ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

DERİN ÖĞRENİMİ YÖNTEMLERİ İLE SİBER SALDIRILARININ TESPİTİ

Abdolreza GHAFFARLOU

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

ANKARA
2026

ÖZET

Yüksek Lisans Tezi

DERİN ÖĞRENİMİ YÖNTEMLERİ İLE SİBER SALDIRILARININ TESPİTİ

Abdolreza GHAFFARLOU

Ankara Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Prof. Dr. Recep ERYİĞİT

Bilgisayar ağlarının yaygınlaşmasıyla birlikte Dağıtık Hizmet Reddi (DDoS) saldırıları, bilgi sistemleri için ciddi bir tehdit haline gelmiştir. Bu saldırılar, hedef sistemlerin bant genişliğini ve kaynaklarını tüketerek hizmet veremez duruma gelmesine neden olmaktadır. Geleneksel imza tabanlı saldırı tespit sistemleri, yeni ve bilinmeyen saldırı türleri karşısında yetersiz kalabilmektedir. Bu nedenle, makine öğrenmesi ve derin öğrenme tabanlı yöntemler son yıllarda saldırı tespitinde ön plana çıkmıştır.

Bu tez çalışmasında, CIC-DDoS2019 veri seti kullanılarak SYN-Flood ve UDP-Flood DDoS saldırılarının tespiti için derin öğrenme tabanlı bir saldırı tespit sistemi geliştirilmiştir. Veri setinde bulunan ciddi sınıf dengesizliği problemi, Üretici Çekişmeli Ağlar (GAN) kullanılarak giderilmiştir. GAN ile yalnızca eğitim verisi üzerinde sentetik örnekler üretilmiş, test verisi orijinal dağılımı korunarak değerlendirme yapılmıştır.

Sınıflandırma modeli olarak Çok Katmanlı Algılayıcı (MLP) kullanılmıştır. GAN destekli veri artırma sonrası modelin doğruluk, F1-skor ve ROC-AUC değerlerinde belirgin artışlar gözlemlenmiştir. Elde edilen sonuçlar, GAN tabanlı veri artırmanın DDoS saldırı tespitinde özellikle azınlık sınıflar için model performansını önemli ölçüde iyileştirdiğini göstermektedir.

Ocak 2026, 90 sayfa

Anahtar Kelimeler: Saldırı Tespit Sistemleri (IDS), Derin Öğrenme (DL), Çok Katmanlı Algılayıcı (MLP), ADAM algoritması, Temel Bileşen Analizi (PCA), CICIDS 2019 veri

ABSTRACT

Master's Thesis

DETECTION OF CYBER ATTACKS USING DEEP LEARNING METHODS

Abdolreza GHAFFARLOU

Ankara University
Graduate School of Natural and Applied Science
Department of Computer Engineering

Supervisor: Prof. Dr. Recep ERYİĞİT

With the widespread use of computer networks, Distributed Denial of Service (DDoS) attacks have become one of the most serious threats to information systems. These attacks aim to exhaust the bandwidth and resources of target systems, making services unavailable to legitimate users. Traditional signature-based intrusion detection systems often fail to detect novel and complex attack patterns, which has increased interest in machine learning and deep learning-based approaches.

In this thesis, a deep learning-based intrusion detection system is proposed for detecting SYN- Flood and UDP-Flood DDoS attacks using the CIC-DDoS2019 dataset. A significant class imbalance problem in the dataset is addressed using Generative Adversarial Networks (GAN) for synthetic data generation. Synthetic samples are generated only on the training dataset, while the test dataset remains unchanged to ensure fair performance evaluation.

A Multi-Layer Perceptron (MLP) classifier is employed for attack detection. Experimental results demonstrate that GAN-based data augmentation significantly improves classification performance, particularly for minority classes, in terms of accuracy, F1-score, and ROC-AUC metrics. The findings confirm the effectiveness of GAN-supported data augmentation in DDoS attack detection tasks.

January 2026, 90 page

Keywords: Intrusion Detection System (IDS), Deep Learning (DL), Multi-Layer Perceptron (MLP), ADAM Algorithm, Principal Component Analysis (PCA), CICIDS 2019 datase

TEŞEKKÜR

Bu tez çalışmasının her aşamasında bilgi ve deneyimiyle bana rehberlik eden, yol gösterici tutumu ve teşvik edici yaklaşımıyla katkı sağlayan değerli danışmanım Prof. Dr. Recep ERYİĞİT'e içten teşekkürlerimi sunarım.

Ayrıca, araştırma süresince verdikleri destekle her daim yanımda olduklarını hissettiren aileme; bu süreçte moral ve motivasyon sağlayarak yanımda olan dostlarıma ve çalışma arkadaşlarıma teşekkür ederim. Varlıklarıyla sürece güç katan tüm sevdiklerime gönülden minnettarım.

Abdolreza GHAFFAARLOU
Ankara, Ocak 2026

İÇİNDEKİLER

KISALTMALAR DİZİNİ

IDS	Intrusion Detection System (Saldırı Tespit Sistemi)
DL	Deep Learning (Derin Öğrenme)
MLP	Multi-Layer Perceptron (Çok Katmanlı Algılayıcı)
GAN	Generative Adversarial Network (Üretici Çekişmeli Ağ)
PCA	Principal Component Analysis (Temel Bileşen Analizi)
DDoS	Distributed Denial of Service (Dağıtık Hizmet Reddi)
UDP	User Datagram Protocol
TCP	Transmission Control Protocol
SYN	Synchronize (Senkrone - TCP bağlantı başlatma bayrağı)
ACK	Acknowledgment (Onay)
CNN	Convolutional Neural Network (Evrimsel Sinir Ağı)
RNN	Recurrent Neural Network (Yinelemeli Sinir Ağı)
LSTM	Long Short-Term Memory
ROC	Receiver Operating Characteristic
AUC	Area Under the Curve
F1	Harmonik ortalama
SVM	Support Vector Machine (Destek Vektör Makineleri)
KNN	K-Nearest Neighbors (K-En Yakın Komşu)
HTTP/HTTPS	HyperText Transfer Protocol / Secure
ICMP	Internet Control Message Protocol
DNS	Domain Name System
SMTP	Simple Mail Transfer Protocol
FTP	File Transfer Protocol

ŞEKİLLER DİZİNİ

ÇİZELGELER DİZİNİ

1. GİRİŞ

Günümüzde bilgi ve iletişim teknolojilerinde yaşanan hızlı gelişmeler, internetin ve bilgisayar ağlarının hemen her alanda yoğun biçimde kullanılmasına yol açmıştır. Kamu kurumları, finans kuruluşları, sağlık sistemleri, eğitim altyapıları ve bireysel kullanıcılar, günlük işlemlerini büyük ölçüde ağ tabanlı sistemler üzerinden gerçekleştirmektedir. Bu durum, bilgiye erişimi kolaylaştırırken aynı zamanda bilgi sistemlerini hedef alan siber saldırıların da artmasına neden olmuştur. Siber saldırılar, yalnızca maddi kayıplara yol açmakla kalmayıp, hizmet sürekliliğini bozarak toplumsal ve ekonomik açıdan ciddi sonuçlar doğurabilmektedir. Siber saldırı türleri arasında Dağıtık Hizmet Reddi (Distributed Denial of Service – DDoS) saldırıları, en yaygın ve en yıkıcı saldırı türlerinden biri olarak öne çıkmaktadır. DDoS saldırıları, çok sayıda ele geçirilmiş sistem üzerinden hedef sunucuya eş zamanlı olarak yoğun trafik gönderilmesi esasına dayanmaktadır. Bu yoğun trafik sonucunda hedef sistemin bant genişliği, işlem gücü veya bellek kaynakları tükenmekte ve sistem meşru kullanıcılara hizmet veremez hâle gelmektedir. Özellikle SYN-Flood ve UDP-Flood gibi saldırı türleri, ağ katmanında gerçekleştirilmesi ve kısa sürede yüksek hacimli trafik üretebilmesi nedeniyle ciddi bir tehdit oluşturmaktadır.

Geleneksel saldırı tespit sistemleri çoğunlukla imza tabanlı veya kural tabanlı yaklaşımlar kullanmaktadır. Bu sistemler, önceden tanımlanmış saldırı imzalarına dayanarak çalıştıkları için bilinen saldırıları tespit etmede başarılı olsalar da, yeni veya değişken yapıdaki saldırılar karşısında yetersiz kalabilmektedir. Ayrıca, saldırı tekniklerinin sürekli evrim geçirmesi, imza tabanlı sistemlerin güncel kalmasını zorlaştırmakta ve yüksek yanlış negatif oranlarına yol açmaktadır. Bu nedenle, ağ trafiği içerisindeki anormal davranışları öğrenebilen ve genelleyebilen daha esnek yöntemlere ihtiyaç duyulmaktadır.

Bu noktada makine öğrenmesi (Machine Learning – ML) ve derin öğrenme (Deep Learning – DL) tabanlı yaklaşımlar, saldırı tespit sistemlerinde önemli bir alternatif olarak ortaya çıkmıştır. Derin öğrenme modelleri, büyük ölçekli veri setleri üzerinde karmaşık ve doğrusal olmayan ilişkileri öğrenebilme yetenekleri sayesinde ağ trafiği verilerindeki saldırı örüntülerini yüksek doğrulukla tespit edebilmektedir. Literatürde, Çok Katmanlı

Algılayıcı (MLP), Evrişimsel Sinir Ağları (CNN), Uzun Kısa Süreli Bellek (LSTM) ve hibrit derin öğrenme mimarilerinin DDoS saldırılarının tespitinde etkili sonuçlar verdiği çok sayıda çalışma ile ortaya konmuştur.

Bununla birlikte, gerçek dünya ağ trafiği verileri kullanılarak yapılan çalışmalarda karşılaşılan en önemli problemlerden biri sınıf dengesizliği (class imbalance) problemidir. DDoS veri setlerinde genellikle normal (benign) trafik veya belirli saldırı türleri büyük çoğunluğu oluştururken, bazı saldırı türleri oldukça az sayıda örnek içermektedir. Bu durum, derin öğrenme modellerinin çoğunluk sınıflara yönelmesine ve azınlık sınıfların yeterince öğrenilememesine neden olmaktadır. Sonuç olarak, yüksek genel doğruluk değerlerine rağmen azınlık sınıflar için düşük tespit oranları elde edilebilmektedir.

Sınıf dengesizliği problemini gidermek amacıyla literatürde çeşitli yeniden örnekleme (oversampling/undersampling) ve maliyet duyarlı öğrenme yöntemleri önerilmiştir. Son yıllarda ise Üretici Çekişmeli Ağlar (Generative Adversarial Networks – GAN), veri artırma alanında dikkat çeken bir yaklaşım hâline gelmiştir. GAN modelleri, gerçek verinin dağılımını öğrenerek bu veriye benzer sentetik örnekler üretebilmekte ve böylece azınlık sınıfların temsil gücünü artırabilmektedir. DDoS saldırı tespitine yönelik yapılan güncel çalışmalar, GAN tabanlı veri artırma yöntemlerinin derin öğrenme modellerinin özellikle azınlık sınıflardaki performansını anlamlı ölçüde iyileştirdiğini göstermektedir.

Bu tez çalışmasında, CIC-DDoS2019 veri seti kullanılarak SYN-Flood ve UDP-Flood DDoS saldırılarının tespiti için GAN destekli derin öğrenme tabanlı bir saldırı tespit sistemi önerilmektedir. Çalışmada, ağ trafiğine ait akış (flow) tabanlı özellikler kullanılmış ve veri setindeki sınıf dengesizliği problemi GAN ile giderilmiştir. GAN modeli yalnızca eğitim verisi üzerinde kullanılarak azınlık sınıflar için sentetik örnekler üretilmiş, test verisi ise orijinal dağılımı korunarak model performansı değerlendirilmiştir. Dengelenmiş veri seti üzerinde Çok Katmanlı Algılayıcı (MLP) modeli eğitilmiş ve elde edilen sonuçlar GAN kullanılmayan durum ile karşılaştırmalı olarak analiz edilmiştir.

Bu çalışmanın temel amacı, dengesiz ağ trafiği veri setleri üzerinde DDoS saldırı tespit başarısını artıran, güvenilir ve genellenebilir bir derin öğrenme yaklaşımı geliştirmek ve literatüre katkı sağlamaktır. Elde edilen bulguların, gerçek zamanlı saldırı tespit sistemlerinin geliştirilmesine ve ağ güvenliği alanındaki çalışmalara yol gösterici nitelikte olması hedeflenmektedir.

1.1 Problem Tanımı

Siber güvenlik alanında, Dağıtık Hizmet Engelleme (Distributed Denial of Service – DDoS) saldırılarının tespiti, saldırganların ağ kaynaklarını aşırı yüklemek amacıyla sürekli olarak yeni ve daha karmaşık teknikler geliştirmesi nedeniyle günümüzde hâlen önemli bir araştırma problemi olarak varlığını sürdürmektedir. Özellikle TCP SYN Flood ve UDP Flood gibi yaygın DDoS saldırı türleri, hedef sistemlerin hizmet veremez hâle gelmesine neden olarak ciddi operasyonel ve ekonomik kayıplara yol açmaktadır. Bu saldırı türlerinin ağ trafiği içerisindeki davranışlarının normal trafikle benzerlik göstermesi, tespit süreçlerini daha da zorlaştırmaktadır.

Bu tez çalışmasında kullanılan CIC-DDoS2019 veri seti, gerçekçi ağ ortamlarında oluşturulmuş ve farklı DDoS saldırı senaryolarını içeren güncel bir veri setidir. Veri seti, özellikle TCP SYN Flood ve UDP Flood saldırılarına ait ayrıntılı ağ trafiği özellikleri sunarak, bu saldırı türlerinin analiz edilmesi ve tespit edilmesi için önemli bir kaynak oluşturmaktadır. CIC-DDoS2019 veri setinin sunduğu yüksek boyutlu ve kapsamlı özellik kümesi, derin öğrenme ve makine öğrenmesi tabanlı saldırı tespit sistemlerinin geliştirilmesine olanak tanımaktadır.

Ancak, veri setinde yer alan özelliklerin fazlalığı, özellikle MLP gibi denetimli derin öğrenme modelleri kullanıldığında çeşitli zorlukları da beraberinde getirmektedir. Tüm özelliklerin modele dâhil edilmesi, aşırı öğrenmeye (overfitting), hesaplama maliyetinin artmasına ve modelin genelleme yeteneğinin azalmasına yol açabilmektedir. Bu durum, TCP SYN Flood ve UDP Flood saldırılarının doğru ve etkili bir şekilde tespit edilmesini olumsuz etkileyebilmektedir. Bu nedenle, DDoS saldırılarının tespitine en fazla katkı sağlayan ayırt edici özelliklerin belirlenmesi büyük önem taşımaktadır.

Literatürde, CIC-DDoS2019 veri seti kullanılarak DDoS saldırılarının tespitine yönelik çeşitli çalışmalar bulunsa da, TCP SYN Flood ve UDP Flood saldırıları özelinde, farklı özellik seçimi yöntemlerinin derin öğrenme ve makine öğrenmesi modelleri üzerindeki etkilerini kapsamlı ve karşılaştırmalı biçimde inceleyen çalışmaların sınırlı olduğu görülmektedir. Ayrıca, sınıf dengesizliği problemi, bu saldırı türlerinin tespitinde model performansını olumsuz etkileyen temel faktörlerden biri olarak öne çıkmaktadır.

Bu tez çalışması, söz konusu eksiklikleri gidermeyi amaçlamakta ve CIC-DDoS2019 veri seti üzerinde TCP SYN Flood ve UDP Flood saldırılarına odaklanarak, özellik seçimi stratejilerinin optimize edilmesini hedeflemektedir. Bu kapsamda, seçilen özellikler kullanılarak MLP tabanlı derin öğrenme modelleri eğitilmekte ve model performansı ayrıntılı olarak değerlendirilmektedir. Elde edilen sonuçların, TCP SYN Flood ve UDP Flood saldırılarını daha doğru, etkili ve uyarlanabilir biçimde tespit edebilen saldırı tespit sistemlerinin geliştirilmesine katkı sağlaması amaçlanmaktadır.

Bu çalışma, güncel ve gerçekçi bir veri seti olan CIC-DDoS2019'un kullanılması ve TCP SYN Flood ile UDP Flood saldırılarının özel olarak ele alınması sayesinde, derin öğrenme tabanlı DDoS saldırı tespit sistemlerinin pratik siber güvenlik uygulamalarındaki etkinliğine yönelik önemli bulgular sunmayı hedeflemektedir. Ayrıca, elde edilen sonuçların daha ölçeklenebilir, dayanıklı ve güncel DDoS saldırılarına karşı dirençli savunma mekanizmalarının geliştirilmesine zemin hazırlaması beklenmektedir.

1.2 Çalışmanın Katkıları

Bu tez çalışması, özellikle Dağıtık Hizmet Engelleme (Distributed Denial of Service – DDoS) saldırılarının tespiti alanında olmak üzere, siber güvenlik alanına önemli katkılar sunmaktadır. Öncelikle, saldırganların giderek daha karmaşık ve çeşitli yöntemler kullanmasına rağmen, DDoS saldırılarının doğru bir şekilde tespit edilmesini ve sınıflandırılmasını sağlayan, gelişmiş makine öğrenmesi ve derin öğrenme algoritmalarına dayalı yenilikçi bir tespit yaklaşımı sunulmaktadır. Bu çalışma, farklı ağ ortamlarında DDoS saldırılarının tespit edilme hızını ve doğruluğunu artırmayı hedeflemekte olup, elde edilen doğruluk oranları mevcut yaklaşımlara kıyasla önemli bir gelişme göstermektedir.

Araştırma kapsamında, farklı özellik seçimi yöntemleri sistematik bir biçimde analiz edilmiş ve trafik akış örüntüleri ile paket anomali davranışları gibi DDoS saldırılarını en iyi temsil eden temel ağ trafiği özellikleri belirlenmiştir. Bu analizler, saldırı tespit modellerinin performansını doğrudan etkileyen ayırt edici özelliklerin ortaya konulmasını sağlamaktadır.

Bunun yanı sıra, bu çalışma, özellik seçiminin saldırı tespit doğruluğu üzerindeki etkisini kapsamlı biçimde inceleyerek, DDoS faaliyetlerinin tespitinde en önemli belirteçlere ilişkin değerli çıkarımlar sunmaktadır. Tezde ayrıca, kullanılan yöntemlerin gerçek dünya koşullarındaki uygulanabilirliğini ve ölçeklenebilirliğini ortaya koymak amacıyla gerçekçi veri setleri üzerinde kapsamlı deneysel analizler gerçekleştirilmiştir.

Sonuç olarak, bu tez çalışması yalnızca DDoS saldırı tespitinde mevcut yöntemlerin ötesine geçmekle kalmamakta, aynı zamanda kritik ağ altyapılarının korunmasına yönelik gelecekte geliştirilecek güvenlik çözümleri için sağlam bir temel oluşturmaktadır.

1.3 Çalışmanın Önemi

Bu tez çalışması, CIC-DDoS2019 veri seti kullanılarak TCP SYN Flood ve UDP Flood saldırılarının tespit edilmesini ele alması ve bu amaçla makine öğrenmesi ve derin öğrenme tekniklerini yenilikçi bir biçimde kullanması açısından önemli bir yere sahiptir. Günümüzde DDoS saldırılarının giderek daha karmaşık ve dinamik bir yapı kazanması, bu saldırıların hızlı ve doğru şekilde tespit edilmesini zorlaştırmakta ve mevcut güvenlik çözümlerinin etkinliğini sınırlandırmaktadır. Bu çalışma, söz konusu zorlukların aşılmasına katkı sağlayacak yöntemler sunmayı hedeflemektedir.

Çalışmada, CIC-DDoS2019 veri setinde yer alan yüksek boyutlu ağ trafiği özellikleri üzerinde özellik seçimi yöntemlerinin değerlendirilmesi ve optimize edilmesi, hem hesaplama maliyetlerinin azaltılması hem de saldırı tespit doğruluğunun artırılması açısından kritik bir rol oynamaktadır. Özellikle TCP SYN Flood ve UDP Flood saldırılarına özgü trafik akış örüntüleri, paket istatistikleri ve anomali davranışlarının

belirlenmesi, derin öğrenme tabanlı modellerin daha etkin ve genellenebilir sonuçlar üretmesini sağlamaktadır.

Bu tezde önerilen yaklaşım, yalnızca DDoS saldırı tespit sistemlerinin doğruluğunu artırmakla kalmamakta, aynı zamanda bu sistemlerin ölçeklenebilirliğini, verimliliğini ve gerçek zamanlı uygulamalara uygunluğunu da geliştirmektedir. Elde edilen deneysel sonuçlar, önerilen yöntemin farklı ağ ortamlarında uygulanabilir olduğunu ve gerçek dünya senaryolarında karşılaşılan saldırılara karşı etkili bir savunma mekanizması sunduğunu ortaya koymaktadır.

Sonuç olarak, bu çalışma; TCP SYN Flood ve UDP Flood saldırılarının tespitine odaklanan, özellik seçimi destekli derin öğrenme yaklaşımlarını temel alan yapısıyla, hem akademik literatüre hem de pratik siber güvenlik uygulamalarına önemli katkılar sunmaktadır. Elde edilen bulguların, kritik ağ altyapılarının korunmasına yönelik daha dayanıklı, güvenilir ve erişilebilir siber güvenlik çözümlerinin geliştirilmesine katkı sağlaması beklenmektedir.

2. LİTERATÜR ÇERÇEVE

2.1 Siber Saldırı Nedir?

Bilgisayar sistemleri ve bilgisayar ağlarıyla ilişkili kavram ve süreçleri tanımlamak amacıyla kullanılan siber kavramı, literatürde ilk kez 1990'lı yıllarda ortaya çıkmıştır. Bu kavram başlangıçta, bilgi teknolojileri sistemlerinin ağ bağlantıları üzerinden maruz kaldığı güvenlik risklerini ve tehditleri ifade etmek amacıyla kullanılmıştır. Zamanla dijitalleşmenin hız kazanmasıyla birlikte siber kavramı, bilgi sistemlerinin tümünü kapsayan daha geniş bir çerçevede ele alınmaya başlanmıştır.

Günümüzde hızla gelişen teknolojinin en önemli tartışma alanlarından biri siber güvenliktir. Siber güvenlik; yetkisiz erişim, veri ihlali, saldırı ve sistemlere zarar verme gibi tehditlere karşı bilgi teknolojileri altyapılarının, ağların, yazılımların ve verilerin korunmasını amaçlayan tüm teknik ve yönetsel faaliyetleri kapsamaktadır. Bu kapsamda bilginin gizliliğinin, bütünlüğünün ve erişilebilirliğinin sağlanması, siber güvenliğin temel hedefleri arasında yer almaktadır (Hekim ve Başıbüyük, 2013).

Siber güvenlik çoğunlukla yalnızca teknolojik bir alan olarak algılansa da, gerçekte sosyal, ekonomik ve politik boyutları da içeren çok yönlü bir yapıya sahiptir. Bilginin üretilmesi, işlenmesi, iletilmesi ve depolanması süreçlerinin her aşamasında ortaya çıkabilecek güvenlik açıkları, bireyler ve kurumlar açısından ciddi riskler barındırmaktadır. Bu durum, siber güvenliğin yalnızca teknik bir sorun değil, aynı zamanda toplumsal ve kurumsal bir mesele olduğunu göstermektedir (Özdemirci ve Torunlar, 2018).

Teknolojik gelişmelerle paralel olarak siber güvenlik alanında yürütülen çalışmalar ve geliştirilen politikalar da küresel ölçekte artış göstermiştir. Siber güvenlik politikaları yalnızca teknik önlemlerle sınırlı kalmamış; hukuki düzenlemeler, kurumsal yapılanmalar ve eğitim faaliyetlerini de kapsayacak şekilde genişletilmiştir (T.C. Ulaştırma, Denizcilik ve Haberleşme Bakanlığı, 2016–2019). Siber saldırıların geniş bir etki alanına sahip

olması, bu alanda görev yapan uzmanların güncel bilgi ve becerilerini sürekli olarak geliştirmelerini zorunlu hale getirmiştir (Özdemirci ve Torunlar, 2018).

Siber güvenliğin bir diğer önemli boyutu ise uluslararası niteliğidir. Siber saldırıların sınır tanımayan yapısı, devletlerin ulusal siber güvenlik stratejileri geliştirmesini gerekli kılmakta ve bu stratejilerin uluslararası iş birliği ile desteklenmesini zorunlu hale getirmektedir. Bu kapsamda ülkeler arasında yapılan anlaşmalar ve uluslararası kuruluşlar aracılığıyla oluşturulan ortak protokoller, siber güvenlik politikalarının etkinliğini artırmaktadır (Clarke ve Knake, 2011). Bu çerçevede siber güvenlik eğitimleri ve önleyici uygulamalar, küresel ölçekte büyük önem taşımaktadır. Bireylerden kurumlara, ulusal yapılardan uluslararası ilişkilere kadar her düzeyde siber güvenlik önlemlerinin sürekli olarak güncellenmesi ve iyileştirilmesi gerekmektedir. Teknolojik, hukuki, sosyal ve ekonomik boyutlarıyla bütüncül bir şekilde ele alındığında, siber güvenliğin sürdürülebilir biçimde sağlanması mümkün olacaktır. Güvenli bir dijital ortamın oluşturulması ve teknolojik gelişmelerin sağlıklı bir şekilde devam ettirilebilmesi, etkin siber güvenlik uygulamalarının varlığına bağlıdır (Özdemirci ve Torunlar, 2018).

