system with enhanced unsw-nb15 dataset based on deep learning techniques", *Journal of Engineering Science and Technology, Elsevier*, 711-727 (2021).

85. Elijah, A. V., Abdullah, A., Ihanjhi, N., Supramaniam, M., and Abdullateef, B., "Ensemble and deep-learning methods for two-class and multi-attack anomaly intrusion detection: An empirical study", *International Journal of Advanced Computer Science and Applications, Elsevier*, 520-528 (2019).

86. Wu, P. P., Guo, H., and Moustafa, N., " Pelican: A deep residual network for network intrusion detection", *50th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W)*, 55-62, (2020).

87. Assiri, A., "Anomaly classification using genetic algorithm-based random forest model for network attack detection", *Computers, Materials and Contunia*, 767-778 (2021).

88. Khammassi, C., and Krichen, S., " A nsga2-lr wrapper approach for feature selection in network intrusion detection", *Computer Networks, Elsevier*, 172, (2020).

89. Bace, R. G., "Intrusion detection.", *Sams Publishing*, (2000).

90. McHugh, J., "Intrusion and intrusion detection. International Journal of Information Security", 1, 14-35, (2001).

91. Gendreau, A. A., and Moorman, M., "Survey of intrusion detection systems towards an end to end secure internet of things.", *In 2016 IEEE 4th international conference on future internet of things and cloud (FiCloud)*, 84-90, IEEE, (2016).

92. Spadaccino, P., and Cuomo, F., "Intrusion detection systems for iot: opportunities and challenges offered by edge computing.", *Journal on Future and Evolving Technologies*, (2022).

93. Owais, S., Snasel, V., Kromer, P., and Abraham, A., "Survey: using genetic algorithm approach in intrusion detection systems techniques." *In 2008 7th Computer Information Systems and Industrial Management Applications*, 300-307, (2008).

94. Patel, A., Taghavi, M., Bakhtiyari, K., and Celestino Júnior, J., "An intrusion detection and prevention system in cloud computing: A systematic review.", *Journal of Networkand Computer Applications*, 36, 25 – 41, (2013).

95. SANS Institute., "Host- vs. Network-Based Intrusion Detection Systems. SANS Institute.",https://www.giac.org/paper/gsec/1377/host-vs-networkbased-intrusion-detection-systems/102574, (2005).

96. Liao, H.-J., Lin, C.-H., Lin, Y.-C., and Tung, K.-Y., "Intrusion detection system: A comprehensive review.", *Journal of Networkand Computer Applications* 36, 16-24, (2013).

97. Gui-xiang, L., and Wie-min, G., "Research of campus network security sytem based on intrusion detection.", *Computer Design and Applications (ICCDA), 2010 International Conference on*, 2096 - 2100, IEEE, (2010).

98. Karataş, G., "Derin öğrenme tabanlı saldırı tespit sistemi", (2020).

99. Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., and Chen, T. "Recent advances in convolutional neural networks.", *Pattern recognition*, 77, 354-377, (2018).

100. Albawi, S., Mohammed, T. A., and Al-Zawi, S., "Understanding of a convolutional neural network.", *In 2017 international conference on engineering and technology (ICET)*, 1-6, IEEE, (2017).

101. Medsker, L. R., and Jain, L. C., "Recurrent neural networks Design and Applications.", *CRC Press*, Australia, 5, 64-67, (2001).

102. Graves, A., and Graves, A., "Long short-term memory. Supervised sequence labelling with recurrent neural networks", 37-45, (2012).

103. Zhou, C., and Paffenroth, R. C., "Anomaly detection with robust deep

autoencoders.", *In Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*, 665-674, (2017).

104. Omari, M. A., Rawashdeh, M., Qutaishat, F., Alshira'H, M., and Ababneh, N., "An intelligent tree-based intrusion detection model for cyber security", *Journal of Network and Systems Management*, (2021).

105. Moualla, S., Khorzom, K., and Jafar, A., "Improving the performance of machine learning-based network intrusion detection systems on the unsw- nb15 dataset", *Computational Intelligence and Neuroscience*, (2021).

106. Altunay H. C., and Albayrak Z., "Network intrusion detection approach based on convolutional neural network", *European Journal of science and Technology*, 22-29, (2021).

107. Popoola, S. I., Adebisi, B., Ande, R., Hammoudeh, M., Anoh, K., and Atayero, A. A., " An integrated rule based intrusion detection system: Analysis on unswnb15 data set and the real time online dataset", *Sensors*, 21, 1397-1418 (2021).

108. Park, N., and Ahn, H. K., " Multi-layer rnn based short-term photovoltaic power forecasting using iot dataset", *AEIT International Annual Conference (AEIT)*, IEEE, (2019).

109. Goodfellow, I., Bengio, Y., and Courville, A., "Deep Learning", (2016).

110. Zhang, A., Lipton, Z., Li, M., and Smola, A., "Dive into Deep Learning", (2021).

111. Nawi, N. M., Atomi, W. H., and Rehman, M. Z., " The effect of data pre-processing on optimized training of artificial neural networks", *Procedia Technology, Elsevier*, 32-39 (2013).

112. Altunay, H. C., and Albayrak, Z., "A hybrid CNN+LSTM-based intrusion detection system for industrial IoT networks.", *Engineering Science and Technology, an International Journal*, Elsevier, 38, 101322, (2023).

ÖZGEÇMİŞ

Hakan Can ALTUNAY, ilkokulu Samsun Denizevleri İlkokulu'nda, ortaokulu ise Samsun 23 Nisan İlköğretim Okulu'nda tamamladı. Samsun Atakum Anadolu Teknik ve Endüstri Meslek Lisesi, Elektronik bölümünden 2001 yılında mezun olduktan sonra, 2002 yılında başladığı Sakarya Üniversitesi Teknik Eğitim Fakültesi Elektronik-Bilgisayar Eğitimi bölümünden 2006 yılında mezun oldu. 2014 yılında Sakarya Üniversitesi Fen Bilimleri Enstitüsü Elektronik – Bilgisayar Eğitimi Ana Bilim dalında yüksek lisansını tamamladı. 2010 yılından bu yana Ondokuz Mayıs Üniversitesi Çarşamba Ticaret Borsası Meslek Yüksekokulu Bilgisayar teknolojileri Bölümü'nde Öğretim Görevlisi olarak çalışmaktadır.

YAYINLAR

1. Altunay, H. C., & Albayrak, Z. (2023). A hybrid CNN+LSTM-based intrusion detection system for industrial IoT networks. *Engineering Science and Technology, an International Journal*, 38(101322).

2. Uluer, A. F., Albayrak, Z., Özalp, A. N., Çakmak, M., & Altunay, H. C. (2022, May). BGP Anomali Tespitinde Hibrit Model Yaklaşımı. In 2022 *30th Signal Processing and Communications Applications Conference* (SIU) (pp. 1-4). IEEE.

3. Altunay, H. C., Albayrak, Z., Özalp, A. N., & Çakmak, M. (2021, June). Analysis of anomaly detection approaches performed through deep learning methods in SCADA systems. In 2021 *3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)* (pp. 1-6). IEEE.

4. Altunay, H. C., & Albayrak, Z., (2021). Network intrusion detection approach based on Convolutional Neural Network. In 2021 *3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA). European Journal of Science and Technology,* (pp. 22-29).