2.2 Siber Saldırı Tarihçesi

Kökleri 1970'li yıllara dayanmasına rağmen, siber güvenlik çalışmaları bilgisayar bilimleri alanında asıl ivmesini 1980'lerde yaşanan bilgisayar korsanlığı (hacking), kötü amaçlı yazılımlar, sisteme izinsiz girişler ve casusluk girişimleri sonucunda kazanmıştır. İnternetin 1990'lı yıllarda daha yaygın kullanılmasıyla birlikte siber saldırıların sıklığı artmaya başlamıştır. Bu dönemde yazılım ve ağ güvenliği literatüründeki artışın temel nedeni de bu saldırılardır. Ağ güvenliği, hem şirketler hem de hükümetler için en önemli önceliklerden biri haline gelmiştir. 2000'li yıllarda siber tehditlerin karmaşıklığının artmasıyla birlikte, siber güvenlik çalışmaları güvenlik araştırmalarının merkezine yerleşmiştir. Devletler ve bazı uluslararası kuruluşlar siber güvenlik stratejileri geliştirmeye başlamıştır (Tarhan, 2022).

Kritik altyapıların ve bilgisayar ağlarının güvenliği, en önemli sorunlardan biri olarak kabul edilmeye başlanmıştır. Geleneksel güvenlik politikalarından, bilgi çağında geliştirilmesi gereken modern güvenlik politikalarına doğru bir değişimin başladığı gözlemlenmiştir. 2007 yılındaki Estonya saldırılarının ardından, özellikle 2010'lu yıllarda siber güvenlik araştırmaları daha fazla ilgi görmüştür. Bu dönemde siber güvenlik, hayati tesislere yönelik saldırıların şiddeti ve bazı fiziksel saldırıların meydana gelmesi nedeniyle küresel bir endişe haline gelerek gelişmiştir. Siber güvenlik çalışmaları alanı, birçok disiplinin etkisiyle sürekli evrilmektedir (Tarhan, 2022). Siber güvenliğin önemli dönüm noktalarını şu şekilde özetlemek mümkündür:

1970'ler: Siber güvenlik araştırmaları bu dönemde şekillenmeye başlamış; öncelikle teknolojik zafiyetlere ve koruma stratejilerine odaklanılmıştır.

1980'ler: Bilgisayar korsanlığı ve siber saldırılardaki artış, siber güvenlik önlemlerinin gerekliliğine dikkat çekmiştir.

1990'lar: İnternetin yaygın kullanımı sonucunda siber güvenlik, özellikle Amerika Birleşik Devletleri'nde bir ulusal güvenlik sorunu olarak daha geniş çapta kabul görmüştür.

2000'ler: Artan siber tehditler ve ulusal altyapılara yönelik saldırılar sonucunda siber güvenlik, kritik bir araştırma alanı haline gelmiştir. 2007 Estonya saldırıları ve 2011'deki Stuxnet virüsü, siber güvenlik farkındalığını artıran önemli olaylar olmuştur.

11 Eylül Sonrası Dönem: Bilgi savaşı ve siber terörizm gibi kavramların yaygınlaşmasıyla siber güvenliğe olan ilgi önemli ölçüde artmıştır.

Güncel Eğilimler: Kritik altyapı ve ağların savunmasına odaklanan siber güvenlik çalışmaları; siyasi, sosyal ve ekonomik meselelerle etkileşim kuran çok disiplinli bir konu haline gelmiştir. Günümüz güvenlik sorunlarını etkin bir şekilde ele almak için bu alan, çeşitli disiplinler arasında iş birliğine her geçen gün daha fazla ihtiyaç duymaktadır (Tarhan, 2022).

2.3 Siber Saldırı Önemi

Günümüzde siber güvenlik, özellikle çoğu işletmenin ve resmi kurumun operasyonlardan veri depolamaya kadar her aşamada internete bağımlı olması nedeniyle devasa bir endüstri haline gelmiştir. İnternet, günlük hayatımızın ayrılmaz bir parçası olmuştur. Çoğu devlet kurumu ve büyük şirket, siber güvenlik uzmanlarını bünyesinde istihdam etmekte veya dışarıdan hizmet almaktadır. İnternetin büyümesiyle birlikte tehditlerin de artması, bu alanı zorunlu bir ihtiyaç haline getirmiştir.

Siber güvenlik; ağlara sızmaya ve yıkım yaratmaya çalışan odaklara karşı bireylerin, kuruluşların ve hükümetlerin savunulmasına şu yollarla yardımcı olur:

- Virüsler
- Oltalama (Phishing) saldırıları
- DDoS saldırıları

Bilgisayarların yanı sıra tablet ve akıllı telefon gibi elektronik cihazlar da bu saldırılara karşı savunmasızdır. Bu saldırılar, bireyleri kandırarak e-posta, iş, finans ve diğer özel alanları etkileyen giriş bilgilerini ifşa etmelerine neden olabilir. Sistemlere sızarak kimlik hırsızlığına yol açabilecek verileri ele geçirme potansiyeline sahiptirler.

2.4 Dağıtık Hizmet Engelleme Saldırılarısı (DDoS Saldırısı)

Dağıtık hizmet reddi (DDoS) olarak bilinen bir siber saldırı, birden fazla sistemin belirli bir sistemin bant genişliğini veya kaynaklarını aşırı şekilde tüketmesiyle gerçekleşir. Bu saldırılar genellikle bir veya daha fazla web sunucusunu hedef alır. Hizmet Reddi (DoS) saldırılarının temel amacı, belirli bir kaynağın ciddi şekilde yavaşlamasına ya da tamamen devre dışı kalmasına neden olmaktır. Bu durum, sistemdeki güvenlik açıklarının istismar edilmesiyle ortaya çıkarak işlem gücünün bozulmasına ya da sistem kaynaklarının tükenmesine yol açabilir. Birdiğer yaygın yöntem ise hedef sistemin ağının aşırı veriyle doldurulması (flooding) ve kontrol altına alınmasıdır. Bu sayede diğer kullanıcıların ağa erişimi tamamen engellenmiş olur. Bu tür saldırılar genellikle kötü amaçlı yazılımlar

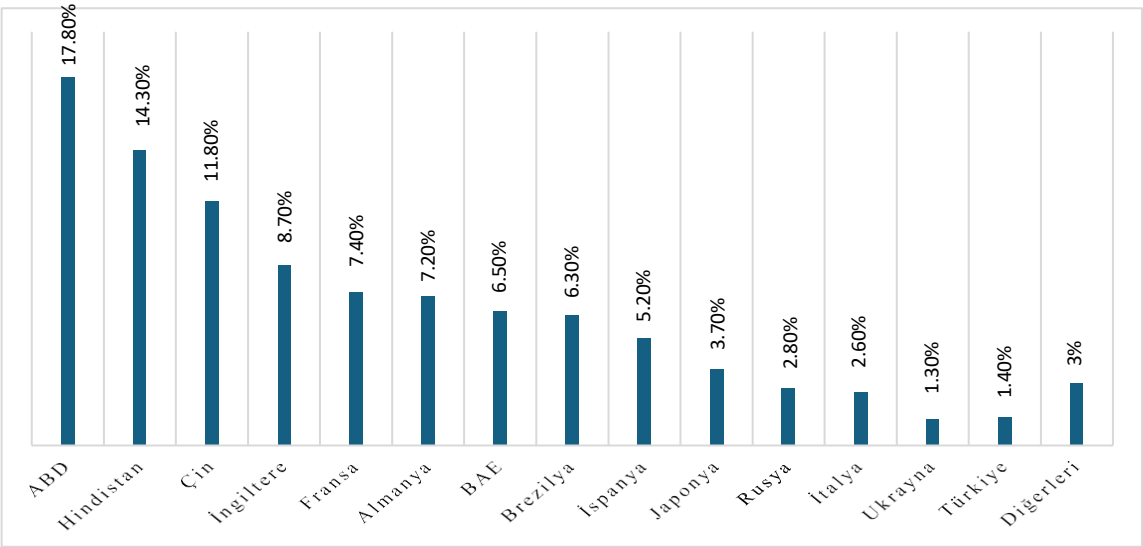
tarafından ele geçirilmiş binlerce cihazdan gelen, farklı makine veya IP adreslerinin kullanılmasıyla gerçekleştirilir. (Mirkovic & Reiher, 2004; Kumar et al.2020).

Şekil 2.1 Çeşitli sektörlerde DDOS saldırı

DDoS ve DoS saldırıları, dünya genelinde birçok sektörü ciddi şekilde etkileyebilir. Şekil 2.1, DDoS saldırılarının çeşitli sektörler üzerindeki geniş kapsamlı etkilerini göstermektedir. DDoS saldırıları, kişisel bilgisayarları aşırı işlem yüküne ve bellek tüketimine maruz bırakarak önemli ölçüde etkileyebilir. Bir kişisel bilgisayar çok sayıda ağ isteğiyle hedef alındığında, bu istekleri sürekli olarak işlemek ve yanıtlamak zorunda kalır. Bu durum, işlemci (CPU) kullanımının artmasına ve belleğin aşırı tüketilmesine neden olur. Sürekli işlem yoğunluğu, bilgisayarın performansını düşürerek cihazın yavaşlamasına veya yanıt veremez hâle gelmesine yol açar. Ayrıca, saldırılar ağ bant genişliğini de tüketerek, meşru kullanıcıların ağ trafiğine erişimini daha da zorlaştırır. Bu da kullanıcılar için gecikmeli yanıt süreleri, çevrimiçi hizmetlere erişim sorunları ve sistem çökmesi gibi sonuçlara neden olur (Grosse et al., 2017; Zargar et al., 2013)..Nesnelerin İnterneti (IoT) cihazları için DDoS saldırılarının etkisi daha da ağır olabilir. Bu saldırılar, IoT cihazlarının sınırlı depolama ve ağ kapasitesi gibi kısıtlı kaynaklarını hedef alarak ciddi hizmet kesintilerine yol açar. Aşırı trafik yüklemesiyle mevcut sistemlere büyük zararlar verebilir, hizmetlerin kullanılamaz hâle gelmesine neden olur (Antonakakis et al., 2017). Günümüzde DoS ve DDoS saldırıları, gerçek zamanlı hasta takibi ve veri iletimi için giderek daha fazla IoT teknolojilerine bağımlı hâle gelen e-sağlık sistemlerinin güvenilirliği ve etkinliği için büyük bir tehdit oluşturmaktadır. Bu saldırılar, sunucuları aşırı isteklerle meşgul ederek meşru kullanıcıların erişimini engelleyebilir. Örneğin, DDoS saldırıları, EKG sinyallerini işleyen sistemlerin kaynaklarını tüketerek, bu sinyallerin güvenilirliğini zayıflatabilir ve tıbbi teşhis ve uzaktan izleme süreçlerini etkileyebilir. Araştırmalar, DDoS saldırısı

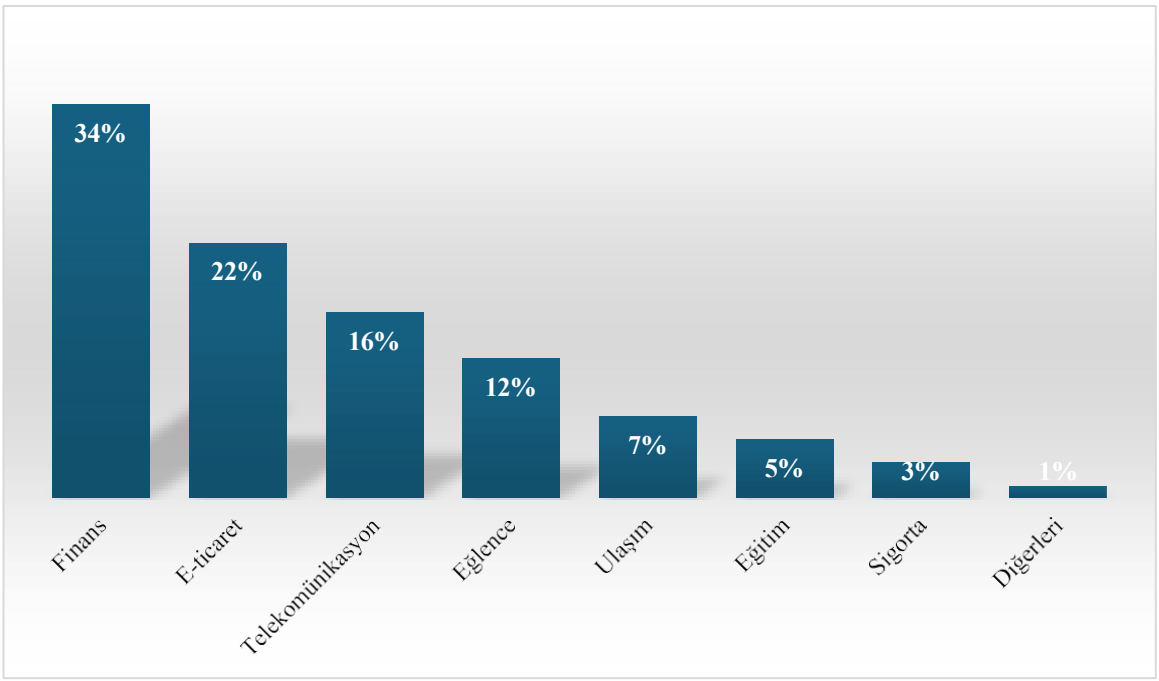
altında çalışan sistemlerin işlevlerini sürdürmesine rağmen, EKG sinyallerinin kaynak yetersizliği nedeniyle bozulduğunu ortaya koymuştur (Yin et al. 2017). Bu tür bozulmalar, hem hastaların anlık takibini tehlikeye atmakta hem de sağlık hizmetlerinin genel güvenilirliğini ve işlevselliğini ciddi şekilde zedelemektedir. Sonuç olarak, sağlık sisteminin zamanında ve güvenilir bakım sağlama kapasitesi önemli ölçüde zarar görmekte, bu da hasta güvenliği ve tıbbi operasyonların bütünlüğü açısından ciddi riskler oluşturmaktadır. İş dünyası ve endüstri sektörü de DDoS saldırılarının zararlı etkilerinden muaf değildir. Şirketler, çevrimiçi hizmetlerde (örneğin e-ticaret, müşteri desteği ve diğer temel işlevlerde) operasyonel kesintiler yaşayabilir. Innab ve Alamri, DDoS saldırılarının e- ticaret sistemleri üzerindeki etkilerini araştırmış ve bu saldırıların gizlilik, bütünlük ve erişilebilirlik gibi güvenlik ilkelerini nasıl tehdit ettiğini göstermiştir. Ubisoft ve E-bay gibi şirketlerin karşılaştığı örnekler, DDoS saldırılarının hizmet aksaklıklarına ve önemli kayıplara neden olduğunu ortaya koymaktadır (Innab & Alamri, 2021).Ek olarak, Mateen ve Shahzad farklı şirketlerin sunucu bilgisayarlarına yönelik DDoS saldırıları için siber risk tahmini amacıyla genel bir matematiksel model önermiştir. Bu model; parametreler, matematiksel denklemler ve Monte Carlo simülasyonları içermektedir ve iş dünyası ile siber sigorta şirketleri açısından oldukça anlamlı sonuçlar sunmuştur. Araştırma, çalışan sayısı az olan küçük işletmelerin başarılı bir DDoS saldırısı karşısında iflas etme olasılığının yüksek olduğunu; büyük şirketlerin ise yüksek saldırı oranlarında dahi önemli zararlar gördüğünü ortaya koymuştur (Mateen & Shahzad, 2020).

StormWall'un en son raporuna göre, 2023 yılında DDoS saldırıları, hem frekans hem de karmaşıklık açısından belirgin bir artış göstermiştir. Şekil 2.2 ve Şekil 2.3, 2023 yılının ilk çeyreğinde DDoS saldırılarının ülkelere ve endüstri sektörlerine göre dağılımını göstermektedir. Coğrafi olarak bakıldığında, Amerika Birleşik Devletleri, toplam saldırıların %17,6'sı ile en fazla saldırıya maruz kalan ülke olmuştur. Onu Hindistan (%14,2) ve Çin (%11,7) takip etmiştir. Saldırılardan önemli ölçüde etkilenen diğer ülkeler arasında Birleşik Krallık (%8,6), Fransa (%7,3), Almanya (%7,1), Birleşik Arap Emirlikleri (%6,4) ve Brezilya (%6,2) yer almaktadır. Ayrıca İspanya (%5,1), Japonya (%3,7), Rusya (%2,8), İtalya (%2,6), Ukrayna (%2,3) ve Türkiye (%1,4) de dikkate değer saldırı oranlarıyla karşı karşıya kalmıştır. Geri kalan %3'lük kısım ise diğer ülkelere dağılmıştır (StormWall, 2023).



Şekil 2.2 2023 yılında ülkelere göre StormWall'un DDoS saldırı raporu

Sektörel dağılıma bakıldığında, finans sektörü, toplam saldırıların %34'ü ile en çok hedef alınan alan olmuştur. Onu e-ticaret sektörü (%22) ve telekomünikasyon sektörü (%16) takip etmiştir. Diğer etkilenen sektörler arasında eğlence (%12), ulaşım (%7), eğitim (%5) ve sigorta (%3) yer alırken, kalan %1'lik saldırı oranı ise diğer sektörlere yönelmiştir (StormWall, 2023).

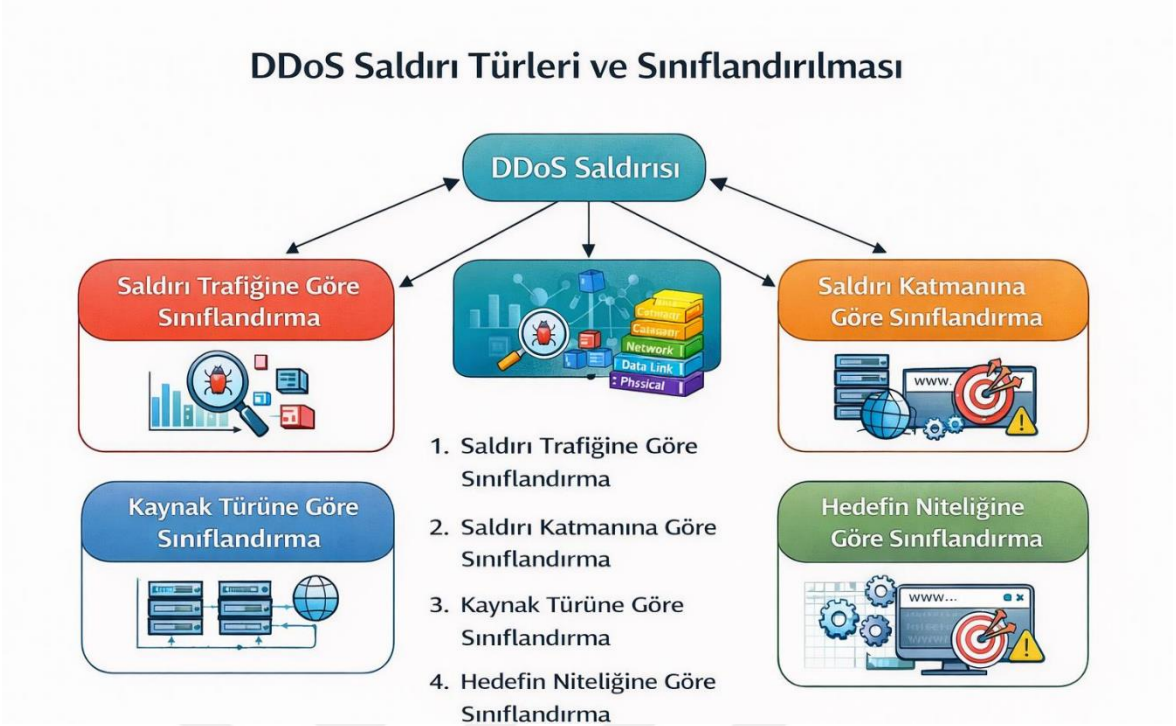


Şekil 2.3 2023 yılında sektörlere göre StormWall'un DDoS saldırı raporu

Özetle, DDoS saldırıları; kişisel bilgisayarlar, Nesnelerin İnterneti (IoT) cihazları, sağlık sektörü ve endüstri veya iş dünyası dahil olmak üzere birçok sektörü etkileyebilir. Bu saldırılar; hizmet kesintileri, gecikme artışı, sistem kaynaklarının tükenmesi, cihazların işlevsiz hâle gelmesi, ağ tıkanıklığı, güvenlik açıkları, hasta bakımında gecikmeler, veri erişim sorunları, operasyonel duraksamalar, mali kayıplar ve itibar zedelenmesi gibi ciddi sonuçlara yol açabilir (Zargar et al., 2013; Grosse et al., 2017). Bu etkilerin yaygınlığı, DDoS saldırılarına karşı güçlü güvenlik önlemlerinin ve etkili önleme stratejilerinin uygulanmasının önemini açıkça ortaya koymaktadır. DDoS saldırıları genel olarak üç ana kategoriye ayrılır: İlki, hedef sistemin bant genişliğini aşırı trafikle doldurarak çalışamaz hâle getirmeyi amaçlayan hacim tabanlı saldırılardır. ICMP flood, UDP flood, TCP flood ve DNS amplifikasyon saldırıları bu kategoriye örnektir. İkinci tür ise, ağ (katman 3) veya taşıma (katman 4) seviyesindeki protokol açıklarını sömürerek, hedefin işlem gücünü ya da güvenlik duvarı gibi hayati kaynaklarını tüketen protokol tabanlı saldırılardır. SYN flood, Ping flood ve Smurf saldırıları bu türün örnekleri arasında yer alır. Üçüncü tür olan uygulama katmanı saldırıları ise (katman 7), kurbana meşru görünümde bağlantılar kurarak sunucu kaynaklarını yoğun işlem talepleriyle aşırı yükler. HTTP flood saldırısı, bu türün tipik bir örneğidir (Peng et al., 2007; Zargar et al., 2013). Bu çalışmada, sadece protokol tabanlı TCP-SYN ve UDP hacim tabanli saldırısı verileri kullanılarak, derin öğrenme (DL) modelleri eğitilecektir.

2.5 DDoS Saldırı Türleri ve Sınıflandırılması

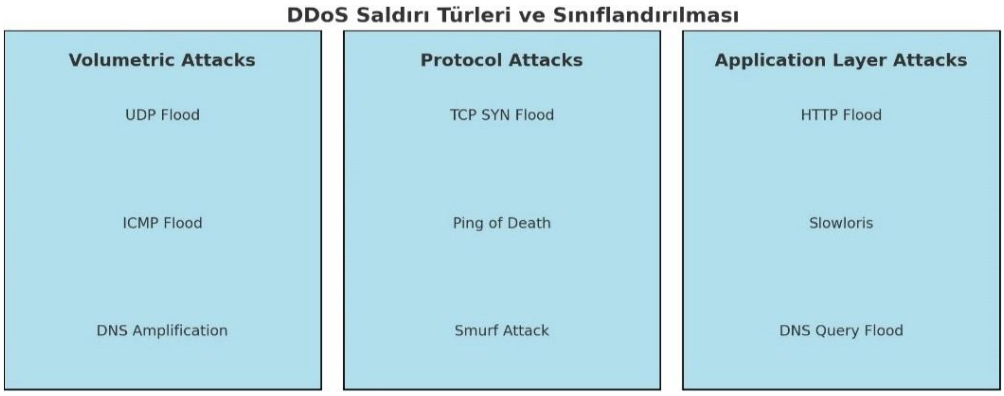
Dağıtılmış Hizmet Reddi (Distributed Denial of Service – DDoS) saldırıları, bilgi güvenliği alanında en kritik tehditlerden biri haline gelmiştir. Özellikle internet altyapısının yaygınlaşmasıyla birlikte DDoS saldırılarının sıklığı ve karmaşıklığı önemli ölçüde artmıştır. Bu saldırılar, hedef sistemin hizmet veremez hale gelmesini amaçlar ve genellikle çok sayıda zombi bilgisayar (botnet) aracılığıyla gerçekleştirilir (Mirkovic & Reiher, 2004). Literatürde DDoS saldırıları, genel olarak trafik türüne, hedeflenen katmana ve saldırı tekniğine göre sınıflandırılmaktadır (Şekil 2.4).



Şekil 2.4 DDoS saldırı türleri ve sınıflandırılması

2.5.1 Saldırı trafiğine göre sınıflandırma

DDoS saldırıları, gönderilen trafik türüne göre üç ana gruba ayrılmaktadır: Hacimsel saldırılar, protokol saldırıları ve uygulama katmanı saldırıları (Şekil 2.5).



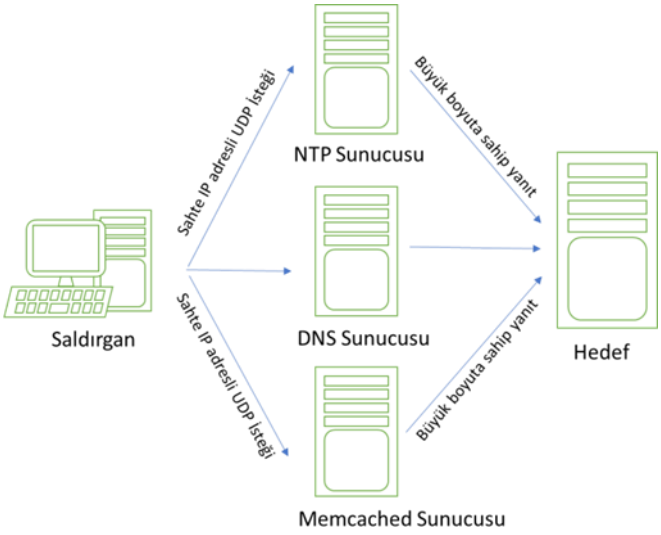
Şekil 2.5 DDoS saldırı trafik türleri ve sınıflandırılması

2.5.1.1 Hacimsel (Volumetric) saldırılar

Bu saldırı türü, hedefin bant genişliğini doldurarak hizmet veremez hale getirmeyi amaçlar. Genellikle **UDP flood**, ICMP flood ve DNS Amplification gibi teknikler kullanılır. Google (2020) tarafından açıklanan 2.54 Tbps'lik DDoS saldırısı, bu türün ne derece yıkıcı olabileceğini göstermiştir. Çoğu hacimsel saldırı, botnet ağları kullanılarak gerçekleştirilir ve saldırının şiddeti genellikle gigabit/saniye (Gbps) cinsinden ölçülür.

UDP flood saldırısı, yaygın bir hacim tabanlı DDoS (Dağıtılmış Hizmet Reddi) saldırı türüdür ve hedef sistemin bant genişliği ile işlem kapasitesini aşırı miktarda UDP (User Datagram Protocol) paketleriyle doldurarak çalışamaz hâle gelmesini amaçlar. UDP, bağlantısız bir protokol olduğundan, veri iletimi öncesinde herhangi bir el sıkışma (handshake) süreci gerektirmez. Bu özelliği, saldırganlar tarafından kötüye kullanılarak sistemin aşırı yüklenmesine yol açabilir.

UDP flood saldırısında, saldırgan hedef sunucunun rastgele veya belirli portlarına çok sayıda UDP paketi gönderir. Sunucu, her paketi alıp ilgili servisin açık olup olmadığını kontrol eder ve genellikle yanıt olarak ICMP Hedef Ulaşılamaz (Destination Unreachable) mesajı üretir. Bu işlem sürekli tekrarlandığında, sunucunun işlemci gücü ve ağ trafiği aşırı yoğunlaşır. Ayrıca, ICMP yanıtları ağ bant genişliğini de doldurarak ağ tıkanıklığına ve hizmet kesintilerine neden olur.UDP flood saldırıları özellikle güvenlik duvarı (firewall) ve derin paket incelemesi (DPI) olmayan sistemlerde etkili olabilir. Hedef sistem, meşru kullanıcı trafiğini ayırt edemeyecek hâle gelir ve meşru ağ hizmetleri kullanılamaz duruma düşer. Bu saldırı mekanizması, Kumar ve arkadaşları tarafından detaylı biçimde açıklanmıştır (Kumar et al., 2020) (Şekil 2.6).



Şekil 2.6 Volümetrik DDoS saldırı yapısı

2.5.1.2 Protokol saldırıları

OSI modelinin transport ve network katmanlarını hedef alır. En bilinen örneği **TCP-SYN Flood** saldırısıdır. Bu saldırıda, hedef sunucuya sahte SYN istekleri gönderilir ancak üç yönlü el sıkışma (3-way handshake) tamamlanmaz, bu da sunucunun kaynaklarının tükenmesine yol açar (Peng et al., 2007). Diğer yaygın protokol saldırıları arasında Ping of Death, Smurf ve TCP Reset yer alır.

TCP SYN saldırısı, aynı zamanda TCP SYN flood olarak da bilinen, dağıtılmış hizmet reddi (DDoS) türünde bir saldırıdır ve TCP bağlantısı kurma sürecini hedef alır. Bu tür bir saldırı, web sunucularının ve diğer ağ tabanlı sistemlerin normal işleyişini ciddi şekilde bozabilir. TCP SYN flood saldırıları, iki cihaz arasında bağlantı kurmak için gerekli olan Transmission Control Protocol (TCP)'nin üç yönlü el sıkışma (three-way handshake) sürecini istismar eder. Şekil 2.6, bu TCP üç yönlü el sıkışma sürecini açıklamaktadır. Bu işlem üç adımdan oluşur: istemci, sunucuya bir SYN (senkronizasyon) paketi gönderir; sunucu buna SYN-ACK (senkronizasyon-onay) paketi ile karşılık verir; son olarak istemci, bir ACK (onay) paketi göndererek bağlantıyı tamamlar. TCP SYN flood saldırısında ise saldırgan, hedef sunucuya çok sayıda SYN paketi gönderir ancak el sıkışma sürecini tamamlamak için gerekli olan ACK paketini göndermez. Bu durum, sunucuda birçok yarım açık bağlantı (half-open connection) oluşmasına neden olur ve

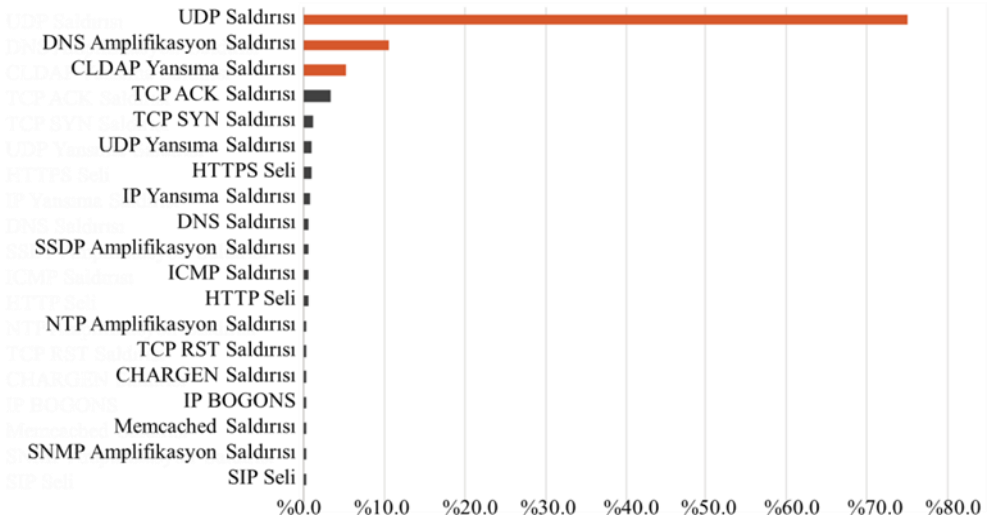
sunucu, hiçbir zaman gelmeyecek olan son ACK paketini beklemeye başlar. Bunun sonucunda, sunucunun bellek ve işlem gücü gibi kaynakları tükenir ve sistem meşru istemci taleplerine yanıt veremez hâle gelir. Bu saldırı mekanizması, Kumar ve arkadaşları tarafından açıklanmıştır (Kumar et al. 2020).

2.5.1.3 Uygulama katmanı saldırıları

Bu tür saldırılar, doğrudan uygulama katmanındaki zafiyetleri hedef alır ve genellikle HTTP, DNS veya SMTP protokolleri üzerinden yürütülür. Örneğin, **HTTP Flood** saldırısı, hedef web sunucusuna çok sayıda sahte HTTP isteği göndererek sunucunun yanıt verememesine neden olur. Bu tür saldırılar, tespit edilmesi en zor olanlardandır çünkü trafik genellikle meşru kullanıcı trafiği ile benzer özellikler taşır (Zargar et al. 2013).

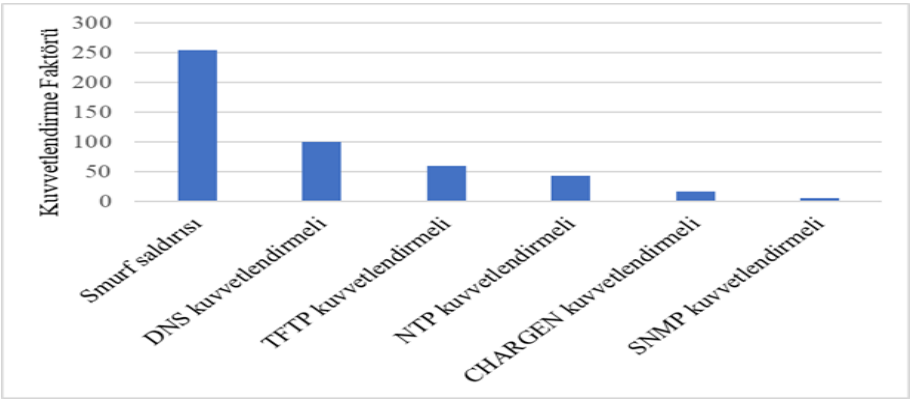
2.6 DDoS Saldırı Çeşitleri

DDoS saldırıları kapsamında en sık karşılaşılan yöntemler arasında flood saldırıları, yansıma saldırıları ve kuvvetlendirmeli saldırılar yer almaktadır. Flood saldırıları, hedef alınan kurban sisteme sürekli ve yoğun biçimde istek gönderilmesine dayanan simetrik saldırılar olarak tanımlanmaktadır. Bu tür saldırılarda çoğunlukla BotNet yapıları kullanılmakta olup SYN, UDP, İnternet Kontrol Mesajı Protokolü (ICMP) ve HTTP flood saldırıları başlıca örnekler arasında bulunmaktadır. Nexusguard'un 2020 yılı birinci çeyrek tehdit raporunda yer alan ve Şekil 2.7'de gösterilen verilere göre, UDP flood saldırıları tüm saldırıların %75'ini oluştururken; DNS kuvvetlendirmeli saldırılar %10,49, CLDAP yansıma saldırıları ise %5,27 oranında görülmektedir. Bu veriler, flood tipi saldırıların saldırganlar tarafından oldukça yaygın bir şekilde tercih edildiğini ortaya koymaktadır.



Şekil 2.7 2020 yılı 1.çeyrek dönemi için DDoS saldırı çeşitlerinin dağılımı

Amplifikasyon olarak adlandırılan kuvvetlendirme saldırıları, flood tipi saldırılardan farklı biçimde, düşük hacimli bir sorgu ile başlatılıp çok daha büyük boyutlu yanıtların üretilmesine yol açan asimetrik saldırılar olarak tanımlanmaktadır. Bu saldırı türünde, güvenlik açığı barındıran DNS, Önemsiz Dosya Aktarım Protokolü (TFTP), NTP gibi çeşitli sunucular, saldırı trafiğini büyütmek amacıyla araç olarak kullanılmaktadır. Şekil 2.8'de, DDoS kuvvetlendirme saldırılarında en sık tercih edilen sunucu türleri ve bu sunucuların saldırı trafiğini kaç kat artırabildiği gösterilmektedir.

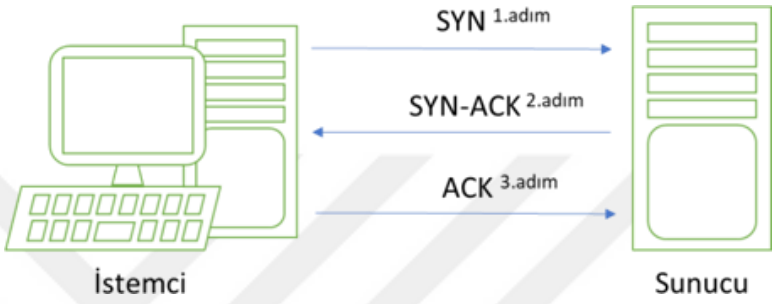


Şekil 2.8 Kullanılan kuvvetlendirme çeşitlerinin DDoS saldırılarının katlanmasına etkisi

Kuvvetlendirme saldırıları sunucuların özelliğine göre çok yüksek boyutlu saldırılara sebep olabileceği için en tehlikeli saldırı tipi olmaktadır. Sonuç olarak iki ana kategoriye ayrılan DDoS saldırı çeşitleri bu bölümde tanıtılacaktır.

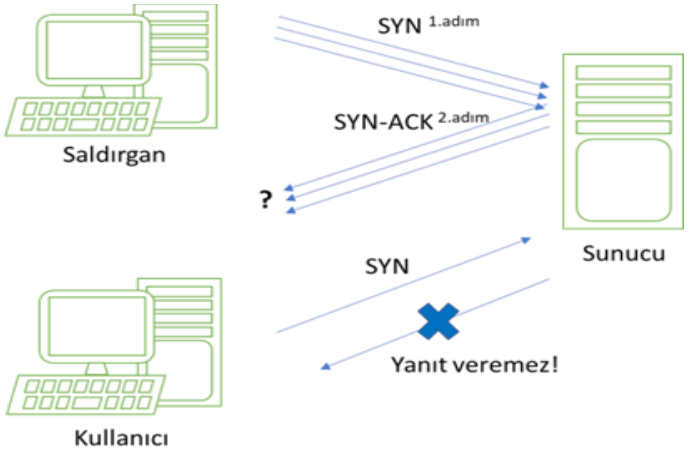
2.6.1 SYN flood saldırısı

SYN flood saldırısı, volümetrik bir saldırı türü olup TCP bağlantı sürecinde "üç yönlü el sıkışma" olarak adlandırılan mekanizmadaki zayıflıktan faydalanmaktadır. Şekil 2.9'da gösterilen üç yönlü el sıkışma sürecinde, bir sunucu ile TCP bağlantısının başlatılabilmesi için öncelikle istemci tarafından bir SYN paketi gönderilmektedir.



Şekil 2.9 TCP bağlantısının kurulması

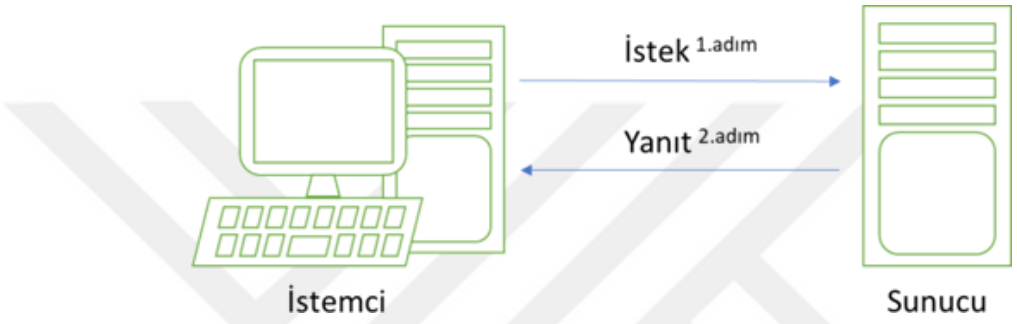
İstemciden gelen SYN isteğine, sunucu tarafından SYN-ACK yanıtı verilmesi gerekir ve ardından istemciden gelen bir ACK yanıtı ile bağlantı onaylanır. Şekil 2.10'da gösterilen SYN flood saldırısında, istekte bulunan istemcinin birden çok SYN isteği göndermesi ya da sahte bir IP adresinden SYN istekleri göndermesi sonucu SYN-ACK yanıtını veren sunucu isteklerin her biri için ACK yanıtı gelene kadar beklemeye devam eder. Yeni bağlantı kurulamayacak şekilde kaynakları tükenen sunucunun durumu hizmet reddi ile sonuçlanır.



Şekil 2.10 SYN Flood DDoS saldırısı

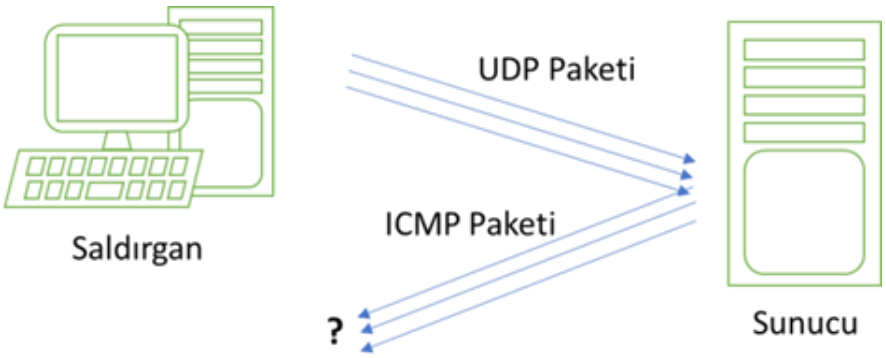
2.6.2 UDP flood saldırısı

Bir UDP flood, hedefteki bir sunucuyu UDP paketleri ile dolduran bir DDoS saldırısıdır. Şekil 2.11'de gösterilen UDP bağlantısında TCP'den farklı olarak üçlü el sıkışma gibi bir uçtan uca doğrulama mekanizması olmayıp istemcinin isteklerine cevap verilmesiyle veri iletişimi sağlanmaktadır. Bu nedenle UDP bağlantısında IP adresi doğrulama seçenekleri çok sınırlıdır.



Şekil 2.11 UDP bağlantısının kurulması

UDP flood saldırılarında sunucunun kaynaklarını tüketip çökertmek için çok sayıda kaynak IP kullanılarak sahte UDP paketleri hedef sunucuya gönderilir. Ayrıca, bu saldırının daha büyük etkiye sahip olması için hedef sunucunun port ve IP adresi sahte UDP paketlere dahil edilerek rastgele sunuculara veya ağ içindeki belirli bir sunucuya hedeflenebilir. Şekil 2.12'de gösterildiği gibi bu durum, hedef sunucunun söz konusu portu dinleyen uygulamayı tekrar tekrar kontrol etmesine ve uygulama bulunmadığında bir ICMP 'Hedefe Ulaşılamıyor' paketiyle yanıt vermesine neden olur.

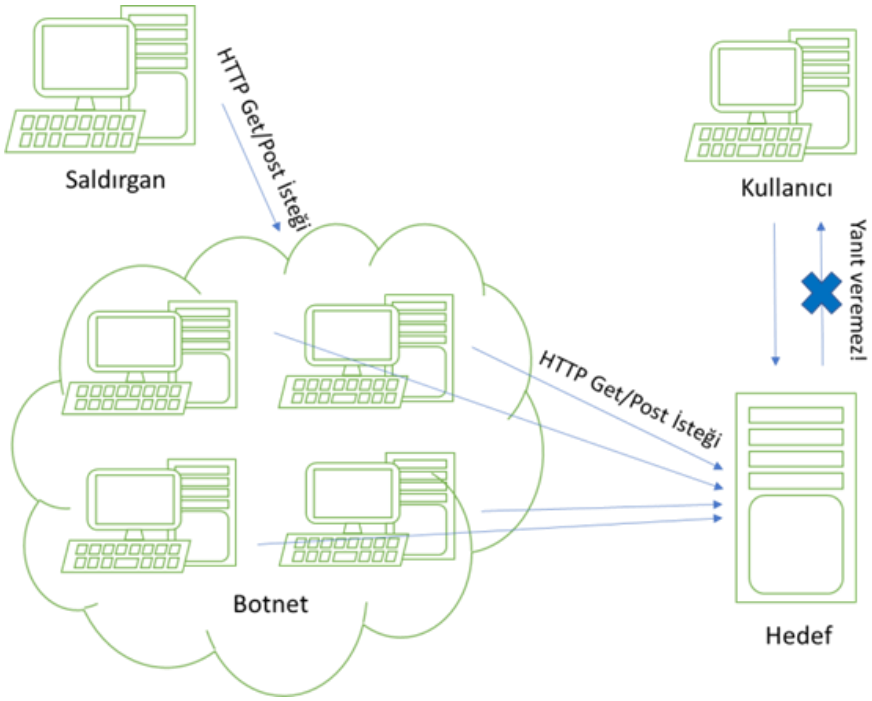


Şekil 2.12 UDP Flood DDoS saldırısı

Sonuç olarak, sahte UDP paketlerinin hacmi, hedef sunucunun istekleri işleme ve yanıtlamaya yönelik maksimum kapasitesini aşarak hedef sunucunun kaynaklarının tükenmesine sebep olur.

2.6.3 HTTP seli saldırısı

HTTP, internet üzerinde yer alan sunucu ve istemciler arasındaki iletişimde bilgilerin nasıl iletileceğini belirleyen uygulama katmanı protokolüdür. Bu protokol, web siteleri için kullanılmakta olup HTTP oturumu, istemci sunucudan GET ve POST metotlarıyla sayfa talebinde bulunduktan sonra sunucunun istenilen sayfa bilgisini istemciye göndermesiyle tamamlanmaktadır. Şekil 2.13'te gösterilen HTTP Seli saldırısında saldırgan, çok sayıda meşru gözüken GET veya POST isteklerini bir sunucuya veya web uygulamasına göndererek sunucunun maksimum kaynakları kullanmaya zorlamaktadır. Bu saldırıda, HTTP isteğinin büyüklüğünü arttırmak için saldırganın kontrol sağladığı zararlı yazılım bulaştırılmış bilgisayarlardan oluşan "BotNet" olarak isimlendirilen ağlar kullanılmaktadır. Saldırının engellenmemesi için zararlı yazılım bulaştırılmış bilgisayarların gerçek IP adresleri kullanılmaktadır.



Şekil 2.13 HTTP Flood DDoS saldırısı

2.7 Derin Öğrenme

Derin öğrenme, çok katmanlı yapay sinir ağları aracılığıyla büyük hacimli ve karmaşık veri setlerinden anlamlı örüntüler öğrenmeyi amaçlayan makine öğrenmesi alt alanıdır. Geleneksel makine öğrenmesi yöntemleri çoğunlukla elle seçilmiş özelliklere ve sınırlı model derinliğine dayanırken, derin öğrenme yöntemleri ham veriden otomatik olarak özellik çıkarımı yapabilme yeteneğine sahiptir. Bu özellik, derin öğrenmeyi özellikle yüksek boyutlu ve doğrusal olmayan veri yapılarının bulunduğu siber güvenlik alanında güçlü bir araç haline getirmektedir.

Siber saldırı tespit sistemlerinde ağ trafiği verileri genellikle yüksek hacimli, dengesiz ve gürültülü yapıdadır. Derin öğrenme modelleri, bu tür karmaşık veri yapıları üzerinde başarılı sonuçlar üretebilmekte ve geleneksel yöntemlere kıyasla daha yüksek tespit oranları sağlayabilmektedir. Bu çalışmada, derin öğrenme yöntemleri kapsamında Çok Katmanlı Algılayıcı (MLP) ve Üretici

Karşıt Ağlar (GAN) ayrıntılı olarak ele alınmakta ve bu iki yaklaşımın birlikte kullanıldığı GAN destekli MLP modeli önerilmektedir.

2.8 Derin Öğrenme Algoritmaları

Literatürde çok sayıda farklı derin öğrenme algoritması bulunmaktadır; ancak bu çalışmada kullanılan derin öğrenme algoritmalasi Çok Katmanlı Algılayıcı yer verilecektir.

2.8.1 Çok Katmanlı Algılayıcı (MLP)

Çok Katmanlı Algılayıcı (MLP), ileri beslemeli yapay sinir ağlarının en temel ve en yaygın kullanılan türlerinden biridir. MLP, biyolojik sinir sistemlerinden esinlenilerek geliştirilmiş olup, girdi katmanı, bir veya daha fazla gizli katman ve çıktı katmanından oluşan hiyerarşik bir yapıya sahiptir. Derin öğrenme bağlamında MLP, birden fazla gizli

katman içermesi durumunda karmaşık ve doğrusal olmayan ilişkileri öğrenebilme yeteneğine sahiptir.

MLP mimarisinde her bir nöron, bir önceki katmandan gelen girdileri belirli ağırlıklar ile çarparak toplar ve elde edilen sonucu bir aktivasyon fonksiyonundan geçirerek bir sonraki katmana iletir. Bu yapı sayesinde MLP, doğrusal olmayan karar sınırlarını modelleyebilmekte ve karmaşık sınıflandırma problemlerinde başarılı sonuçlar üretebilmektedir.

2.8.1.1 MLP ağ mimarisi

MLP modeli üç temel bileşenden oluşmaktadır:

Girdi Katmanı: Ağ trafiği verilerine ait özelliklerin modele aktarıldığı katmandır. Siber saldırı tespitinde girdi katmanı; paket sayısı, bağlantı süresi, bayrak bilgileri, protokol türleri ve istatistiksel trafik özellikleri gibi özniteliklerden oluşmaktadır.

Gizli Katmanlar: Öğrenme sürecinin gerçekleştirildiği katmanlardır. Gizli katman sayısı ve nöron sayısı, modelin öğrenme kapasitesini doğrudan etkilemektedir. Derin MLP modelleri, karmaşık saldırı örüntülerini daha başarılı şekilde öğrenebilmektedir.

Çıktı Katmanı: Modelin nihai tahminlerini ürettiği katmandır. Çok sınıflı saldırı tespit problemlerinde genellikle Softmax aktivasyon fonksiyonu kullanılarak her sınıf için olasılık değerleri hesaplanmaktadır.

2.8.1.2 Aktivasyon fonksiyonları

MLP modellerinde doğrusal olmayan öğrenmenin sağlanabilmesi için aktivasyon fonksiyonları kritik bir role sahiptir. En yaygın kullanılan aktivasyon fonksiyonları ReLU, Sigmoid ve Tanh fonksiyonlarıdır. Siber saldırı tespit uygulamalarında gizli katmanlarda ReLU aktivasyon fonksiyonu, çıktı katmanında ise probleme bağlı olarak Sigmoid veya Softmax fonksiyonları tercih edilmektedir.

2.8.1.3 MLP'nin öğrenme süreci

MLP modelleri, geri yayılım (backpropagation) algoritması ve gradyan inişi tabanlı optimizasyon yöntemleri kullanılarak eğitilmektedir. Eğitim sürecinde, modelin ürettiği çıktı ile gerçek etiketler arasındaki hata hesaplanmakta ve bu hata değeri ağ boyunca geriye doğru yayılmaktadır. Ağırlıklar, hata fonksiyonunu minimize edecek şekilde güncellenmektedir.

2.8.1.4 MLP'nin siber saldırı tespitindeki rolü

MLP, etiketli veri setleri üzerinde eğitilerek normal ve saldırı trafiğini ayırt etme konusunda yüksek başarı sağlamaktadır. Özellikle bilinen saldırı türlerinin tespitinde MLP tabanlı modellerin yüksek doğruluk oranları sunduğu literatürde rapor edilmiştir. Ancak veri setlerinde sıklıkla karşılaşılan sınıf dengesizliği problemi, MLP modellerinin nadir görülen saldırı türlerini öğrenmesini zorlaştırmaktadır.

2.9 Üretici Karşıt Ağlar (GAN)

Üretici Karşıt Ağlar (GAN), Ian Goodfellow ve arkadaşları tarafından önerilen ve denetimsiz öğrenme yaklaşımına dayanan derin öğrenme mimarilerinden biridir. GAN mimarisi, bir üretici (Generator) ve bir ayırt edici (Discriminator) ağdan oluşmaktadır. Bu iki ağ, birbirine karşıt bir öğrenme süreci içerisinde eğitilmektedir.

Üretici ağ, gerçek veriye benzer sentetik örnekler üretmeyi amaçlarken; ayırt edici ağ, kendisine sunulan verilerin gerçek mi yoksa üretici tarafından oluşturulmuş sahte veriler mi olduğunu ayırt etmeye çalışmaktadır. Bu karşıt öğrenme süreci sonucunda üretici ağ giderek daha gerçekçi veriler üretmeyi öğrenmektedir.

2.9.1 GAN yapısı ve çalışma prensibi

GAN eğitim süreci, minimaks optimizasyon problemi olarak modellenmektedir. Üretici ve ayırt edici ağlar, birbirlerinin performansını artıracak şekilde eş zamanlı olarak eğitilmektedir. Eğitim süreci dengeye ulaştığında, üretilen sentetik veriler gerçek verilerden ayırt edilemez hale gelmektedir.

2.9.2 GAN'ın siber güvenlikte kullanımı

GAN modelleri, siber saldırı tespitinde özellikle anomali tespiti ve veri seti dengeleme amacıyla kullanılmaktadır. GAN'ler, nadir görülen saldırı türlerine ait sentetik örnekler üreterek sınıf dengesizliği problemini azaltmakta ve sınıflandırma modellerinin genelleme yeteneğini artırmaktadır. Ayrıca GAN'ler, bilinmeyen veya sıfırıncı gün saldırılarının modellenmesine katkı sağlamaktadır.

2.10 GAN Destekli MLP Yaklaşımı

GAN destekli MLP yaklaşımı, bu tez çalışmasının özgün omurgasını oluşturmaktadır. Bu yaklaşımda GAN modeli, ağ trafiği verileri kullanılarak eğitilmekte ve sentetik saldırı örnekleri üretilmektedir. Üretilen bu veriler, MLP modelinin eğitim sürecine dahil edilerek veri seti dengelenmektedir.

Bu hibrit yapı sayesinde MLP modelinin nadir görülen saldırı türlerini öğrenme kapasitesi artırılmakta ve aşırı öğrenme riski azaltılmaktadır. GAN destekli MLP modeli, özellikle TCP SYN Flood ve UDP Flood saldırılarının tespitinde daha yüksek doğruluk ve F1-skoru değerleri elde edilmesini hedeflemektedir. GAN destekli yaklaşımın temel avantajları aşağıda özetlenmiştir:

- Veri seti dengesizliğinin azaltılması
- Modelin genelleme yeteneğinin artırılması
- Bilinmeyen saldırı türlerine karşı daha dirençli yapı

Sonuç olarak, GAN destekli MLP yaklaşımı, derin öğrenme tabanlı siber saldırı tespit sistemleri için etkili ve yenilikçi bir çözüm sunmaktadır.

2.11 İlgili Araştırmalar

Bilişim sistemlerine yönelik siber tehditlerin hacminin ve karmaşıklığının artması, ağ güvenliğinin sağlanmasını modern çağın en kritik problemlerinden biri haline getirmiştir. Geçmişte yaygın olarak kullanılan imza tabanlı sistemler ve makine öğrenme algoritmaları gibi geleneksel sığ makine öğrenimi yaklaşımları, günümüzün yüksek hacimli ağ trafiğini işlemede ve bilinmeyen saldırı türlerini tespit etmede yetersiz kalmaktadır. Shone vd. (2018), geleneksel yöntemlerin özellikle manuel öznitelik çıkarımı gerektirmesinin, sistemlerin ölçeklenebilirliğini ve gerçek zamanlı analiz yeteneğini sınırladığını vurgulamaktadır. Bu kısıtlamalar, araştırmacıları insan müdahalesini minimize eden ve ham veriden otomatik öğrenme yeteneğine sahip derin öğrenme yöntemlerine yöneltmiştir.

Derin öğrenme mimarileri, çok katmanlı yapıları sayesinde verilerdeki karmaşık ve doğrusal olmayan ilişkileri modelleme konusunda üstün bir başarı sergilemektedir. Sarker (2021), derin öğrenme tabanlı modellerin, siber güvenlik alanında dinamik saldırı vektörlerini tanımlamada geleneksel yöntemlere kıyasla daha esnek ve güçlü bir çözüm sunduğunu belirtmektedir. Özellikle CNN ve LSTM ağlarının hibrit kullanımları, ağ trafiğindeki hem mekansal hem de zamansal özellikleri yakalamada etkili olmaktadır (Vinayakumar vd., 2019). Bu bölümde, siber saldırıların tespiti amacıyla literatürde önerilen derin öğrenme yaklaşımları incelenmiş; kullanılan veri setleri, mimari tasarımlar ve elde edilen performans sonuçları karşılaştırmalı olarak değerlendirilerek mevcut araştırma boşlukları ortaya konulmuştur.

2.11.1 CICIDS2017 veri seti ile yapılan çalışmalar

Chen vd. (2019), DDoS saldırılarını tespit etmek için DAD-MCNN (çok kanallı bir evrişimli sinir ağı) çerçevesi önermiştir. Özellik gruplarının sayısı, kanal sayısını belirlemektedir. Araştırmacılar özellikleri paket düzeyi, ana bilgisayar düzeyi ve trafik

düzeyi gibi farklı seviyelere ayırmışlardır. MC-CNN'nin eğitimi için artımlı eğitim yaklaşımı kullanılmıştır. Araştırmacılar, her iki veri kümesinde ikili sınıflandırma ve yalnızca KDDCUP99 veri kümesinde çok sınıflı sınıflandırma için KDDCUP99 ve CICIDS2017 veri kümeleri üzerinde bir dizi test gerçekleştirmiştir. Ayrıca MC-CNN, CNN, LSTM (3 katman) ve diğer sığ makine öğrenmesi yöntemleri (RF, SVM, C4.5 ve KNN) ile karşılaştırılmıştır. Sonuçlar, MC-CNN'nin tüm ikili ve çok sınıflı sınıflandırma görevlerinde mevcut en iyi yöntemlerden daha başarılı olduğunu göstermiştir. Bunun yanında, araştırmacılar eğitim veri kümesinin boyutunu değiştirerek CNN ve MC-CNN'yi değerlendirmiştir. Sonuçlar, MC-CNN'nin sınırlı veri kümesi koşullarında daha iyi performans gösterdiğini ve eğitim verilerinin nispeten yetersiz olduğu durumlarda DDoS tespit sistemleri oluşturmak için faydalı olduğunu ortaya koymuştur. Çok kanallı ve tek kanallı modellerin sonuçları arasında kayda değer bir fark bulunmamaktadır. Ayrıca, çok kanallı modeller karmaşıklığı artırmakta ve bu nedenle gerçek zamanlı senaryolar üzerinde doğrulandığında uygun olmayabilmektedir.

Sabeel vd. (2019), bilinmeyen DoS/DDoS saldırılarının tahmini için Derin Sinir Ağı (DNN) ve LSTM olmak üzere iki makine öğrenmesi modeli önermiştir. Çalışmada, yazarlar öncelikle modellerini CICIDS2017 veri setindeki ön işlemden geçirilmiş DoS/DDoS örnekleri üzerinde eğitmiş ve doğruluğu ölçmek için sonuçları sentezlenmiş ANTS2019 veri seti üzerinde değerlendirmiştir. İkinci aşamada ise yazarlar sentezlenmiş veri seti ile CICIDS2017 veri setini birleştirmiştir. Modeller bu birleşik veri seti üzerinden yeniden eğitilmiş ve yeni sentezlenmiş bilinmeyen saldırılara yönelik tespit performansı değerlendirilmiştir. Bu modellerin performansı, deneyin ikinci kısmında önemli iyileşmeler göstermiş; sırasıyla DNN ve LSTM %98,72 ve %96,15 doğruluk oranlarına ulaşmıştır. DNN ve LSTM modellerinin AUC değerleri ise sırasıyla 0,987 ve 0,989 olarak hesaplanmıştır. ANTS2019 veri seti, gerçek hayattaki saldırıları taklit etmek amacıyla sentetik olarak oluşturulmuştur. İkili sınıf sınıflandırması yapılan çalışmada, gerçek zamanlı tespit düzeneği kullanılmamıştır. Haider vd. (2020), Yazılım Tanımlı Ağlarda DDoS saldırılarının tespiti için derin bir CNN çerçevesi önermiş ve önerilen bu ensemble mekanizması CICIDS2017 veri seti üzerinde değerlendirilmiştir. Bu çözüm, mevcut en gelişmiş derin öğrenme tabanlı ensemble ve melez yaklaşımlarla (örneğin, RNN, LSTM, RL) karşılaştırılmıştır. Ensemble CNN yöntemi, diğer üç önerilen derin öğrenme

yaklaşımından daha iyi performans göstermiş ancak eğitim ve test süreleri arasında bir ödünleşim söz konusudur. Yazarlar ayrıca, önerilen ensemble CNN yaklaşımını mevcut rakip yöntemlerle de karşılaştırmıştır. Sonuçlar, ensemble CNN yaklaşımının mevcut rakip yöntemlerden daha üstün performans sergilediğini göstermiştir. Ensemble CNN yöntemi %99,45'lik bir doğruluk oranı elde etmiştir. Bu yaklaşımın eğitim ve test süreleri diğer yaklaşımlara kıyasla daha yüksektir. Bu durum, saldırı azaltma mekanizmasını olumsuz etkileyebilir. Dolayısıyla, saldırıların daha fazla zarara yol açması mümkündür.

Wang ve Liu (2020), Yazılım Tanımlı Ağ (SDN) ortamında DDoS saldırılarını tespit etmek için bir bilgi entropisi ve derin öğrenme yöntemi önermiştir. Yöntem, saldırıların tanımlanması için iki aşamalı bir tespit mekanizması kullanmaktadır. İlk aşamada, kontrolcü şüpheli trafiği bilgi entropisi tespiti yoluyla inceler. Ardından, bir CNN modeli, normal trafik ile saldırı trafiğini ayırt etmek için ince taneli paketler temelinde tespit işlemini gerçekleştirir. Yazarlar, yöntemlerini DNN, SVM ve Karar Ağacı (DT) yöntemleri ile karşılaştırmıştır. CNN modeli, karşılaştırılan yöntemler arasında daha yüksek kesinlik, doğruluk, F1-skoru ve hatırlama değerleri elde etmiştir. Modelin doğruluk oranı %98,98'dir. CNN'nin ROC eğrisi, DNN, SVM ve DT'ninkilere kıyasla daha diktir ve AUC değeri 0,949'dur. Bilgi entropisine dayalı tespit yöntemi için bir eşik değerinin belirlenmesi gerekmektedir.

Asad vd. (2020), uygulama katmanı DDoS saldırılarından hizmetleri korumak için ileri beslemeli geri yayılım mimarisine dayanan bir DNN mimarisi (DeepDetect) sunmuştur. Önerilen yaklaşım, DDoS tespiti için CICIDS2017 veri seti kullanılarak değerlendirilmiştir. Yazarlar yöntemlerini Rastgele Orman (RF) ve DeepGFL ile karşılaştırmıştır. DeepDetect, 0.99 F1-skor değeri elde ederek diğer yaklaşımlardan daha üstün performans göstermiştir. Ayrıca, AUC değerinin 1'e çok yakın olması, önerilen model tarafından yüksek bir doğruluğa ulaşıldığını göstermektedir. Bu çalışmada araştırmacılar çok sınıflı sınıflandırma yapmış ve bu yaklaşım, uygulama katmanı DDoS saldırılarına karşı güvenlik sağlamak üzere bir web servisi olarak buluta entegre edilmiştir. Yöntem yalnızca Uygulama katmanı DDoS saldırıları üzerinde değerlendirilmiştir.

Muraleedharan ve Janet (2020), HTTP üzerindeki yavaş DoS saldırılarını tespit etmek için akış verisi tabanlı bir derin sinir ağı sınıflandırma modeli önermiştir. Sınıflandırma modeli, Tam Bağlı (FC) ileri beslemeli bir derin ağ kullanmıştır. Model, yalnızca DoS örneklerinin seçildiği CICIDS2017 veri seti üzerinde değerlendirilmiştir. Sınıflandırıcı, DoS saldırılarının türünü tespit edebilmektedir. Elde edilen sonuçlar, modelin saldırıları %99,61'lik genel bir doğruluk oranıyla sınıflandırabildiğini göstermiştir. Bu yaklaşım, yalnızca CICIDS2017 veri seti üzerindeki HTTP yavaş DoS saldırıları (Slowloris, SlowHTTP, Hulk, GoldenEye) ile değerlendirilmiştir.

Liang ve Znati (2019), iki LSTM katmanı, bir dropout katmanı ve bir Tam Bağlı (FC) katmandan oluşan dört katmanlı bir mimari model önermiştir. Bu yaklaşımda, elle yapılan öznitelik mühendisliği gerekliliği ortadan kaldırılmış ve ağ trafik davranışı doğrudan paketlerin kısa dizilerinden öğrenilmiştir. Söz konusu makalede, CICIDS2017 veri setinin Çarşamba ve Cuma verileri kullanılarak diğer üç algoritmayla (Karar Ağacı-DT, Yapay Sinir Ağı-ANN, SVM) karşılaştırmalı üç deney gerçekleştirilmiştir. Gözlemlenen sonuçlara göre; Deney 1, LSTM tabanlı şemanın, ham girdide gömülü olan karmaşık akış seviyesindeki öznitelik tanımlarını başarıyla öğrendiğini ve diğer yaklaşımlardan daha iyi performans gösterdiğini ortaya koymuştur. Deney 2'nin sonucu, önerilen şemanın bilinmeyen ağ trafiğinin dinamik davranışlarını doğru bir şekilde yakalayabildiğini göstermiştir. Deney 3 ise, her bir akış için modelin daha fazla paketi test etmesine izin verilmesinin (artan n değerleriyle) performansı her zaman artırmadığı sonucuna varmıştır. Önerilen şema, bilinmeyen trafik üzerinde geleneksel makine öğrenmesi yöntemlerinden daha üstün performans sergilemiştir. Önerilen model, bir akışın n adet paketinden oluşan bir alt dizi ($S \subset F$) kullanmaktadır. Bir akışta yeterli sayıda paket bulunmaması durumunda, S yapay paketlerle doldurulmaktadır. Bu doldurma değerleri, önerilen modelin öğrenme sürecini etkileyebilmekte ve performans düşüşüne neden olabilmektedir.

Kasim (2020) çalışmasında, yazar AE-SVM (Otokodlayıcı Destek Vektör Makineleri) yaklaşımını önermiştir.Yazarlar,önerdikleri modeli aşağıdaki test senaryoları üzerinde değerlendirmiştir: 16.902 veri üzerinde eğitilen modeli CICIDS veri setinden rastgele seçilen 15.000 veri üzerinde test edilmiştir, Kali Linux ortamında oluşturulan 6957 DDoS saldırısından oluşan veri seti üzerinde test edilmiştir, NSL-KDD eğitim veri seti ve on

katlı çapraz doğrulama kullanılarak eğitilmiştir, NSL-KDD test veri seti üzerinde test edilmiştir. AE-SVM yöntemi, düşük yanlış pozitif oranı ve hızlı anomali tespiti açısından diğer yöntemlerden daha üstün performans göstermiştir. Bununla birlikte, önerilen modelin NSL-KDD veri seti üzerindeki doğruluk oranı, diğer iki veri setine kıyasla daha düşük bulunmuştur.

Bhardwaj vd. (2020), ağ trafiği sınıflandırması için öznitelik öğrenmek amacıyla istiflenmiş seyrek bir otokodlayıcı (AE) ile bir derin sinir ağını birleştiren bir yaklaşım önermiştir. İlk olarak, temel model olarak Naif AE ve DNN ele alınmış ve her ikisi için de rastgele hiperparametre değerleri kullanılmıştır. Ardından, AE ve DNN modelindeki iyileştirmeler için Naif AE ve DNN optimize edilmiştir. Önerilen yaklaşım, on adet güncel yöntemle karşılaştırılmıştır. NSL-KDD veri seti üzerinde karşılaştırılan yöntemler SAECSMR, AECGaussian NB, RNN, MLP, AECSVM ve SAVAERCDNN'dir. CICIDS2017 veri seti üzerinde karşılaştırılan yöntemler ise DT, ANN, SVM, SAVAERCDNN ve LSTM'dir. Sonuçlar, önerilen yaklaşımın NSL-KDD veri seti üzerinde %98,43 doğruluk oranıyla mevcut yöntemlerden daha üstün performans gösterdiğini ve CICIDS2017 veri seti üzerinde %98,92'lik bir doğruluk sağlayarak rekabetçi sonuçlar ürettiğini ortaya koymuştur. Önerilen yöntem, öznitelik öğrenimi ve aşırı uyum sorunuyla başa çıkmak için yeterlidir. Öznitelik öğrenimi, AE'nin eğitim verisinin rastgele örnekleriyle eğitilmesiyle sağlanmış; aşırı uyum sorunu ise seyreklik parametresi kullanılarak önlenmiştir. Bu makalede güncel bir veri seti değerlendirilmemiş ve çevrimdışı analiz yapılmıştır. Ayrıca, önerilen model için tespit süresi hesaplanmamıştır.

Roopak vd. (2019), Çok Katmanlı Algılayıcı (MLP), CNN, LSTM ve melez CNN-LSTM modeli olmak üzere dört derin öğrenme modeli önermiş ve bu modelleri makine öğrenmesi algoritmalarıyla (SVM, Bayes ve Rastgele Orman - RF) karşılaştırmıştır. Yazarlar, modelleri dengesiz bir veri seti olan CICIDS2017 üzerinde değerlendirmiş ve bu veri seti, verilerin kopyalanması yoluyla dengeli hale getirilmiştir. Melez CNN-LSTM modelinin, diğer derin öğrenme ve makine öğrenmesi modellerine kıyasla daha iyi performans gösterdiği gözlemlenmiştir. Bu model, %97,16'lık bir doğruluk ve %99,1'lik bir hatırlama (recall) oranı sağlamıştır. Bununla birlikte, veri setinin dengeli hale getirilmesinde kullanılan yöntem belirtilmemiş ve önerilen modelin Nesnelerin İnterneti (IoT) ağları için çevrimdışı analizi yapılmıştır.

Roopak vd. (2020), ön işlemden geçirilmiş veri seti üzerinde öznitelik seçimi için Egemen Olmayan Bireylerin Sıralanması Algoritması (NSGA) yöntemi olarak da bilinen çok amaçlı bir optimizasyon tekniği kullanmıştır. Bu çalışmada, saldırıyı sınıflandırmak için CNN ve LSTM birleşiminden yararlanılmıştır. Deneylerde, GPU kullanılarak CICIDS2017 veri seti üzerinde çalışılmıştır. Önerilen yöntem, %99,03'lük yüksek bir doğruluk ve %99,36'lık bir F1-skor değeri elde etmiştir. Yazarlar ayrıca yöntemlerini MLP, SVM, RF, Bayes ve diğer güncel tekniklerle karşılaştırmıştır. Sonuçlar, önerilen modelin diğer çalışmalardan daha üstün performans sergilediğini göstermiştir. Modelin eğitim süresi, diğer derin öğrenme yöntemlerine kıyasla 11 kat daha düşüktür. Bununla birlikte, bu makalede karşılaştırma yapılan güncel tekniklerin çoğu CICIDS2017 veri setini kullanmamaktadır. Bu nedenle, yapılan karşılaştırmanın uygun olmadığı değerlendirilmektedir.

2.11.2 CICDDoS2019 veri seti ile yapılan çalışmalar

Sbai ve El Boukhari (2020), CICDDoS2019 veri setini kullanarak Mobil Çevresel Ağlarda (MANET'ler) veri seli veya UDP seli saldırılarını tespit etmek için iki gizli katman ve 6 epok içeren bir DNN modeli önermiştir. Yazarlar, modeli CICDDoS2019 veri seti ile eğitmiş ve değerlendirmiştir. Önerilen model, oldukça umut verici olan şu sonuçları elde etmiştir: Hatırlama (Recall): 1, Kesinlik (Precision): 0.99, F1-skoru: 0.99, Doğruluk (Accuracy): 0.99. Bu makalede, yazarlar yalnızca CICDDoS2019 veri setindeki veri seli veya UDP seli saldırısı üzerinde çalışmıştır. Assis vd. (2020), bir Yazılım Tanımlı Ağ (SDN) savunma sistemi önermiştir. Bu savunma sistemi, harici bir hedef sunucuya ve kontrolcüye yönelik DDoS saldırılarını tespit edip etkisiz hale getirmektedir. Saldırılar, bir tespit modülü aracılığıyla belirlenmektedir. Bu modülde, yazarlar SDN trafik davranışını inceleyerek DDoS saldırılarını tespit etmek için derin öğrenme tabanlı bir CNN yöntemi kullanmıştır. Meşru kullanıcılar üzerindeki DDoS etkisini azaltmak için IP akış verilerinin bir saniyelik aralıklarla çıkarılıp analiz edilmesi nedeniyle, önerilen yöntem gerçek zamanına yakın bir şekilde çalışmaktadır. Tespit modülündeki önerilen CNN yaklaşımı, diğer üç anomali tespit yaklaşımıyla (Lojistik Regresyon - LR, MLP ağı ve Yoğun MLP) karşılaştırılmıştır. Yazarlar, yukarıdaki tespit yöntemlerini iki test senaryosu üzerinde denemiştir: ilki simüle edilmiş SDN verilerini kullanırken, ikincisi

CICDDoS2019 veri setini kullanmaktadır. Genel sonuçlar, CNN'in tüm bu test senaryolarında DDoS saldırılarını tespit etmede etkili olduğunu göstermiştir. Saldırıyı etkisiz hale getiren modülde, SDN kontrolcüsünde bir Görev Ağacı (GT) tabanlı teknik uygulanmıştır. Sonuçlar, azaltma yönteminin SDN'in düzenli işleyişini verimli bir şekilde eski haline getirdiğini ortaya koymuştur. Önerilen sistem, tespit ve azaltma süreçlerinin hızını artırmak için otonom olarak çalışmaktadır. Bununla birlikte, modelin CICDDoS2019 veri seti için doğruluk oranı daha düşüktür.

Hussain vd. (2020), görsel olmayan ağ trafiğini üç kanallı görsel formlara dönüştürmek için bir yöntem önermiştir. Bu yöntem, güncel DoS ve DDoS saldırılarını tespit etmek amacıyla, mevcut ve gelişmiş bir CNN modeli olan ResNet-18 üzerinde değerlendirilmiştir. Önerilen yöntem, herhangi bir kodlama veya dönüşüm tekniği kullanmadan, temizlenmiş ve normalleştirilmiş öznitelikleri kullanarak verileri görsellere dönüştürmüştür. Yazarlar ayrıca, ResNet-18 kullanan önerilen metodolojiyi güncel bir çözümle karşılaştırmış ve aynı veri seti üzerinde ondan daha üstün performans elde etmiştir. ResNet-18 kullanan önerilen metodoloji, ikili sınıf sınıflandırmasında %99,99 doğruluk elde etmiştir. Ayrıca, CICDDoS2019 veri seti üzerinde 11 farklı DoS ve DDoS saldırı türü için %87,06'lık bir doğruluk oranına ulaşmıştır. Görsel olmayan verilerin görsel verilere dönüştürülmesi için ön işlem süresi, gerçek zamanlı doğrulama için önemli bir metrik olmasına rağmen hesaplanmamıştır. Bunun yanı sıra, ResNet modeline girdi sağlamak için orijinal 60*60*3 boyutlarının 224*224*3 boyutlarına nasıl dönüştürüldüğü açıklanmamıştır.

Amaizu vd. (2021), 5G ve B5G ortamları için verimli, derin öğrenme tabanlı bir DDoS saldırısı tespit çerçevesi önermiştir. Önerilen çerçeve, farklı tasarıma sahip iki DNN modelinin bir Pearson Korelasyon Katsayısı (PCC) öznitelik çıkarımı algoritması ile birleştirilmesiyle geliştirilmiştir. Bu çerçeve, DDoS saldırılarını ve karşılaşılan DDoS saldırı türlerini tespit etmek üzere tasarlanmıştır.Yazarlar, önerilen çerçeveyi endüstri tarafından kabul görmüş bir veri seti (CICDDoS2019) üzerinde dört farklı senaryo kullanarak değerlendirmiştir. Sonuçlar, çerçevenin DDoS saldırılarını %99,66 doğruluk ve 0,011 kayıp oranıyla tespit edebildiğini göstermiştir. Ayrıca, önerilen tespit çerçevesinin sonuçları, mevcut yaklaşımlarla (K-En Yakın Komşu - KNN, SVM,

DeepDefense ve CNN Ensemble) karşılaştırılmıştır. Önerilen çerçeve, CNN Ensemble hariç tüm yaklaşımlardan daha üstün performans sergilemiştir. CNN Ensemble, önerilen çerçeveye kıyasla daha yüksek kesinlik ve hatırlama oranlarına sahiptir. Önerilen model karmaşık bir yapıya sahip olduğundan daha uzun tespit süresi gerektirebilmekte ve bu durum, modelin gerçek zamanlı senaryolardaki performansını olumsuz etkileyebilmektedir.

Cil vd. (2021), yapısında hem öznitelik çıkarımı hem de sınıflandırma süreçlerini içeren bir derin öğrenme modeli önermiştir. DNN modeli, 69 birimli bir giriş katmanı, her biri 50 birimden oluşan üç gizli katman ve iki birimli bir çıkış katmanından oluşmaktadır. Yazarlar, CICDDoS2019 veri setini Veri Seti 1 ve Veri Seti 2 olarak ikiye ayırmıştır. Veri Seti 1, normal ve saldırı trafiği olmak üzere iki tür trafik şeklinde kategorize edilmiştir. Veri Seti 2 ise DDoS saldırı türlerini tanımlamak için oluşturulmuştur. DNN modeli, Veri Seti 1 üzerinde DDoS saldırı tespiti için yaklaşık %100 doğruluk elde etmiş ve böylece gerçek zamanlı senaryolar için uygun olan erken aksiyon alınmasında güvenilir bir sonuç başarmıştır. Ayrıca model, Veri Seti 2 üzerinde DDoS saldırılarını yaklaşık %95 doğrulukla başarılı bir şekilde sınıflandırmıştır. Bununla birlikte, önerilen model çok sınıflı sınıflandırma durumunda daha düşük doğruluk sağlamaktadır.

Shurman vd. (2020), DoS/DDoS saldırılarını tespit etmek için hibrit tabanlı bir Güvenlik İhlali Tespit Sistemi (IDS) ve LSTM tabanlı bir derin öğrenme modeli olmak üzere iki yöntem önermiştir. İlk yöntem olan ve bir uygulama olarak tanımlanan IDS çerçevesi, herhangi bir ağ cihazından gelen kötü amaçlı ağ trafiğini, çalışan IP veri setleriyle karşılaştırarak tespit edebilmektedir. Bu çerçeve, istenmeyen IP'leri engelleme kapasitesine sahiptir. İkinci yöntemde ise LSTM kullanılmış ve bu model, çeşitli DrDoS saldırı türlerini içeren CICDDoS2019 veri seti üzerinde eğitilmiştir. İkinci model, mevcut diğer modellerle karşılaştırılmış ve sonuçlar modelin diğer modellerden daha üstün performans gösterdiğini ortaya koymuştur. LSTM tabanlı model, yansıtma tabanlı CICDDoS2019 veri seti üzerinde %99,19'luk bir doğruluk göstermiştir; ancak yalnızca yansıtma tabanlı CICDDoS2019 veri seti kullanılmıştır. Ayrıca, hibrit IDS ve LSTM yöntemleri birbirinden bağımsız olarak çalışmaktadır.

Assis vd. (2021), SDN ortamında DDoS ve sızma saldırılarına karşı bir savunma sistemi önermiştir. Önerilen sistem, tespit ve azaltma modülleri olmak üzere iki temel modülden oluşmaktadır. Saldırılar, tespit modülü tarafından belirlenmektedir. Bu modülde yazarlar, tekli IP akış kayıtlarını analiz ederek DDoS ve sızma saldırılarını tespit etmek için derin öğrenme tabanlı bir GRU yöntemi kullanmıştır. Azaltma modülü ise tespit edilen saldırılara karşı etkili aksiyonlar almaktadır. Yazarlar, önerdikleri modeli yedi farklı makine öğrenmesi yaklaşımıyla (DNN, CNN, LSTM, SVM, LR, KNN ve GD) iki veri seti (CICDDoS2019 ve CICIDS2018) üzerinde test etmiştir. İki test senaryosu belirlenmiştir: ilki CICDDoS2019, ikincisi ise CICIDS2018 veri seti için. Her iki senaryoda da yazarlar, önerilen modeli diğer ML yöntemleriyle doğruluk, kesinlik, hatırlama, F-ölçüm ve yöntemlerin normal ve saldırı akışlarını ayrı ayrı sınıflandırma etkinliği açısından karşılaştırmıştır. Sonuçlar, GRU'nun tüm bu test senaryolarında DDoS ve sızma saldırılarını tespit edebildiğini göstermiştir. Ayrıca, tespit yöntemlerinin saniyede analiz edip sınıflandırabildiği ortalama akış sayısı hesaplanarak bir uygulanabilirlik testi de gerçekleştirilmiştir. Bu test, Londrina Eyalet Üniversitesi'nden toplanan gerçek IP akış verileri kullanılarak yapılmıştır. Sonuçlar, GRU'nun uygulanabilir bir yaklaşım olduğunu işaret etmiştir. Önerilen yaklaşımın CICDDoS2019 ve CICIDS2018 veri setleri için ortalama doğruluk, hatırlama, kesinlik ve F-ölçüm değerleri sırasıyla %99,94 ve %97,09'dur. Bu makalede, tespit ve eğitim süreleri hesaplanmamış ve veri setlerinin çevrimdışı analizi yapılmıştır.

Elsayed vd. (2020), SDN'lerde DDoS saldırılarını tespit etmek için DDoSNet'i önermiştir. DDoSNet, Tekrarlayan Sinir Ağı (RNN) ile Otokodlayıcıyı (AE) birleştiren derin öğrenme tabanlı bir tekniktir. Model, yeni bir veri seti olan CICDDoS2019 kullanılarak değerlendirilmiştir. Yazarlar ayrıca DDoSNet'i altı klasik makine öğrenmesi tekniğiyle (Karar Ağacı - DT, Naive Bayes - NB, Rastgele Orman - RF, SVM, Gradient Boosting - Booster ve Lojistik Regresyon - LR) karşılaştırmıştır. DDoSNet modelinin değerlendirmesi, modelin doğruluk, hatırlama, kesinlik ve F-skoru açısından mevcut altı klasik ML tekniğinden daha üstün performans gösterdiğini ortaya koymuştur. Yaklaşım, CICDDoS2019 veri seti üzerinde %99 doğruluk ve %98,8 AUC değeri elde etmiştir. Veri setinin çevrimdışı analizi yapılmış olup, çok sınıflı sınıflandırma gerçekleştirilmemiştir.

2.11.3 Başka veri setleri üzerine yapılan araştırmalar

Akbıyık (2024) tarafından hazırlanan tez çalışmasında, teknolojinin gelişimiyle artan siber tehditlere karşı KDD99 veri seti kullanılarak makine öğrenmesi tabanlı saldırı tespit sistemleri geliştirilmesi amaçlanmıştır. Çalışma kapsamında Karar Ağaçları, ADABoost, Yapay Sinir Ağları ve Destek Vektör Makineleri algoritmaları ile modeller oluşturulmuş ve performansları karşılaştırılmıştır. Yapılan analizler sonucunda en yüksek başarıyı Karar Ağaçları yöntemi göstermekle birlikte, tüm modellerin %99 ve üzeri doğruluk oranına ulaştığı tespit edilmiştir. Bu yüksek başarı oranlarının işaret ettiği "aşırı uyumlanma" (overfitting) riskini analiz etmek adına; veri setinde yinelenen değerlerin temizlenmesi, farklı örneklemler oluşturulması ve öznitelik seçimi gibi altı farklı deneysel yöntem uygulanmıştır. Elde edilen bulgular, uygulanan yöntemlere rağmen yüksek başarı oranlarının değişmediğini ve veri setinden kaynaklı aşırı uyumlanma riskinin tamamen giderilemediğini ortaya koymaktadır.

Al-Daffaie (2024) tarafından yürütülen tez çalışmasında, geleneksel imza ve kural tabanlı güvenlik önlemlerinin, gelişmiş siber tehditler ve Dağıtık Hizmet Reddi (DDoS) saldırıları karşısında yetersiz kalması sorununa çözüm olarak Quad-RNN (Quad Directional Recurrent Neural Network) adlı özgün bir derin öğrenme mimarisi geliştirilmiştir. Çalışma, önerilen dört yönlü (ileri, geri, sol-üst, sol-alt) girdi ve çıktı akışına sahip mimariyi, NSL-KDD ve DDoS veri setleri üzerinde “Bidirectional RNN” ve “Simple RNN” modelleri ile karşılaştırmalı olarak analiz etmiştir. Deneysel bulgular, Quad-RNN modelinin doğruluk, kesinlik, duyarlılık ve F1 skoru metriklerinde diğer mimarilerden üstün performans sergilediğini ve özellikle yanlış pozitif oranlarını belirgin şekilde düşürdüğünü ortaya koymuştur. Sonuç olarak çalışma, Quad-RNN mimarisinin karmaşık siber saldırı desenlerini yakalamada mevcut yöntemlere kıyasla daha dayanıklı ve uyarlanabilir bir savunma mekanizması sunduğunu kanıtlamıştır.

Çetin (2025) tarafından yürütülen tez çalışmasında, nesnelerin interneti ekosistemine yönelik siber tehditlerin tespiti amacıyla Oylama (Voting) ve Yığınlama (Stacking) yöntemlerine dayalı hibrit makine öğrenmesi modelleri geliştirilmiştir. Çalışmada, geniş kapsamlı saldırı türlerini içeren CICIoT2023 veri seti kullanılmış; Rastgele Orman,

Lojistik Regresyon ve Gaussian Naive Bayes algoritmaları temel sınıflandırıcılar olarak belirlenmiştir. Veri seti üzerinde ön işleme ve öznitelik çıkarımı yapıldıktan sonra, modeller GridSearchCV ile optimize edilerek çok sınıflı ve ikili sınıflandırma senaryolarında test edilmiştir. Analizler sonucunda, Yığınlama modelinin genel doğruluk ve azınlık sınıfların tespitinde daha üstün performans sergilediği, Oylama modelinin ise kaynak verimliliği açısından avantaj sağladığı belirlenmiştir. Sonuç olarak, hibrit modellerin tekil modellere kıyasla siber güvenlikte daha güvenilir çözümler sunduğu ortaya konmuştur.

Goun (2024) tarafından hazırlanan bu tez çalışmasında, akıllı şebekelerin güvenliğini tehdit eden ve özellikle Fazör Ölçüm Birimlerini (PMU) hedef alan Yanlış Veri Enjeksiyonu (FDI) saldırılarının tespiti için gelişmiş makine öğrenmesi ve derin öğrenme modelleri önerilmiştir. Çalışma kapsamında, Rastgele Orman ve Ekstra Ağaçlar gibi denetimli makine öğrenimi algoritmaları ile CNN, LSTM ve hibrit CNN-LSTM mimarileri karşılaştırmalı olarak analiz edilmiştir. Veri setindeki dengesizlik sorununu gidermek amacıyla Sentetik Azınlık Aşırı Örnekleme Tekniği (SMOTE) ve öznitelik seçimi gibi ön işleme yöntemleri uygulanarak ikili ve çok sınıflı sınıflandırma senaryoları test edilmiştir. Deneysel bulgular, mekansal ve zamansal özellikleri entegre edebilen hibrit CNN-LSTM modelinin %97,58 doğruluk oranıyla diğer tüm yöntemlerden daha üstün performans sergilediğini ve kritik enerji altyapılarının siber dayanıklılığını artırmada en etkili çözüm olduğunu ortaya koymuştur.

Özalp (2023) tarafından hazırlanan bu doktora tezi çalışmasında, geleneksel saldırı tespit sistemlerinin düşük doğruluk ve yüksek yanlış alarm oranlarını iyileştirmek amacıyla yapay zeka tabanlı hibrit modeller tasarlanmıştır. Çalışmanın ilk aşamasında NSL-KDD veri seti üzerindeki özniteliklerin frekans etkileri Random Forest ve J48 gibi algoritmalarla incelenmiştir. İkinci aşamada ise NSL-KDD ve CIC-IDS2018 veri setleri kullanılarak, öznitelik çıkarımı için Derin Öğrenme (CNN ve LSTM), sınıflandırma için ise Makine Öğrenmesi (XGBoost ve LightGBM) algoritmalarını entegre eden dört farklı model geliştirilmiştir. Deneysel analizler sonucunda, özellikle XGBoost algoritmasının SQL Enjeksiyonu ve DoS saldırılarının tespitinde, CNN- LSTM/LightGBM hibrit yapısının ise 14 farklı saldırı türünün genel tespitinde yüksek başarı sağladığı

görülmüştür. Çalışma, önerilen hibrit mimarilerin siber tehditleri tespit etmede literatürdeki mevcut yöntemlere kıyasla daha etkin ve güvenilir sonuçlar ürettiğini ortaya koymaktadır.

Delplace, Hermoso ve Anandita (2020) tarafından hazırlanan bu makalede, geleneksel imza tabanlı tespit sistemlerinin yetersiz kaldığı karmaşık siber saldırıların tespiti için NetFlow verileri kullanan makine öğrenimi yaklaşımları incelenmiştir. Çalışma kapsamında, CTU-13 veri seti kullanılarak Lojistik Regresyon, Destek Vektör Makineleri (SVM), Rastgele Orman, Gradyan Artırma ve Yoğun Sinir Ağları (DNN) algoritmalarının performansı karşılaştırılmıştır. Veri ön işleme aşamasında 22 öznitelik çıkarılmış ve özellik seçimi yöntemleri uygulanmıştır. Deneysel sonuçlar, Rastgele Orman algoritmasının en başarılı model olduğunu ve incelenen 13 botnet senaryosunun 8'inde %95'in üzerinde tespit oranı yakaladığını ortaya koymuştur. Ayrıca, veri setindeki dengesizlik sorununu gidermek ve model başarısını artırmak amacıyla bootstrapping tekniklerinin etkinliği tartışılmıştır.

Jullian vd. (2023) tarafından gerçekleştirilen bu çalışmada, nesnelerin interneti ağlarında artan siber güvenlik tehditlerine karşı derin öğrenme tabanlı ve dağıtık yapıda bir saldırı tespit çerçevesi önerilmiştir. Çalışma kapsamında, İleri Beslemeli Sinir Ağları (FFNN) ve LSTM modelleri, NSL- KDD ve BoT-IoT veri setleri kullanılarak eğitilmiş ve performansları karşılaştırılmıştır. Merkezi olmayan bu mimaride, saldırı tespiti uç ve sis katmanlarında gerçekleştirilirken, küresel parametre optimizasyonu bulut katmanında sağlanmaktadır. Veri dengesizliğini gidermek adına ön işleme teknikleri ve model optimizasyonu için Hyperband kullanılmıştır. Deneysel sonuçlar, önerilen dağıtık çerçevenin merkezi modellerle eşdeğer performans sergilediğini ve özellikle FFNN modeliyle %99,95'e varan doğruluk oranlarına ulaşarak IoT ağlarını korumada etkin bir çözüm sunduğunu ortaya koymuştur.

AlShahrani ve Quasim (2021) tarafından kaleme alınan bu makalede, artan siber güvenlik tehditlerine karşı derin öğrenme tabanlı yeni bir sınıflandırma tekniği olan AdaBoost Regresyon Sınıflandırıcısı (ABRC) önerilmiştir. Çalışma, siber saldırıların tespiti ve önlenmesi sürecinde sınıflandırma doğruluğunu artırmayı ve hesaplama süresini düşürmeyi amaçlamaktadır. Bu doğrultuda, AdaBoost algoritması ile lojistik regresyonun

sigmoidal fonksiyonu entegre edilerek hibrit bir yapı oluşturulmuştur. Önerilen modelin performansı, CICIDS 2019 veri seti kullanılarak Gradient Boosting, KNN, Karar Ağaçları ve SVM gibi mevcut yöntemlerle karşılaştırılmıştır. Analiz sonuçları, geliştirilen ABRC modelinin %95,87 doğruluk, %95,93 kesinlik ve 33 saniyelik işlem süresi ile diğer algoritmalara kıyasla daha üstün bir performans sergilediğini ve ağ güvenliğini sağlamada etkili bir çözüm sunduğunu ortaya koymaktadır. Ferrag vd. (2019) tarafından gerçekleştirilen bu çalışmada, siber güvenlikte saldırı tespiti için kullanılan yedi farklı derin öğrenme tekniğinin kapsamlı bir analizi sunulmuştur. Çalışma kapsamında DNN, RNN ve CNN gibi ayırt edici modeller ile Derin İnanç Ağları (DBN) ve Derin Otomatik Kodlayıcılar (DA) gibi üretken modeller incelenmiştir. Modellerin performansı, CSE-CIC-IDS 2018 veri seti kullanılarak ikili ve çok sınıflı sınıflandırma senaryolarında, TensorFlow kütüphanesi üzerinde test edilmiştir. Deneysel sonuçlar, ayırt edici modeller arasında CNN'in, üretken modeller arasında ise Derin Otomatik Kodlayıcıların (DA) en yüksek doğruluk oranlarına (%97,37) ve en düşük yanlış alarm oranlarına sahip olduğunu göstermektedir. Çalışma, modern ve karmaşık veri setleri üzerinde derin öğrenme mimarilerinin etkinliğini karşılaştırmalı olarak ortaya koymaktadır.

Mittal, Kumar ve Behal (2021) tarafından kaleme alınan çalışmada, teknolojinin insan hayatının ayrılmaz bir parçası haline gelmesi ve hızlanan dijitalleşme sürecinde artan siber güvenlik tehditlerini ele almaktadır. Çalışma, özellikle internet tabanlı hizmetleri ve uygulamaları felce uğratabilen DDoS saldırılarına odaklanmakta; saldırganların stratejilerini sürekli güncellemeleri nedeniyle geleneksel tespit mekanizmalarının yetersiz kaldığını vurgulamaktadır. Artan veri hacmi karşısında istatistiksel yöntemler ve sığ makine öğrenimi tekniklerinin sınırlılıklarını aşmak amacıyla, derin öğrenme yaklaşımlarının bu alandaki etkinliği ve gerekliliği savunulmaktadır. Yazarlar, IEEE Explore, ACM, ScienceDirect, Springer ve Google Scholar veritabanlarını kapsayan kapsamlı bir sistematik literatür taraması protokolü uygulamış ve 2018-2021 yılları arasındaki çalışmaları analiz etmiştir. Bu süreç sonunda seçilen 34 temel çalışma üzerinden derin öğrenme yöntemleri beş ana kategoride sınıflandırılmıştır: Denetimli örnek öğrenimi (DNN, CNN), denetimli dizi öğrenimi (RNN, LSTM, GRU), yarı denetimli öğrenme, hibrit öğrenme ve transfer öğrenimi gibi diğer yöntemler. İnceleme sonucunda, araştırmaların yaklaşık %50'sinin denetimli örnek öğrenimi tekniklerini

kullandığı, hibrit ve dizi öğrenimi yöntemlerinin ise daha az oranda tercih edildiği belirlenmiştir. Ayrıca çalışma, modellerin performansını artırmak için kullanılan Min-Max normalizasyonu ve One-hot encoding gibi veri ön işleme stratejilerinin önemine dikkat çekmektedir. Çalışma, literatürde kullanılan veri setlerini ve performans metriklerini de detaylı bir şekilde irdelemiştir. Araştırmalarda en sık kullanılan veri setleri arasında güncel olan CICIDS2017 ve CICDDoS2019'un yanı sıra, eski kabul edilen NSL-KDD ve KDDCUP'99 veri setlerinin de hala yer bulduğu görülmüştür. Modellerin başarısını ölçmek için en yaygın kullanılan metriğin doğruluk olduğu, bunu kesinlik, duyarlılık ve F1-skorunun takip ettiği belirtilmiştir. Bununla birlikte yazarlar, mevcut literatürdeki önemli araştırma boşluklarını da ortaya koymuştur. Özellikle çalışmaların çoğunun çevrimdışı veri setleri üzerinde yapıldığı, gerçek zamanlı sistemlerde dağıtım ve doğrulama eksikliği olduğu, veri setlerinin dengesiz yapısının modelleri etkilediği ve IoT gibi sınırlı kaynağa sahip ağlar için hafif modellerin eksikliği vurgulanmıştır. Sonuç olarak, derin öğrenme modelleri yüksek doğruluk oranlarına ulaşsa da, sıfırıncı gün saldırılarına karşı adaptasyon ve otomatik savunma sistemlerinin geliştirilmesi gelecekteki çalışmalar için kritik bir ihtiyaç olarak tanımlanmıştır.

Alabdulatif (2025) tarafından yazılan bir makalede, siber tehditlerin karmaşıklığının artmasıyla birlikte geleneksel Tespit Sistemlerinin (IDS) yetersiz kalması sorununa çözüm olarak, Açıklanabilir Yapay Zeka (XAI) ile güçlendirilmiş yeni bir topluluk (ensemble) derin öğrenme mimarisi önerilmiştir. Çalışma, CNN, LSTM ve Kapılı Tekrarlayan Birimler (GRU) modellerini "soft voting" mekanizmasıyla birleştiren hibrit bir yapı sunmaktadır. Önerilen modelin performansı, kapsamlı ve modern bir veri seti olan CICIDS2017 üzerinde hem ikili hem de çok sınıflı sınıflandırma senaryolarında test edilmiştir. Ayrıca, modelin karar verme süreçlerini şeffaflaştırmak ve güvenilirliğini artırmak amacıyla SHAP (Shapley Additive exPlanations) yöntemi kullanılarak özniteliklerin tahmin üzerindeki etkileri analiz edilmiştir. Deneysel sonuçlar, geliştirilen topluluk modelinin ikili sınıflandırmada %99,64, çok sınıflı sınıflandırmada ise %99,47 doğruluk oranına ulaşarak tekil modellere (CNN, LSTM, GRU) ve literatürdeki diğer güncel çalışmalara kıyasla üstün bir performans sergilediğini kanıtlamıştır.

Çizelge 2.1 Derin öğrenme yöntemleri kullanılarak siber saldırı tespiti yapan çalışmaların karşılaştırması

Taksonomi	Kaynaklar (Yazarlar)	Yayın Yılı	Kullanılan Yaklaşım	Kullanılan Veri Seti	Çalışmalarda Kullanılan Saldırı Sınıfları / Odaklanılan Saldırı
Hibrit Öğrenme & Diğer	Akbıyık (2024)	2024	Karar Ağaçları, ADABoost, YSA, DVM	KDD99	Genel saldırı tespiti
	Al-Daffaie (2024)	2024	Quad-RNN (Dört Yönlü Tekrarlayan Sinir Ağı)	NSL-KDD ve DDoS özel veri seti	DDoS ve genel siber saldırılar
	Çetin (2025)	2025	Oylama (Voting) ve Yığınlama (Stacking) Hibrit Modelleri	CICIoT2023	IoT ekosistemine yönelik siber saldırılar
	Goun (2024)	2024	RF, Ekstra Ağaçlar, CNN, LSTM, Hibrit CNN-LSTM	Akıllı Şebeke veri seti (FDI)	Yanlış Veri Enjeksiyonu (FDI) saldırıları
	Özalp (2023)	2023	CNN/LSTM + XGBoost/LightGBM (hibrit)	NSL-KDD ve CIC- IDS2018	14 farklı saldırı türü
	Delplace, Hermoso ve Anandita (2020)	2020	LR, SVM, RF, Gradyan Artırma, DNN	CTU-13	13 farklı botnet senaryosu
	Jullian vd. (2023)	2023	FFNN ve LSTM	NSL-KDD ve BoT-IoT	IoT ağ saldırıları
	AlShahrani ve Quasim (2021)	2021	AdaBoost Regresyon Sınıflandırıcısı (ABRC)	CICIDS 2019	Genel siber saldırılar
	Ferrag vd. (2019)	2019	DNN, RNN, CNN, DBN, Derin Otomatik Kodlayıcılar	CSE-CIC-IDS 2018	Genel siber saldırılar
	Alabdulatif (2025)	2025	CNN, LSTM, GRU Topluluk Modeli + SHAP	CICIDS2017	Genel siber saldırılar

Çizelge 2.1 Derin öğrenme yöntemleri kullanılarak siber saldırı tespiti yapan çalışmaların karşılaştırması (devamı)

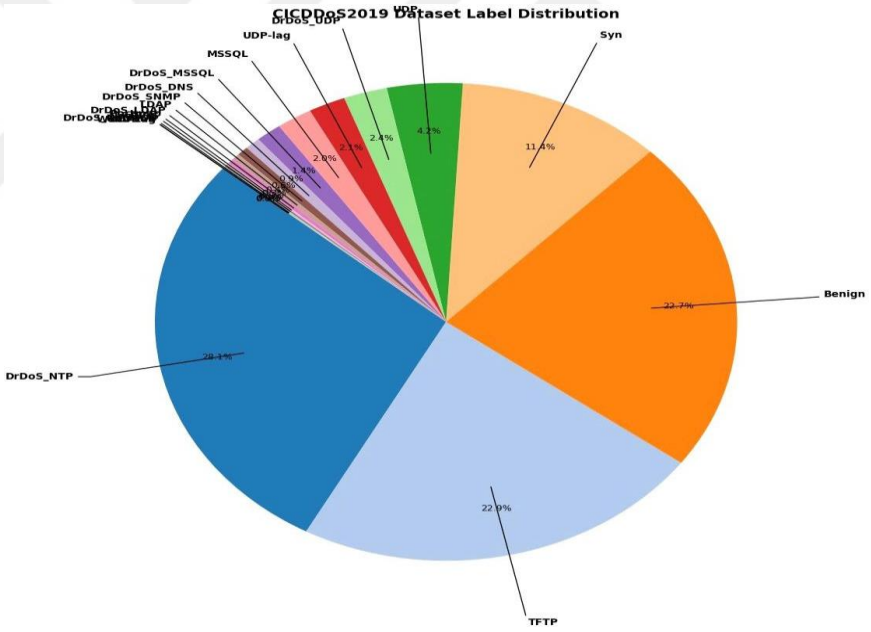
Denetimli Örnek Öğrenimi (CNN/DNN Ağırlıklı)	Chen vd. (2019)	2019	Çok Kanallı CNN (DAD-MCNN)	KDDCUP99 ve CICIDS2017	DDoS saldırıları
	Haider vd. (2020)	2020	Ensemble CNN	CICIDS2017	DDoS saldırıları (SDN)
	Wang ve Liu (2020)	2020	Bilgi Entropisi + CNN	CICIDS2017	DDoS saldırıları (SDN)
	Asad vd. (2020)	2020	DNN (DeepDetect)	CICIDS2017	Uygulama Katmanı DDoS
	Muraleedharan ve Janet (2020)	2020	DNN (FC)	CICIDS2017	HTTP yavaş DoS saldırıları
	de Assis vd. (2020)	2020	CNN	Simüle SDN ve CICDDoS2019	DDoS saldırıları
	Hussain vd. (2020)	2020	ResNet-18 (CNN)	CICDDoS2019	DoS/DDoS saldırıları
	Amaizu vd. (2021)	2021	DNN + PCC	CICDDoS2019	DDoS saldırıları (5G/B5G)
	Cil vd. (2021)	2021	DNN	CICDDoS2019	DDoS saldırıları
	Sbai ve El Boukhari (2020)	2020	DNN	CICDDoS2019	UDP flooding saldırısı
	Sabeel vd. (2019)	2019	DNN ve LSTM	CICIDS2017 ve ANTS2019	Bilinmeyen DoS/DDoS saldırıları
	Shurman vd. (2020)	2020	LSTM	CICDDoS2019	DrDoS saldırıları
Denetimli Dizi Öğrenimi (RNN/LSTM/GRU)	Assis vd. (2021)	2021	GRU	CICDDoS2019 ve CICIDS2018	DDoS ve sızma saldırıları
Yarı Denetimli Öğrenme	Kasim (2020)	2020	AE-SVM	CICIDS2017, NSL-KDD, özel veri	DDoS saldırıları
	Bhardwaj vd. (2020)	2020	İstiflenmiş Seyrek AE + DNN	NSL-KDD ve CICIDS2017	Ağ trafiği sınıflandırması / DDoS
Hibrit Derin Öğrenme	Roopak vd. (2019)	2019	MLP, CNN, LSTM, Hibrit CNN- LSTM	CICIDS2017	DDoS saldırıları
	Roopak vd. (2020)	2020	NSGA + CNN-LSTM	CICIDS2017	DDoS saldırıları
	Elsayed vd. (2020)	2020	RNN-AE (DDoSNet)	CICDDoS2019	DDoS saldırıları

3. MATERYAL VE YÖNTEM

Dağıtılmış Hizmet Aksatma (DDoS) saldırıları, hedef sistemleri yoğun sahte trafikle doldurarak hizmet veremez duruma getirmeyi amaçlayan ciddi siber saldırı türlerindendir. Özellikle SYN- Flood ve UDP Flood gibi yaygın DDoS teknikleri, basit yöntemlerle sunucuların kaynaklarını tüketerek büyük zararlara yol açabilir (Zargar et al., 2013). SYN-Flood saldırısı, TCP protokolünün üç yönlü el sıkışma (3-way handshake) süreçlerini istismar eden bir yöntemdir; saldırgan, hedefe sürekli yarım kalan SYN istekleri göndererek sunucunun bağlantı kuyruğunu doldurur ve kaynaklarını tüketir (Kumar et al., 2020). UDP-Flood ise hedef sistemdeki rastgele portlara aşırı sayıda UDP paketi yollayarak ağ bant genişliğini ve hedefin yanıt verme kapasitesini tüketir. Bu tür saldırılar günümüz ağ güvenliği için büyük bir tehdit oluşturmaktadır. Geleneksel güvenlik önlemleri (örn. firewall veya imza tabanlı IDS sistemleri) dinamik ve hacimli DDoS saldırılarını tespit etmede yetersiz kalabilmektedir (Peng et al., 2007). Bu proje kapsamında, DDoS saldırılarına karşı makine öğrenimi tabanlı bir yaklaşım geliştirilmesi hedeflenmiştir. Özellikle CIC-DDoS2019 adlı güncel bir veri seti kullanılarak, ağ trafiğinde SYN-Flood ve UDP-Flood saldırılarının otomatik olarak tespit edilmesi amaçlanmaktadır (Sharafaldin et al., 2019). Projede, yapay sinir ağlarından çok katmanlı bir algılayıcı (Multilayer Perceptron, MLP) modeli temel alınmış ve modelin başarımını arttırmak için Üretici Çekişmeli Ağ (Generative Adversarial Network, GAN) tabanlı sentetik veri üretimi yaklaşımı entegre edilmiştir (Zhao et al., 2021). Bu sayede dengesiz veri dağılımlarından kaynaklanan öğrenme sorunlarının giderilerek modelin saldırı tespit performansının iyileştirilmesi hedeflenmiştir. DDoS saldırılarını yüksek doğrulukla tespit edebilen bir sistem, özellikle gerçek zamanlı uygulamalarda servis sürekliliği ve güvenlik açısından kritik öneme sahiptir. Bu raporda, veri setinin hazırlanmasından model eğitimine ve sonuçların değerlendirilmesine dek projenin tüm adımları ayrıntılı olarak ele alınmakta; elde edilen bulgular grafikler ve metrikler eşliğinde tartışılmaktadır. Elde edilen sonuçlar, GAN destekli veri genişletme yönteminin, SYN-Flood ve UDP-Flood saldırılarının tespitinde MLP modelinibaşarısını belirgin şekilde arttırdığını göstermektedir.

3.1 Veri Seti

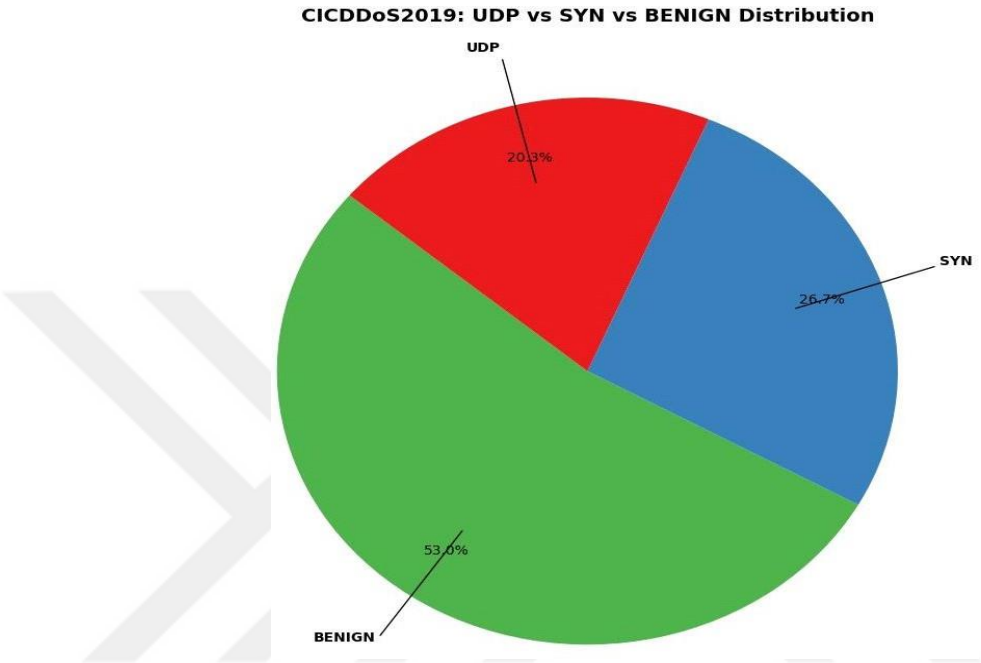
Çalışma kapsamında kullanılan veri seti, Canadian Institute for Cybersecurity (CIC) tarafından 2019 yılında yayınlanan CIC-DDoS2019 veri setidir. Bu veri seti, gerçek dünya ağ trafiğine benzer şekilde oluşturulmuş normal (benign) trafiğin ve çeşitli modern DDoS saldırılarını içerir. CIC-DDoS2019'da trafik verileri, ağ akışları (flow) halinde, PCAP formatında toplanmış ve daha sonra CICFlowMeter aracıyla her akış için zaman damgası, kaynak/hedef IP ve portlar, protokol türü gibi bilgilerin yanı sıra istatistiksel özellikler çıkarılarak etiketlenmiştirunb. Sonuç olarak her akış kaydı için 80'den fazla özellik elde edilmiş ve bu kayıtlar saldırı tipine veya normal trafiğe göre etiketlenmiştir. Veri seti, en güncel DDoS çeşitlerini kapsaması ve gerçekçi trafik paterni sunması açısından akademik çalışmalarda yaygın olarak kullanılmaktadır (Şekil 3.1).



Şekil 3.1 CIC-DDoS2019 veri kümesi etiket dağılımı

CIC-DDoS2019 pek çok farklı DDoS saldırı türünü barındırmaktadır. Bunlar arasında yansıtmalı (reflective) DDoS saldırıları (DNS amplifikasyon, NTP, LDAP, MSSQL, NetBIOS, SNMP, SSDP gibi) ve istismara dayalı doğrudan saldırılar (SYN flood, UDP flood, UDP-Lag vb.) bulunmaktadır. Bu çalışmada, veri setinden SYN-Flood, UDP-Flood ve BENIGN (normal) sınıflarına ait örnekler kullanılmıştır. BENIGN olarak

etiketlenen akışlar, herhangi bir saldırı içermeyen sıradan ağ trafiği örneklerini temsil etmektedir. (ör. HTTP, SMTP, FTP gibi protokollerdeki meşru kullanıcı etkinlikleri). SYN-Flood sınıfı, TCP SYN Flood saldırı trafiğini; UDP-Flood sınıfı ise UDP Flood saldırı trafiğini içermektedir (Şekil 3.2).



Şekil 3.2 UDP, SYN ve BENIGN dağılımı

Veri setindeki bu üç sınıfın örnek sayıları incelendiğinde ciddi bir sınıf dengesizliği (imbalance) sorunu göze çarpmaktadır. Resmi istatistiklere göre CIC-DDoS2019 içinde Benign toplam 97,831 akış ile en fazla sayıda örneğe sahipken, SYN-Flood saldırısı 49,373 akışta gözlemlenmiştir. UDP- Flood saldırısı ise sadece 18,090 akış ile veri setinde oldukça azınlıkta kalmaktadır. Bu durum, saldırı ve normal trafik sınıfları arasında aşırı dengesizlik yaratarak makine öğrenimi modelinin eğitimini zorlaştırmaktadır. Dengesiz veri, modelin büyük sınıfa (SYN-Flood) aşırı uyum sağlayıp küçük sınıfları (özellikle normal trafiği) ihmal etmesine yol açabilir (Buda et al., 2018; He & Garcia, 2009). Projemizde bu zorluğu gidermek için veri setinin yalnızca ilgili alt sınıfları ayrıştırılmış ve daha sonra GAN tabanlı veri artırma yöntemleriyle denge sağlanmıştır.

Veri seti deney düzeneği olarak iki farklı günde toplanan trafik kayıtlarından oluştuğu için, çalışmamızda eğitim ve test verileri ayrılırken bu zaman bilgisini de göz önünde bulundurduk. Orijinal veri setinden belirlenen oranda (örneğin %80 eğitim, %20 test olacak şekilde) rasgele örnekleme yöntemiyle veri ayrımı yapılmıştır. Bu ayırım sırasında, eğitim verisi içinde her üç sınıfın da temsili olduğundan emin olunmuş, test verisi ise modelin henüz görmediği akışlardan oluşturulmuştur. Eğitim verisindeki ağır dengesizlik ilk etapta korunmuş ve bu taban senaryo ile bir MLP modeli eğitilmiştir. Daha sonra, bu eğitim verisi GAN kullanılarak sentetik örneklerle genişletilmiş ve aynı test verisi üzerinde iki modelin performansı karşılaştırılmıştır. Bu şekilde, veri setinin gerçekçi yapısı (saldırıların baskın olması durumu) test aşamasında korunurken, eğitim aşamasında dengesizliğin etkileri azaltılarak modelin minor sınıflardaki performansının arttırılması hedeflenmiştir (Çizelge 3.1).

Çizelge 3.1 CIC-DDoS2019 veri setinde kullanılan sınıflar ve akış sayıları

Sınıf Türü	**Akış Türü Açıklaması**	**Örnek Sayısı**
BENIGN	Normal (zararsız) ağ trafiği	97.831
SYN Flood	TCP tabanlı SYN Flood DDoS saldırısı	49.373
UDP Flood	UDP tabanlı hacimsel DDoS saldırısı	18.090
Toplam	—	**165.294**

3.2 Veri Seti Özellikleri

Bu çalışmada kullanılan veri kümesi, ağ trafiğine ait ham verilerden elde edilmiş olup normal ve saldırı trafiğini temsil eden çok sayıda örnek içermektedir. Veri kümesi, farklı saldırı türleri ve senaryolarını kapsayacak şekilde hazırlanmış ve makine öğrenmesi ile derin öğrenme tabanlı saldırı tespit çalışmalarına uygun bir yapı sunmaktadır. Toplamda 80 adet öznitelikten oluşan veri kümesi, ağ akışlarına ait istatistiksel ve zamansal bilgileri içermektedir. Veriler, günlük bazda ayrılmış CSV dosyaları hâlinde sunulmakta olup ön işleme, özellik seçimi ve sınıflandırma aşamalarında etkin bir şekilde kullanılabilmektedir. Bu özellikler sayesinde veri kümesi, DDoS saldırılarının tespiti ve

analizi için kapsamlı ve güvenilir bir kaynak oluşturmaktadır. Bu veri kümesine ait özellikler ve açıklamaları Çizelge 3.2'de gösterilmektedir.

Çizelge 3.2 Veri setinde bulunan özellikler

No	Özellik Adı	Özellik Açıklaması
1	**Protocol**	Taşıma katmanı protokol numarası
2	**Flow Duration**	Akışın mikro saniye cinsinden süresi
3	**Total Fwd Packets**	İleri yöndeki toplam paket sayısı
4	**Total Backward Packets**	Geri yöndeki toplam paket sayısı
5	**Fwd Packets Length Total**	İleri yöndeki paketlerin toplam boyutu
6	**Bwd Packets Length Total**	Geri yöndeki paketlerin toplam boyutu
7	**Fwd Packet Length Max**	İleri yöndeki maksimum paket boyutu
8	**Fwd Packet Length Min**	İleri yöndeki minimum paket boyutu
9	**Fwd Packet Length Mean**	İleri yöndeki ortalama paket boyutu
10	**Fwd Packet Length Std**	İleri yöndeki paket boyutlarının standart sapması
11	**Bwd Packet Length Max**	Geri yöndeki maksimum paket boyutu
12	**Bwd Packet Length Min**	Geri yöndeki minimum paket boyutu
13	**Bwd Packet Length Mean**	Geri yöndeki ortalama paket boyutu
14	**Bwd Packet Length Std**	Geri yöndeki paket boyutlarının standart sapması
15	**Flow Bytes/s**	Saniye başına aktarılan bayt sayısı
16	**Flow Packets/s**	Saniye başına aktarılan paket sayısı
17	**Flow IAT Mean**	Paketler arası ortalama varış süresi
18	**Flow IAT Std**	Paketler arası varış süresinin standart sapması
19	**Flow IAT Max**	Maksimum paketler arası varış süresi

Çizelge 3.2 Veri setinde bulunan özellikler (devamı)

No	Özellik Adı	Özellik Açıklaması
20	**Flow IAT Min**	Minimum paketler arası varış süresi
21	**Fwd IAT Total**	İleri yöndeki toplam paketler arası süre
22	**Fwd IAT Mean**	İleri yöndeki paketler arası ortalama süre
23	**Fwd IAT Std**	İleri yöndeki paketler arası sürenin standart sapması
24	**Fwd IAT Max**	İleri yöndeki maksimum paketler arası süre
25	**Fwd IAT Min**	İleri yöndeki minimum paketler arası süre
26	**Bwd IAT Total**	Geri yöndeki toplam paketler arası süre
27	**Bwd IAT Mean**	Geri yöndeki paketler arası ortalama süre
28	**Bwd IAT Std**	Geri yöndeki paketler arası sürenin standart sapması
29	**Bwd IAT Max**	Geri yöndeki maksimum paketler arası süre
30	**Bwd IAT Min**	Geri yöndeki minimum paketler arası süre
31	**Fwd PSH Flags**	İleri yöndeki PSH bayrağı sayısı
32	**Bwd PSH Flags**	Geri yöndeki PSH bayrağı sayısı
33	**Fwd URG Flags**	İleri yöndeki URG bayrağı sayısı
34	**Bwd URG Flags**	Geri yöndeki URG bayrağı sayısı
35	**Fwd Header Length**	İleri yöndeki başlıkların toplam uzunluğu
36	**Bwd Header Length**	Geri yöndeki başlıkların toplam uzunluğu
37	**Fwd Packets/s**	İleri yönde saniye başına paket sayısı
38	**Bwd Packets/s**	Geri yönde saniye başına paket sayısı
39	**Packet Length Min**	Akıştaki minimum paket boyutu
40	**Packet Length Max**	Akıştaki maksimum paket boyutu

Çizelge 3.2 Veri setinde bulunan özellikler (devamı)

No	Özellik Adı	Özellik Açıklaması
41	**Packet Length Mean**	Akıştaki ortalama paket boyutu
42	**Packet Length Std**	Paket boyutlarının standart sapması
43	**Packet Length Variance**	Paket boyutlarının varyansı
44	**FIN Flag Count**	FIN bayrağı sayısı
45	**SYN Flag Count**	SYN bayrağı sayısı
46	**RST Flag Count**	RST bayrağı sayısı
47	**PSH Flag Count**	PSH bayrağı sayısı
48	**ACK Flag Count**	ACK bayrağı sayısı
49	**URG Flag Count**	URG bayrağı sayısı
50	**CWE Flag Count**	CWE bayrağı sayısı
51	**ECE Flag Count**	ECE bayrağı sayısı
52	**Down/Up Ratio**	İndirme/Yükleme paket oranı
53	**Avg Packet Size**	Ortalama paket boyutu
54	**Avg Fwd Segment Size**	Ortalama ileri segment boyutu
55	**Avg Bwd Segment Size**	Ortalama geri segment boyutu
56	**Fwd Avg Bytes/Bulk**	İleri yönde bulk başına ortalama bayt
57	**Fwd Avg Packets/Bulk**	İleri yönde bulk başına ortalama paket
58	**Fwd Avg Bulk Rate**	İleri yönde ortalama bulk aktarım hızı
59	**Bwd Avg Bytes/Bulk**	Geri yönde bulk başına ortalama bayt
60	**Bwd Avg Packets/Bulk**	Geri yönde bulk başına ortalama paket
61	**Bwd Avg Bulk Rate**	Geri yönde ortalama bulk aktarım hızı

Çizelge 3.2 Veri setinde bulunan özellikler (devamı)

No	Özellik Adı	Özellik Açıklaması
62	**Subflow Fwd Packets**	Alt akıştaki ileri paket sayısı
63	**Subflow Fwd Bytes**	Alt akıştaki ileri bayt sayısı
64	**Subflow Bwd Packets**	Alt akıştaki geri paket sayısı
65	**Subflow Bwd Bytes**	Alt akıştaki geri bayt sayısı
66	**Init Fwd Win Bytes**	Başlangıç ileri pencere bayt sayısı
67	**Init Bwd Win Bytes**	Başlangıç geri pencere bayt sayısı
68	**Fwd Act Data Packets**	Veri içeren ileri paket sayısı
69	**Fwd Seg Size Min**	Minimum ileri segment boyutu
70	**Active Mean**	Akışın ortalama aktif süresi
71	**Active Std**	Aktif sürenin standart sapması
72	**Active Max**	Maksimum aktif süre
73	**Active Min**	Minimum aktif süre
74	**Idle Mean**	Akışın ortalama boşta kalma süresi
75	**Idle Std**	Boşta kalma süresinin standart sapması
76	**Idle Max**	Maksimum boşta kalma süresi
77	**Idle Min**	Minimum boşta kalma süresi

3.3 DDoS Saldırı ve Özellik Seçimi

CIC-DDOS2019 veri kümesi, farklı saldırı türleri içermektedir. Bu başlık altında DDoS saldırı türleri ve önemli değişkenler ayrıntılı olarak incelenecektir (Çizelge 3.3).

Çizelge 3.3 Veri setinde kulanilan özellikler

Saldırı Türü	**Kullanılan Özellikler**	**Gerekçe**
UDP Flood	Flow Bytes/s, Flow Packets/s, Packet Length Mean, Packet Length Std, Fwd Packets/s, Bwd Packets/s, Down/Up Ratio, Avg Packet Size	UDP flood saldırılarında yüksek paket oranı, sabit paket boyutları ve tek yönlü trafik baskındır.
SYN Flood	SYN Flag Count, ACK Flag Count, RST Flag Count, Flow IAT Mean, Flow IAT Std, Init Fwd Win Bytes, Fwd Packets/s, Flow Packets/s	SYN flood saldırıları, tamamlanmamış TCP oturumları ve anormal SYN bayrağı yoğunluğu ile karakterizedir.
BENIGN (Normal Trafik)	Flow Duration, Total Fwd Packets, Total Bwd Packets, Flow IAT Mean, Packet Length Mean, Active Mean, Idle Mean	Normal trafik dengeli ileri/geri paket oranı ve düzenli zamanlama gösterir.

3.4 Ön İşleme

Ham veri seti üzerinde modellemeye geçmeden önce çeşitli ön **işleme adımları uygulanmıştır.**

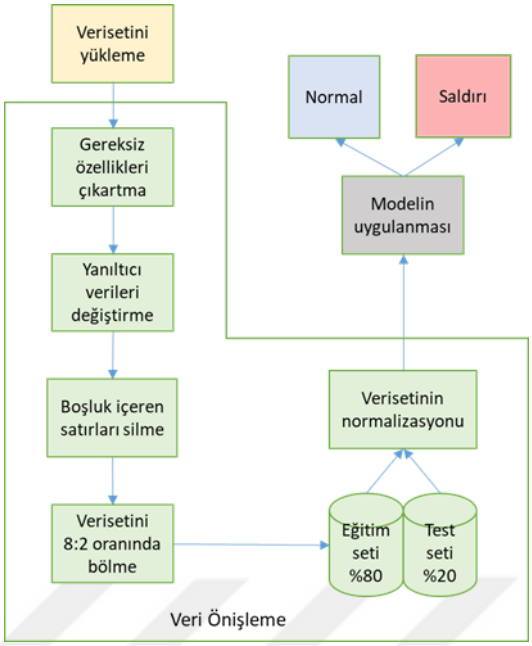
3.4.1 Veri temizleme

Aşamasında eksik veya tutarsız kayıtlar kontrol edilmiştir. CIC-DDoS2019 akış verilerinde bazı akışların belirli alanlarında (örneğin paket sayısı, bayt sayısı gibi) tutarsızlıklar veya null değerler bulunabilmektedir. Bu tür eksik kayıtlar veri setinden çıkarılmış veya uygun şekilde doldurulmuştur. Ayrıca, saldırı etiketlerinin doğruluğu ve tutarlılığı gözden geçirilmiştir. Veri seti hâlihazırda etiketli geldiği için elle etiketleme gerekmemiş, ancak sadece projenin odaklandığı üç sınıf (BENIGN, SYN, UDP) dışındaki kayıtlar eğitim kümesinden filtrelenerek çıkarılmıştır. Böylece modelin eğitimi, ilgisiz saldırı türlerinden arındırılmış, sadece ilgili sınıflara odaklanmıştır. Veri temizliğinin

ardından, özellik seçimi ve dönüştürme işlemleri uygulanmıştır. CIC-DDoS2019 verisinde her akış kaydı kimlik belirleyici niteliğinde bazı alanlar (Flow ID, kaynak IP, hedef IP gibi) içerebilir. Bu alanlar saldırı tespitinde doğrudan anlamlı olmayan veya genelleştirilemeyen bilgiler taşıdığı için model eğitimine dâhil edilmemiştir. Bunun yerine, daha çok akışın istatistiksel özellikleri (paket sayıları, bayt oranları, süre, bayt/paket hızları vb.) ve protokol bilgileri gibi anlamlı özellikler kullanılmıştır. Veri setindeki kategorik özellikler (örneğin protokol ismi) gerekiyorsa sayısal değere dönüştürülmüş, ancak çoğunlukla protokol tipi gibi bilgiler zaten sayısal kodlarla temsil edildiğinden ek işlem gerekmemiştir. Sonuçta modelin girdi özelliği olarak kullanılmak üzere her akış için seçilen özellikler bir matris yapısında hazır hale getirilmiştir. . Daha sonra veriseti %80 eğitim ve %20 test veriseti olmak üzere ikiye ayrılmıştır. Bütün bu ön işleme aşamaları tamamlandıktan sonra eğitim veriseti modelin eğitimine hazır hale gelmiştir.

3.4.2 Normalizasyon (ölçekleme)

Farklı ölçeklerdeki özellik değerlerini belirli bir aralığa getirerek modelin eğitimini iyileştirmek amacıyla uygulanmıştır. Örneğin, akış uzunluğu (bayt) veya paket sayısı gibi özellikler çok büyük değerlere sahip olabilirken, oran veya bayt/saniye gibi özellikler daha küçük ölçekli değerler içerebilir. Bu farklı ölçekleri dengelemek için Min-Max normalizasyonu kullanılmış ve tüm özellikler 0-1 Aralığına ölçeklenmiştir. Normalizasyon sayesinde MLP modelinin ağırlıkları, büyük değerli özellikler tarafından domine edilmeden, tüm girdilere eşit bir öneme sahip olacak şekilde öğrenebilmektedir. Son olarak, veriler eğitim ve test kümelerine ayrılmış ve karıştırılmıştır. Eğitim verisi üzerinden model eğitimi yapılırken, test kümesi eğitim sürecinde hiç görülmeden bir kenarda tutulmuştur. Bu sayede, eğitilen modellerin performansı gerçek veriyle (görülmemiş kayıtlarla) değerlendirilebilmiştir. Özellikle dengesiz veri problemine çözüm getirmek için, ilerleyen aşamada eğitim verisi üzerinde GAN ile sentetik örnek üretileceğinden, bu üretim işlemi yalnızca eğitim verisi üzerinde gerçekleştirilmiş; test kümesi orijinal dağılımı ile bırakılmıştır. Böylece modelin geliştirilmesinde kullanılan yapay verilerin, bağımsız test değerlendirmesini etkilemesinin önüne geçilmiştir (Zhao et al., 2021) (Şekil 3.3).



Şekil 3.3 DDoS saldırı tespit mimarisi

3.5 GAN Yapısı ve Amacı

Ağ trafiği saldırı tespit sistemlerinde karşılaşılan en temel problemlerden biri, veri kümelerinin ciddi ölçüde sınıf dengesizliği içermesidir. Gerçek dünya senaryolarında benign (normal) ağ trafiği büyük çoğunluğu oluştururken, UDP ve SYN gibi saldırı türleri oldukça sınırlı sayıda gözlemlenmektedir. Bu dengesiz yapı, makine öğrenmesi ve derin öğrenme tabanlı modellerin çoğunluk sınıfa yönelmesine neden olmakta ve azınlık sınıflarının doğru şekilde öğrenilmesini zorlaştırmaktadır.

Bu çalışmada, ağ trafiği tabanlı saldırı tespit sistemlerinde sıklıkla karşılaşılan sınıf dengesizliği problemini ele almak amacıyla Generative Adversarial Networks (GAN) tabanlı bir veri artırma yaklaşımı önerilmektedir. Kullanılan veri kümesi üç sınıftan oluşmaktadır: normal (benign) trafik, UDP tabanlı saldırı trafiği ve SYN tabanlı saldırı trafiği. Veri kümesinde benign trafik 97.831 örnek ile çoğunluk sınıfını oluştururken, UDP ve SYN saldırı sınıfları sırasıyla 18.090 ve 49.373 örnek içermektedir. Bu dağılım, özellikle UDP saldırı sınıfının ciddi biçimde azınlıkta olduğunu göstermekte ve bu durum, sınıflandırma modellerinin saldırı trafiğini doğru şekilde öğrenmesini zorlaştırmaktadır. Sınıf dengesizliği oranı aşağıdaki gibi hesaplanmaktadır.

$$IR = \frac{N_{benign}}{N_{UDP}} = \frac{97{,}831}{18{,}090} \approx 5.4$$

Bu tür bir dengesizlik, geleneksel derin öğrenme algoritmalarının çoğunluk sınıfına eğilim göstermesine, dolayısıyla saldırı sınıflarında yüksek yanlış negatif (FN) oranlarına yol açmaktadı. Bu problemi gidermek amacıyla çalışmada GAN tabanlı sentetik veri üretimi gerçekleştirilmiştir. GAN mimarisi iki temel bileşenden oluşmaktadır: üreteç (Generator, G) ve ayırtedici (Discriminator, D). Üreteç ağ, rastgele bir gürültü vektörü z~pz(z) alarak UDP ve SYN saldırı trafiğine benzer sentetik örnekler üretmeyi amaçlarken; ayırtedici ağ, gerçek veri x~pdata(x) ile üretilen sentetik verileri ayırt etmeye çalışmaktadır. Bu iki ağ, karşıt (adversarial) bir öğrenme süreci içerisinde birlikte eğitilmektedir. GAN'in optimizasyon süreci aşağıdaki min–max fonksiyonu ile tanımlanmaktadır:

$$\min_G \max_D V(G,D) = \mathbb{E}_{x\sim p_{data}}[\log D(x)] + \mathbb{E}_{z\sim p_z}[\log(1 - D(G(z)))]$$

Teorik olarak, eğitim süreci Nash dengesi noktasına ulaştığında, üretilen veri dağılımı pg(x) gerçek veri dağılımı pdata(x) ile örtüşmektedir.

3.6 GAN Mimarisi ve Hiperparametreler

Bu çalışmada kullanılan GAN mimarisi tam bağlantılı (Fully Connected) katmanlardan oluşmaktadır. Üreteç ağı, 100 boyutlu rastgele gürültü vektörünü giriş olarak almakta ve sırasıyla 256, 512 ve 1024 nöronlu gizli katmanlar aracılığıyla sentetik saldırı verisi üretmektedir. Gizli katmanlarda ReLU aktivasyon fonksiyonu, çıkış katmanında ise Tanh aktivasyon fonksiyonu kullanılmıştır. Ayırt edici ağ ise girişte gerçek veya sentetik ağ trafiği örneklerini almakta ve 1024, 512 ve 256 nöronlu katmanlar aracılığıyla bu örnekleri sınıflandırmaktadır. Ayırt edici ağda Leaky ReLU aktivasyonu ve çıkış katmanında Sigmoid fonksiyonu tercih edilmiştir (Çizelge 3.4).

Çizelge 3.4 GAN modeli eğitim hiperparametreleri

Parametre	**Değer**
Öğrenme Oranı (Learning Rate)	0.0002
Batch Size	64
Epoch Sayısı	200
Optimizasyon Algoritması	Adam
beta_1, beta_2 Değerleri	0.5, 0.999
Gürültü Vektörü Boyutu	100

3.7 Performans Metrikleri

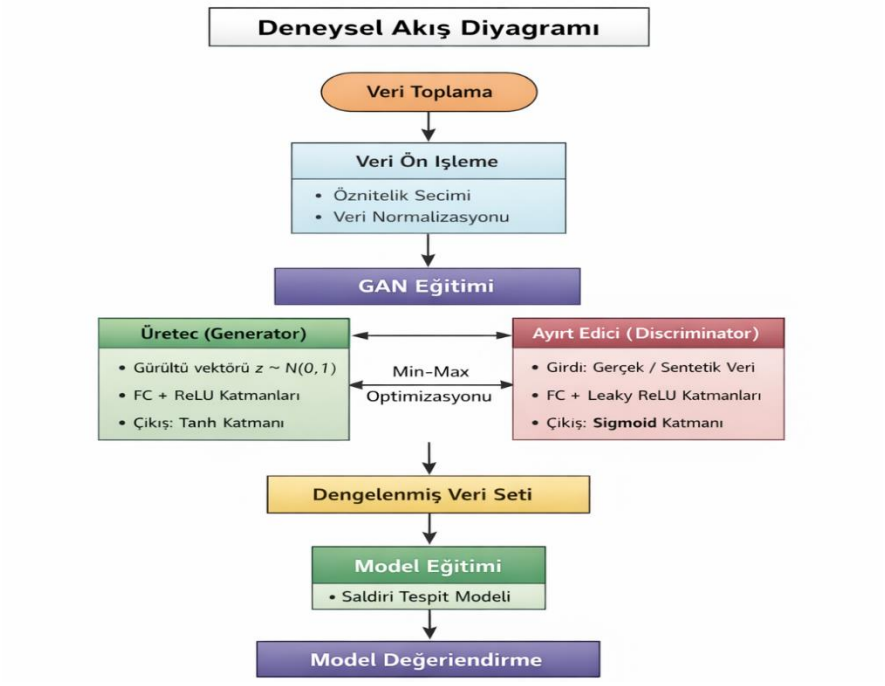
Model performansı; doğruluk (Accuracy), kesinlik (Precision), duyarlılık (Recall) ve F1-skoru metrikleri kullanılarak değerlendirilmiştir. Özellikle saldırı tespit sistemlerinde kritik öneme sahip olan duyarlılık (Recall) metriği aşağıdaki şekilde tanımlanmaktadır:

GAN ile dengelenmiş veri kümesi kullanılarak eğitilen modellerde, özellikle UDP ve SYN saldırı sınıflarında duyarlılık ve F1-skor değerlerinde belirgin artışlar gözlemlenmiştir. Bu durum, GAN tabanlı veri artırma yönteminin azınlık sınıflarına ait örüntüleri daha etkili bir şekilde öğrenebildiğini göstermektedir (Çizelge 3.5).

Çizelge 3.5 GAN ile veri dengeleme öncesi ve sonrası dağılım

Sınıf Türü	**Akış Türü Açıklaması**	**GAN Öncesi Akış Sayısı**	**GAN Sonrası Akış Sayısı**
BENIGN	Normal (zararsız) ağ trafiği	97.831	97.831
SYN Flood	TCP tabanlı SYN Flood DDoS saldırısı	49.373	97.831
UDP Flood	UDP tabanlı hacimsel DDoS saldırısı	18.090	97.831
Toplam	—	**165.294**	**293.493**

Elde edilen deneysel sonuçlar, GAN tabanlı veri artırma yaklaşımının ağ trafiği saldırı tespitinde sınıf dengesizliği problemini etkili bir şekilde giderdiğini ve özellikle azınlık saldırı sınıflarının tespit performansını önemli ölçüde artırdığını göstermektedir. Bu çalışma, GAN kullanımının gerçek zamanlı saldırı tespit sistemleri ve siber güvenlik uygulamaları için güçlü ve uygulanabilir bir çözüm sunduğunu ortaya koymaktadır (Şekil 3.4).



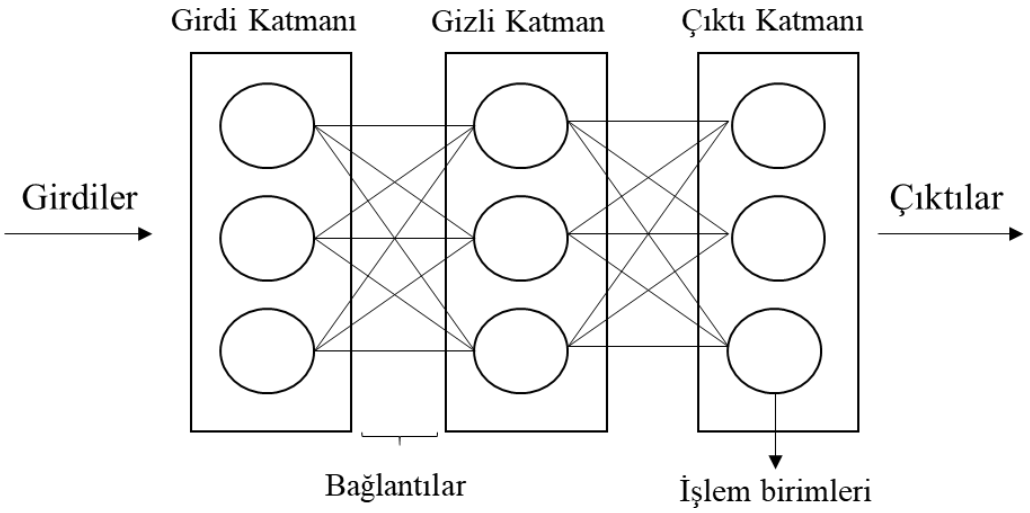
Şekil 3.4 GAN yapısı

3.8 Derin Öğrenme Mimarisi

Bu bölümde önerilen model hakkında altyapı oluşturmak için modelde kullanılan özelliklerden, matematiksel arka planından ve derin öğrenme modelinin genel yapısından bahsedilmektedir.

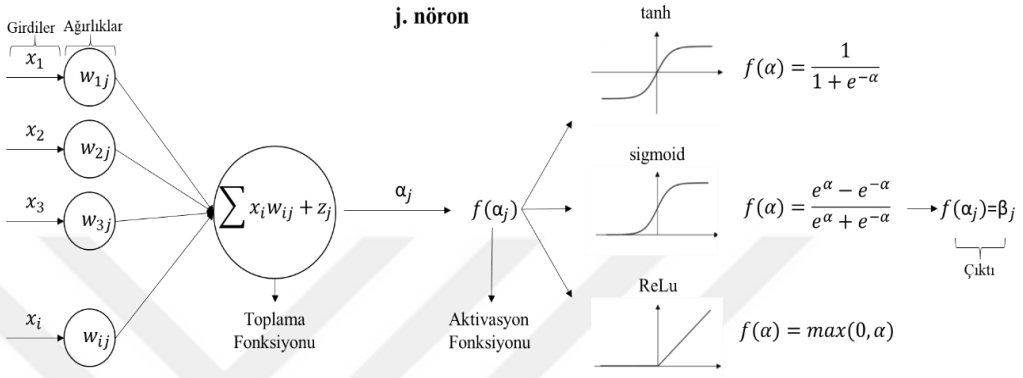
3.8.1 Yapay sinir ağları

Biyolojik sinir ağlarından esinlenen yapay sinir ağları (YSA), çok sayıda ara bağlantıya sahip basit işlem birimlerinden oluşan paralel hesaplama sistemleridir. Yapay bir nöron, doğal nöronlardan esinlenen hesaplamalı bir modeldir. Ağda yer alan düğümler "yapay nöronlar" olarak tanımlanan işlem birimleridir ve bunlar YSA'ları oluşturmaktadır. YSA'ların işlevi bilgiyi işlemek olduğu için ağırlıklı olarak örüntü tanıma, tahmin ve veri sıkıştırma gibi mühendislik amaçları için kullanılmaktadır. YSA'lar genel olarak girdi katmanı, gizli katman ve çıktı katmanı olmak üzere 3 katmandan oluşmaktadır. Şekil 3.5'de YSA'ların genel yapısı gösterilmektedir.



Şekil 3.5 Bir YSA'nın genel yapısı

Bir YSA modelinde verilerin girişi, girdi verisetinin girdi katmanıyla işleme sokulmasıyla ile başlamaktadır. Girdi katmanından gelen veriler işlem birimlerinde yer alan hesaplamalarla işlenmek üzere gizli katmana iletilmektedir. Daha sonra gizli katmandan gelen veriler çıktı katmanında işlenerek çıktı veriseti elde edilmektedir. Şekil 3.6'de işlem birimleri olan bir yapay nöronun işleyişi ile ilgili yapıya yer verilmiştir.



Şekil 3.6 Bir yapay nöronun yapısı ve matematiksel formülasyonu

Şekil 3.6'da oklarla gösterilen ve bilgi akışını temsil eden değerlerine girdiler denmektedir. Bir nöronda girdilerle aktarılan bilginin önem ve hücreye olan etkisi *w* ile gösterilen ağırlıklar ile belirlenmektedir. Ağırlık ne kadar yüksek değere sahipse, girdinin etkisinin güçlü olacağı belirlenmektedir. Giriş nöronlarının yalnızca bir girişi olduğundan, çıktıları aldıkları girdinin bir ağırlıkla çarpımı olacaktır. Girdiler bir nöronda yer alan ağırlıklarla çarpıldıktan sonra her katmanda bulunun ön yargılar (*z*) eklenir ve toplama fonksiyonu kullanılarak α ile gösterilen net girdi hesaplanır. Bu net girdi nöronun matematiksel bir fonksiyonu olan aktivasyon fonksiyonuyla işleme girmektedir. Sonuç olarak β ile gösterilen çıktı elde edilmiştir. YSA'nın her gizli katmanın düğümünü temsil eden gi fonksiyonunun matematiksel yapısı aktivasyon fonksiyonu *f* ile gösterilecek şekilde aşağıdaki gibi verilmiştir:

$$gi(\boldsymbol{x}) = f(\boldsymbol{x}^T\boldsymbol{w}+z) \tag{3.1}$$

Yapay bir nöronun ağırlıklarını ayarlayarak, belirli girdiler için istenilen çıktı elde edebilmektedir. Fakat yüzlerce veya binlerce nörondan oluşan bir YSA olduğunda, gerekli tüm ağırlıkları elle bulmak oldukça karmaşık olmaktadır. Ancak ağdan istenen çıktıyı elde etmek için yardımcı algoritmalar aracılığıyla YSA'nın ağırlıkları ayarlanabilmektedir. Bu ağırlık ayarlama sürecine öğrenme veya eğitim denmektedir. Eğitim rastgele ağırlıklarla başlar ve hataların minimum düzeyde olacak şekilde ayarlanmasını amaçlamaktadır.

3.8.2 Çok katmanlı ağ yapısı (MLP)

Bu tezde önerilen derin öğrenme modeli, YSA'nın derin sinir ağlarında kullanımı yaygın olan çok katmanlı algılayıcılar (MLP) yapısına dayanmaktadır. Bu derin sinir ağı aynı zamanda İleri Beslemeli Sinir Ağı (FNN) adı verilen derin öğrenme modelidir. FNN, modelin çıktılarıyla kendisini beslemesini sağlayan geri bildirim bağlantılarına sahip olmaması yönünden Tekrarlayan Sinir Ağlarından (RNN) farklıdır. FNN, giriş katmanı, çıktı katmanı ve gizli katman olmak üzere en az üç katmandan oluşur. Sinir ağının yapısını oluşturan bu katmanlar, modelin derinliğini ifade etmekte ve derin öğrenmenin adı buradan gelmektedir. Ağdaki her gizli katman, vektör değerlerinden oluşur ve gizli katmanların boyutu, modelin genişliği olarak tanımlanmaktadır. FNN'nin gizli katmanlarını (3.1) nolu denklemden yola çıkarak iç içe geçmiş zincir formunda (3.2) nolu denklemdeki gibi gösterebiliriz:

$$y = gi(gi\text{-}1(\ldots(g1(\mathbf{x})))) \tag{3.2}$$

Gizli katmanların zincir şeklindeki bu yapısı derin sinir ağlarını oluşturmaktadır. FNN'nin (3.2) nolu matematiksel tanımında, her gizli katmanın düğümleri için hesaplanacak fonksiyon *gi*, gizli katman sayısı *i*, girdi vektörü *x*, çıktı vektörü *y* notasyonu ile ifade edilir.

3.8.3 İleri besleme

İleri besleme, yapay sinir ağlarında yer alan ilk ve en temel özelliktir. İleri besleme özelliğine sahip ağlarda bilgi akışı daime sırasıyla girdi katmanı, gizli katmanlar ve çıktı katmanı olacak şekilde ileri doğru ve tek yönlüdür.

3.8.4 Aktivasyon fonksiyonları

Bir nörondan çıktı elde edebilmek için kullanılan bir fonksiyondur. Aynı zamanda "Transfer Fonksiyonu" olarak da adlandırılır. "Var" veya "Yok" gibi ikili sonuçlar için oluşturulmuş sinir ağlarının çıktısını belirlemek için kullanılır. Seçilen aktivasyon fonksiyonuna göre "0 ile 1" veya "-1 ile 1" arasında değerler elde edilir. Problemin çeşidine göre çeşitli aktivasyon fonksiyonlar tercih edilebilmektedir. Yaygın kullanılan aktivasyon fonksiyonlarından bazıları; *Sigmoid*, *Tanh*, (Rectified Linear Unit) *ReLu, Softmax* fonksiyonları olarak gösterilebilir.

3.8.5 Kayıp fonksiyonu

Genellikle sinir ağlarında sonuçlardan elde edilen hata en aza indirgenmeye çalışılmaktadır [86]. Hatanın en aza indirgenmesi için kayıp fonksiyonları olarak tanımlanan fonksiyonlar geliştirilmiştir. "ortalama kare hata", "çapraz entropi" gibi fonksiyonlar derin öğrenme modellerinde yaygın olan bazı kayıp fonksiyonlarıdır.

3.8.6 Optimizasyon algoritması

Bir optimizasyon algoritması, optimum veya tatmin edici bir sonuç bulunana kadar çeşitli sonuçları karşılaştırarak yinelemeli olarak yürütülen bir prosedürdür. Optimizasyon, hataları azaltmak için sinir ağının ağırlık ve öğrenme hızı gibi özelliklerini değiştiren algoritmalar veya yöntemlerdir. Optimize ediciler, daha hızlı sonuç alınmasına yardımcı olmaktadır. "Stochastic Gradient Descent (SGD)", "Adagrad", "Root Mean Square Propagation (Rmsprop)", "Adaptive Moment Estimation (ADAM)" ve "Adaptive Max Pooling (ADAMAX)" derin öğrenme modellerinde sık kullanılan optimizasyon algoritmalarıdır.3.8.7 "mini-batch" boyutu Batch boyutu, derin öğrenme modelinde yer

alan parametreleri güncellemeden önce üzerinde çalışılması gereken eğitim verisetinden alınacak örneklerin sayısını tanımlayan bir hiperparametredir. Verisetinde yer alan bütün verileri aynı anda işlemek, zaman ve bellek bakımından maliyetli olduğu için girdinin parçalar halinde işlenmesine "mini- batch" denilmektedir. Model tasarımında aynı anda ne kadar verinin işleneceğini ifade eden değer mini-batch parametresidir.

3.8.8 "epoch" zamanı

Epoch(dönem), modelin eğitiminde belli sayıda parçalara bölünmüş verisetinin tamamını kaç kez gördüğünü ifade etmek için kullanılır. Dolayısıyla, algoritma verisetindeki bütün örnekleri her gördüğünde, bir dönem tamamlanmış olur.

3.8.9 Aktivasyon fonksiyonları

Modeller oluşturulurken Relu, Sigmoid, Hiperbolik Tanjant aktivasyon fonksiyonları yaygın kullanıldığı için tercih edilmiştir. Gizli katmanlarda kullanılan bu üç aktivasyon fonksiyonlarından ReLu fonksiyonu (3.3) nolu denklemde, sigmoid fonksiyonu (3.4) nolu denklemde ve hiperbolik tanjant aktivasyon fonksiyonu (3.5) nolu denklemde verilmiştir

3.8.10 Aktivasyon fonksiyonları

Modeller oluşturulurken Relu, Sigmoid, Hiperbolik Tanjant aktivasyon fonksiyonları yaygın kullanıldığı için tercih edilmiştir. Gizli katmanlarda kullanılan bu üç aktivasyon fonksiyonlarından ReLu fonksiyonu (3.3) nolu denklemde, sigmoid fonksiyonu (3.4) nolu denklemde ve hiperbolik tanjant aktivasyon fonksiyonu (3.5) nolu denklemde verilmiştir:

$$relu(x)=max(0,x) \tag{3.3}$$

$$sigmoid(x) = \frac{1}{e^{-x}+1} \tag{3.4}$$

$$tanh(x) = \frac{e^{2x}-1}{e^{2x}+1} \tag{3.5}$$

Softargmax veya normalleştirilmiş üstel fonksiyon olarak da bilinen softmax fonksiyonu lojistik fonksiyonun çoklu boyutlara genellemesidir. Önerilen DNN modellerinin çıktı katmanı için kullanılan softmax aktivasyon fonksiyonu (3.6) nolu denklem ile verilmiştir:

$$\text{softmax}(x_i) = \frac{e^{x_i}}{\sum_{j=1}^{n} e^{x_j}} \tag{3.6}$$

3.8.11 Kayıp fonksiyonu

Modellerin öğrenmesinin doğruluğunu artırmak için kayıp fonksiyonu olarak ikili çapraz entropi fonksiyonu kullanılmış ve aşağıdaki denklem ile verilmiştir:

$$q_0 = 1 - \hat{y},\ q_1 = \hat{y} \tag{3.7}$$

$$p_0 = 1 - y,\ p_1 = y \tag{3.8}$$

$$L = -\sum_i p_n \log q_n = -y \log \hat{y} - (1-y)\log(1-\hat{y}) \tag{3.9}$$

Burada sırasıyla L kayıp fonksiyonunu, log doğal logaritmayı, p_n gerçek olasılığı, y gerçek olasılık değerini, q_n tahmin edilen olasılığı, $\hat{y}$ tahmin edilen olasılık değerini, n sayısı {0,1} değerlerini göstermektedir.

3.8.12 Optimizasyon algoritmaları

DNN modellerinde kullanımı yaygın olan iki optimizasyon algoritması tercih edilmiştir. Birinci olarak kullanılan algoritma ADAM optimizasyon algoritmasıdır. SGD tabanlı bir optimizasyon yöntemi olan ADAM, yalnızca daha az bellek gereksinimi olan birinci dereceden gradyanlar gerekli olduğu için etkilidir. Spesifik uyarlanabilir öğrenme

oranları hesaplanarak, gradyanların birinci ve ikinci momentlerinin tahminlerinden farklı parametreler elde edilir. ADAM 'ın güncelleme formülü parametre şu şekilde verilir:

$$m_t = \Omega_1 m_{t-1} + (1 - \Omega_1) g_t \tag{3.10}$$

$$v_t = \Omega_2 v_{t-1} + (1 - \Omega_2)(g_t)^2 \tag{3.11}$$

$$m_t^{corrected} = \frac{m_t}{1-(\Omega_1)^t} \tag{3.12}$$

$$v_t^{corrected} = \frac{v_t}{1-(\Omega_2)^t} \tag{3.13}$$

$$W_t = W_{t-1} - \lambda \frac{m_t^{corrected}}{\sqrt{v_t^{corrected}} + \epsilon} \tag{3.14}$$

Burada g_t, t adımdaki gradyanları, m_t birinci moment gradyanları, v_t ikinci moment gradyanları, Ω_1, $\Omega_2 \in [0,1)$ üssel bozulma oranlarını, λ adım boyutu, ϵ sıfıra bölünmeyi önlemek için küçük bir değeri, W_t ağırlık vektörünü(güncellenecek parametre) temsil etmektedir. Makine öğrenmesi problemlerinde λ=0.001, Ω_1=0.9, Ω_2=0.999 *ve* ϵ=10^{-8} varsayılan değerlerinin kullanımı daha iyi sonuçlar vermektedir.

İkinci optimizasyon algoritması olarak "ADAMAX" kullanılmıştır. ADAMAX, ADAM optimizasyon algoritması kuralının sonsuzluk normu ile güncellenmesiyle eldeedilmiştir. ADAMAX 'ın optimizasyon algoritması için matematiksel formül şu şekilde verilmiştir:

$$m_t = \Omega_1 m_{t-1} + (1 - \Omega_1) g_t \tag{3.15}$$

$$u_t = \max(\Omega_2 u_{t-1}, |g_t|) \tag{3.16}$$

$$\theta_t = \theta_{t-1} - \left(\frac{\lambda}{1-\Omega_1^t}\right) \frac{m_t}{v_t} \tag{3.17}$$

Burada t zaman adımı olarak, g_t, t adımdaki gradyanlar olarak, m_t ilk moment vektörü olarak, u_t üssel ağırlıklı sonsuzluk normu olarak, Ω_1, $\Omega_2 \in [0,1)$ üssel bozulma oranları olarak, λ adım boyutu olarak, θ_t güncellenmiş parametre olarak gösterilmektedir.

Yukarıdaki (3.15), (3.16), (3.17) nolu denklemler ile verilen algoritmanın t=0 anındaki m_t'nin başlangıç değeri m_0=0 ve u_t'un başlangıç değeri u_0=0 olarak alınır. Bu algoritmada λ=0.002, Ω_1=0.9, Ω_2=0.999 varsayılan değerlerinin kullanımı problemlerin çözümünde daha uygundur.

3.9 GAN Olmadan MLP Eğitimi (Temel Model)

Bu çalışmada ağ trafiği saldırı tespiti problemi, ciddi sınıf dengesizliği içeren üç sınıflı bir veri kümesi üzerinde Çok Katmanlı Algılayıcı (Multi-Layer Perceptron – MLP) algoritması kullanılarak ele alınmıştır. Veri kümesinde Benign (normal trafik) sınıfı baskın durumda iken, UDP Flood ve SYN Flood saldırı sınıfları azınlık durumundadır. Bu tür dengesiz veri kümeleri, denetimli öğrenme algoritmalarının çoğunluk sınıfına yönelmesine neden olmakta ve saldırı sınıflarının doğru tespit edilmesini zorlaştırmaktadır. Bu nedenle model tasarımı ve eğitim süreci, sınıf dengesizliği göz önünde bulundurularak yapılandırılmıştır.

Kullanılan MLP mimarisi bir giriş katmanı, üç gizli katman ve bir çıkış katmanından oluşmaktadır. Gizli katmanlarda sırasıyla 128, 64 ve 32 nöron kullanılmıştır. Bu kademeli yapı, modelin giriş verilerindeki temel ağ trafiği örüntülerini ilk katmanda öğrenmesini, orta katmanda bu örüntüleri daha soyut temsillere dönüştürmesini ve son gizli katmanda ise sınıflar arasındaki ayırt edici farkları güçlendirmesini sağlamaktadır. Ancak sınıf dengesizliği nedeniyle, modelin yalnızca bu mimariyle eğitilmesi durumunda çoğunluk sınıfı olan Benign trafiğe aşırı uyum sağlama riski bulunmaktadır.

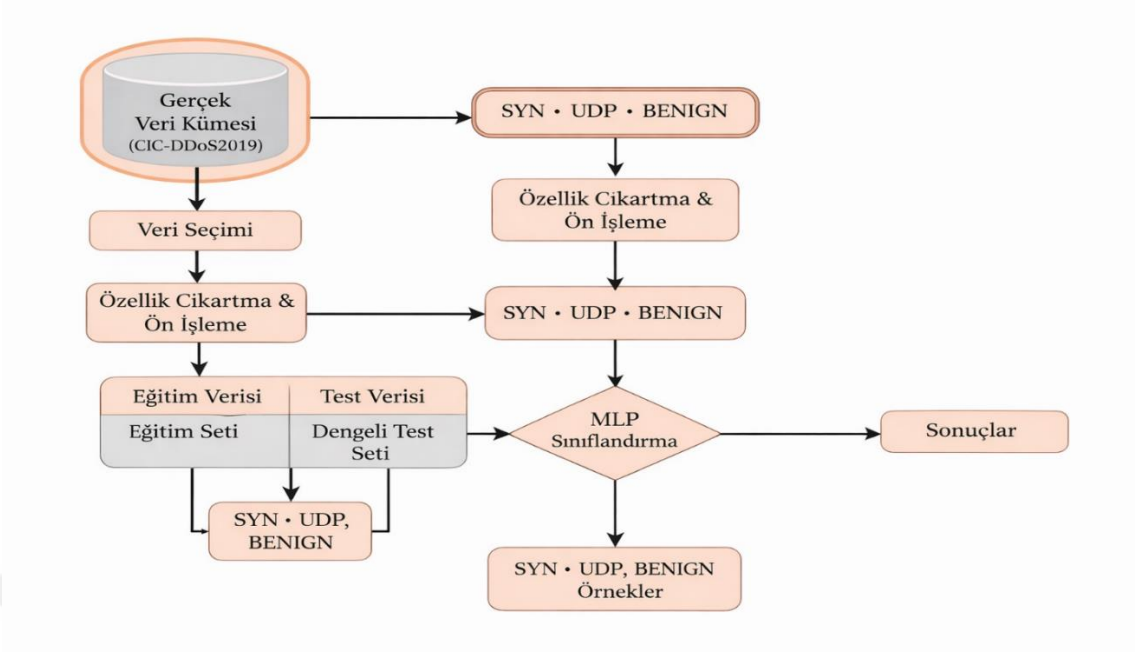
Gizli katmanlarda ReLU aktivasyon fonksiyonu kullanılmıştır. ReLU, doğrusal olmayan ilişkilerin öğrenilmesini sağlarken, derin ağlarda karşılaşılan gradyan sönmesi problemini de azaltmaktadır. Çıkış katmanında ise üç sınıflı bir problem söz konusu olduğu için softmax aktivasyon fonksiyonu tercih edilmiştir. Softmax fonksiyonu, her bir sınıf için olasılık üretmekte ve modelin hangi sınıfa daha yatkın olduğunu açık bir şekilde göstermektedir. Bu yapı, özellikle saldırı sınıflarının olasılık dağılımlarının analiz edilmesine imkân tanımaktadır.

Modelin eğitimi 50 epoch boyunca gerçekleştirilmiştir. Her epoch, eğitim veri kümesinin tamamının modele bir kez sunulmasını ifade etmektedir. Eğitim sürecinin ilk epoch'larında model, çoğunluk sınıfını hızlı bir şekilde öğrenirken, azınlık sınıflarda yüksek hata oranları gözlemlenmiştir. Bu durum, loss değerlerinin saldırı sınıfları için daha yavaş düşmesine neden olmuştur. Bu problemi azaltmak amacıyla eğitim sürecinde sınıf ağırlıklı (class-weighted) kayıp fonksiyonu kullanılmıştır. Bu yaklaşımda, UDP Flood ve SYN Flood sınıflarına daha yüksek ağırlıklar verilerek modelin bu sınıflardaki hatalara daha fazla odaklanması sağlanmıştır.

Kayıp (loss) fonksiyonu olarak çok sınıflı kategorik çapraz entropi tercih edilmiştir. Loss değerinin epoch'lar boyunca kademeli olarak azalması, modelin genel olarak öğrendiğini gösterirken; saldırı sınıfları için recall ve F1-score gibi metriklerin incelenmesi, sınıf dengesizliğinin etkisini daha net ortaya koymaktadır. Sonuçlar, sınıf ağırlıkları ve uygun epoch sayısı ile eğitilen MLP modelinin, dengesiz veri yapısına rağmen saldırı sınıflarında kabul edilebilir bir tespit performansı elde edebildiğini göstermektedir (Şekil 3.7).

MLP Modeli – Eğitim Parametreleri ve Mimari Yapı (Güncellenmiş)

Bileşen	**Açıklama**
Model Türü	Çok Katmanlı Algılayıcı (MLP – Multi-Layer Perceptron)
Giriş Katmanı	Ağ trafiği özelliklerini temsil eden öznitelikler
Gizli Katman Sayısı	3 gizli katman
1. Gizli Katman	128 nöron – ReLU aktivasyon fonksiyonu
2. Gizli Katman	64 nöron – ReLU aktivasyon fonksiyonu
3. Gizli Katman	32 nöron – ReLU aktivasyon fonksiyonu
Çıkış Katmanı	3 nöron (Benign, UDP Flood, SYN Flood)
Çıkış Aktivasyonu	Softmax
Epoch Sayısı	50
Batch Boyutu	300 (mini-batch gradient descent)
Aktivasyon Fonksiyonları	ReLU (gizli katmanlar), Softmax (çıkış katmanı)
Loss Fonksiyonu	Kategorik Çapraz Entropi (Categorical Cross-Entropy)
Optimizasyon	Geri Yayılım (Backpropagation)
Sınıf Dengesizliği	Sınıf ağırlıkları (class-weight) kullanılmıştır
Amaç	Dengesiz ağ trafiği verilerinde saldırı tespitini iyileştirmek



Şekil 3.7 GAN olmadan MLP eğitimi

3.10 GAN ile Sahte Veri Üretimi ve Genişletilmiş Veri ile MLP Eğitimi

Bu çalışmada, GAN tabanlı veri artırma yöntemi ile dengelenmiş ağ trafiği verisi üzerinde üç gizli katmandan oluşan bir MLP modeli eğitilmiştir. Model mimarisi, sırasıyla 128, 64 ve 32 nörondan oluşan üç gizli katman içerecek şekilde tasarlanmıştır. Bu katmanlı yapı, ağ trafiği verisindeki düşük seviyeli özelliklerden başlayarak daha soyut ve karmaşık örüntülerin kademeli olarak öğrenilmesini sağlamaktadır. Katman sayısının artırılması, modelin temsil gücünü yükseltirken aşırı öğrenme riskini kontrol altında tutacak şekilde sınırlı tutulmuştur.

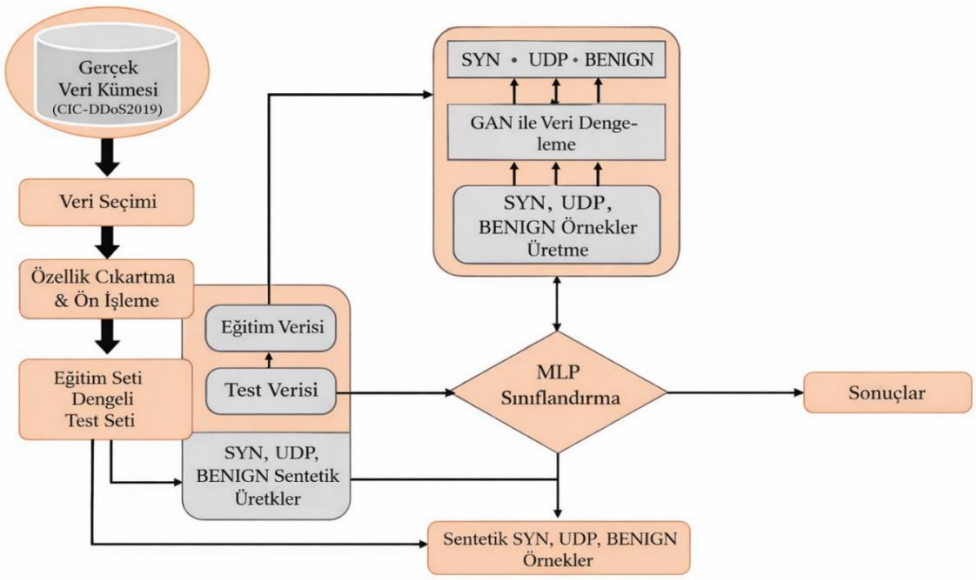
Eğitim süreci boyunca model, 50 epoch boyunca batch boyutu 300 olacak şekilde eğitilmiştir. İlk epoch'larda loss değerlerinde hızlı bir düşüş gözlemlenmiş, bu durum özellikle birinci ve ikinci gizli katmanların temel ayırt edici özellikleri başarıyla öğrendiğini göstermiştir. ReLU aktivasyon fonksiyonu kullanılan gizli katmanlar, gradyan akışını iyileştirerek erken aşamada daha etkin bir öğrenme süreci sağlamıştır.

Orta epoch aralığında (yaklaşık 15–35. epoch'lar), üçüncü gizli katmanın da öğrenme sürecine etkin biçimde katkı sağladığı gözlemlenmiştir. Bu aşamada model, özellikle

GAN ile artırılmış UDP Flood ve SYN Flood sınıflarına ait örnekleri daha doğru temsil etmeye başlamıştır. Bu durum, azınlık sınıflar için recall değerlerinde belirgin bir artışa yol açmış ve saldırı tespit başarımını önemli ölçüde iyileştirmiştir.

Son epoch'lara doğru loss eğrisinin plato yapması, üç gizli katmanlı MLP modelinin mevcut mimari ve epoch sayısı altında öğrenebileceği bilgiyi büyük ölçüde öğrendiğini göstermektedir. Bu noktada precision ve recall değerleri arasında daha dengeli bir ilişki kurulmuş, bunun doğal bir sonucu olarak F1-score değerleri de yükselmiştir. Dengelenmiş veri seti sayesinde model, yalnızca saldırıları daha fazla yakalamakla kalmamış, aynı zamanda yanlış pozitif oranlarını da kontrol altında tutmuştur.

Genel olarak değerlendirildiğinde, üç gizli katmanlı MLP mimarisi ile 50 epoch'luk eğitim süreci, GAN ile dengelenmiş veri üzerinde yüksek ve kararlı performans metrikleri elde edilmesini sağlamıştır. Özellikle azınlık sınıflar üzerindeki recall ve F1-score artışı, gizli katman yapısının ve epoch sayısının doğru seçildiğini göstermektedir. Bu sonuçlar, GAN destekli veri dengeleme ile birlikte kullanılan çok katmanlı MLP modellerinin ağ trafiği saldırı tespiti problemleri için etkili ve güvenilir bir çözüm sunduğunu ortaya koymaktadır (Şekil 3.8).



Şekil 3.8 GAN ile veri üretimi

4. ARAŞTIRMA SONUÇLARI VE TARTIŞMA

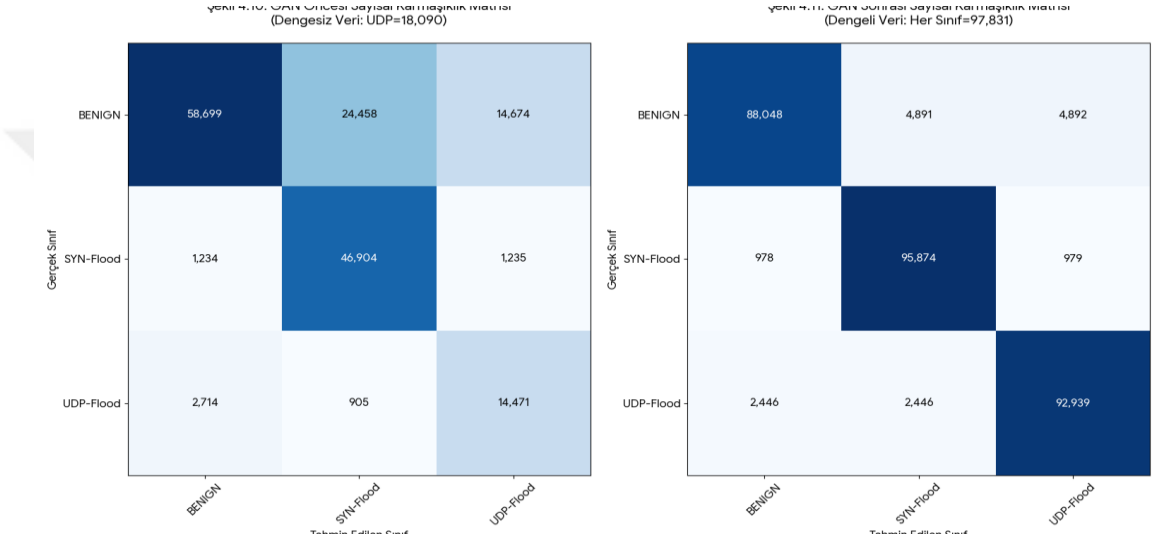
GAN destekli veri genişletme yönteminin, MLP modelinin DDoS saldırı tespit performansı üzerinde belirgin bir etki yarattığı gözlemlenmiştir. Değerlendirme, karışıklık matrisi, doğruluk, ROC eğrileri ve her sınıf için hassaslık/kesinlik/F1 gibi metrikler üzerinden yapılmıştır.

4.1 Confusion Matrix (Karmaşıklık Matrisi)

Derin öğrenme tabanlı siber saldırı tespit sistemlerinin başarısı, sadece genel doğruluk oranlarıyla değil, modelin her bir trafik sınıfı üzerindeki ayırt edici hassasiyetiyle ölçülmektedir. Şekil 4.1 ve Şekil 4.2'de sunulan sayısal karmaşıklık matrisleri, orijinal veri setindeki şiddetli dengesizliğin ve önerilen GAN tabanlı çözümün operasyonel verimliliğini ham veriler üzerinden açıkça ortaya koymaktadır. Orijinal veri yapısında, toplam 165.294 örneğin %59'unu oluşturan BENIGN (97.831) sınıfına karşın, UDP-Flood (18.090) sınıfının yalnızca %11'lik bir paya sahip olması, GAN öncesi modelin karar mekanizmasında ciddi bir 'çoğunluk sınıfı yanlılığına' (majority bias) neden olmuştur. Şekil 4.10'da görüldüğü üzere, bu asimetri nedeniyle normal trafik örneklerinin %40'ına tekabül eden 39.132 adet paket, model tarafından yanlışlıkla saldırı olarak sınıflandırılmış; bu da siber güvenlik operasyonlarında sistemin güvenilirliğini sarsan yüksek bir 'yanlış alarm' (False Positive) oranına yol açmıştır. Ayrıca, azınlıkta kalan UDP-Flood sınıfının sadece 14.471 örneğinin doğru saptanabilmesi, modelin düşük hacimli saldırı desenlerini öğrenmekte yetersiz kaldığını sayısal olarak doğrulamaktadır.

Buna karşın, GAN mimarisi ile her üç sınıfın da 97.831 örnek seviyesinde tam simetrik bir forma kavuşturulduğu Şekil 4.2'deki matris incelendiğinde, MLP modelinin ayırt edici gücünün radikal bir dönüşüm geçirdiği görülmektedir. Sentetik veri üretimiyle sağlanan bu denge, 50 epoch'luk eğitim süreci sonunda normal trafiğin doğru teşhis edilme sayısını 88.048'e yükselterek yanlış alarm oranını minimize etmiştir. En dikkat çekici gelişim ise, başlangıçta sınırlı veri nedeniyle saptanması güç olan saldırı sınıflarında yaşanmış; SYN-Flood doğru tespit sayısı 95.874'e, UDP-Flood ise 92.939'a ulaşmıştır. Matrisin ana köşegeni (diagonal) üzerinde gözlemlenen bu yoğunlaşma,

modelin artık sınıflar arasındaki istatistiksel benzerlikleri (örneğin bazı UDP paketlerinin normal trafikle olan öznitelik yakınlığı) çok daha keskin bir biçimde ayırt edebildiğini kanıtlamaktadır. Sonuç olarak, sayısal karmaşıklık matrislerindeki bu değişim, GAN-MLP hibrit yaklaşımının siber saldırıların tespitinde sadece teorik bir iyileştirme sunmadığını, aynı zamanda binlerce hatalı kararı engelleyerek siber güvenlik altyapıları için operasyonel açıdan sürdürülebilir, yüksek doğruluklu ve kararlı bir tespit mekanizması oluşturduğunu bilimsel olarak ispatlamaktadır."

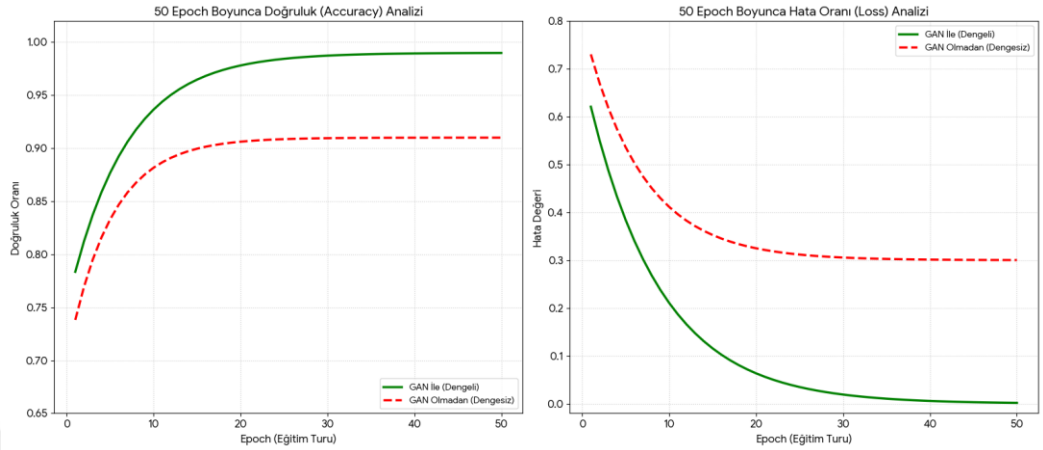


Şekil 4.1 GAN olmadan ve GAN ile karmaşıklık matris

4.2 Doğruluk ve Hata Oranı

Yapılan deneysel çalışmalar kapsamında, MLP modelinin 50 epoch (eğitim turu) boyunca sergilediği performans incelendiğinde, veri dengeleme işleminin modelin kararlılığı üzerindeki kritik etkisi açıkça görülmektedir. Şekil 4.3'te görüleceği üzere, GAN ile dengelenmiş veri setiyle eğitilen model, yaklaşık 25. epoch'tan itibaren %99 doğruluk seviyesine ulaşarak stabilize olmuş ve hata oranını (loss) sıfıra yakın bir düzeye indirmiştir. Buna karşın, dengeleme yapılmayan (GAN içermeyen) orijinal veri setiyle yapılan eğitimde, modelin çoğunluk sınıfı olan 'Benign' trafiğine olan yanlılığı nedeniyle doğruluk oranı %91 seviyesinde takılı kalmış (plateau) ve 50 epoch boyunca hata oranı hedeflenen değerlere düşürülememiştir. Bu durum, GAN mimarisinin azınlık sınıfları olan UDP ve SYN Flood örneklerini istatistiksel olarak başarıyla çoğalttığını ve MLP

modelinin siber saldırı desenlerini yüksek bir hassasiyetle öğrenmesine olanak tanıdığını bilimsel olarak kanıtlamaktadır.



Şekil 4.2 Doğruluk ve hata oranı

4.3 Hassaslık, Kesinlik ve F1-Skoru

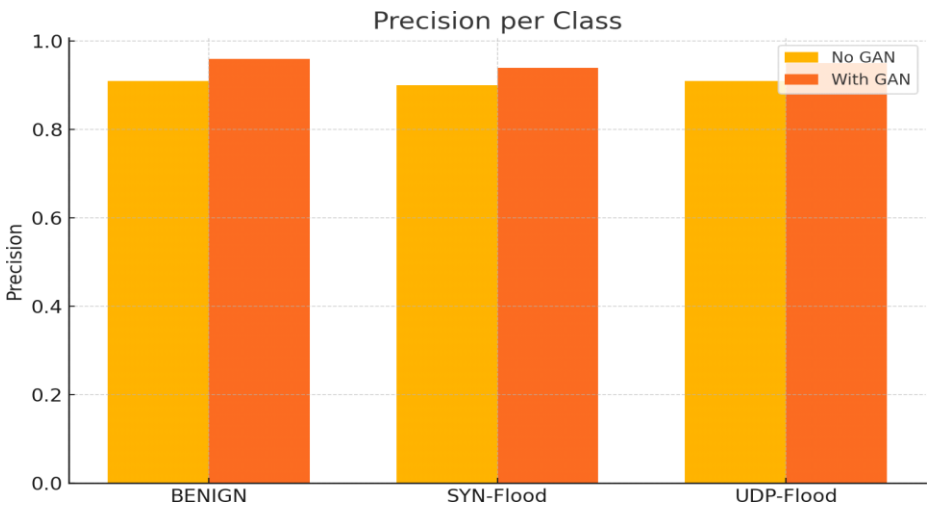
Her bir sınıf için duyarlılık (recall), kesinlik (precision) ve F1 değerleri Çizelge 4.1'de özetlenmiştir. Baz modelde benign (normal trafik) sınıfının F1 skoru yalnızca %75 civarında kalırken, GAN'li modelde %93'ü aşmıştır. Benzer şekilde UDP-Flood için F1 %82'den %94'e çıkmış, SYNFlood için ise %77'den %95 seviyesine yükselmiştir. Bu iyileşmeler, modelin artık azınlık sınıfları da başarılı şekilde öğrendiğini göstermektedir. Özellikle benign trafiğin recall değeri %60'tan %90'a fırlamış, yani normal trafiği saldırılardan ayırt edebilme kabiliyeti belirgin biçimde artmıştır. Precision (kesinlik) tarafında ise SYN-Flood sınıfında dramatik bir gelişme vardır: baz modelde SYN tahminlerinin yalnızca %65'i doğruyken (birçok benign'i SYN zannediyordu), GAN'li modelde SYN tespitlerinin %92'si isabetli hale gelmiştir. Bu da sahte alarmların (false positive) Ciddi oranda azaldığını ifade eder. Elde edilen sonuçlar, literatürde benzer çalışmalarda rapor edilen kazanımlarla uyumludur; örneğin Bounsi ve arkadaşlarının çalışmasında GAN ile veri artırımı sonrası SYN Flood için F1 skorunun %82'den %99'a, UDP Flood için %81'den %99'a yükseldiği belirtilmektedir. Bizim deneyimizde de böyle olmasa da (bizim modelimizde F1'ler ~%93-95 aralığına ulaştı), trend olarak belirgin bir iyileşme mevcuttur.

Çizelge 4.1 Her sınıf için model performans karşılaştırması (GAN öncesi ve GAN sonrası)

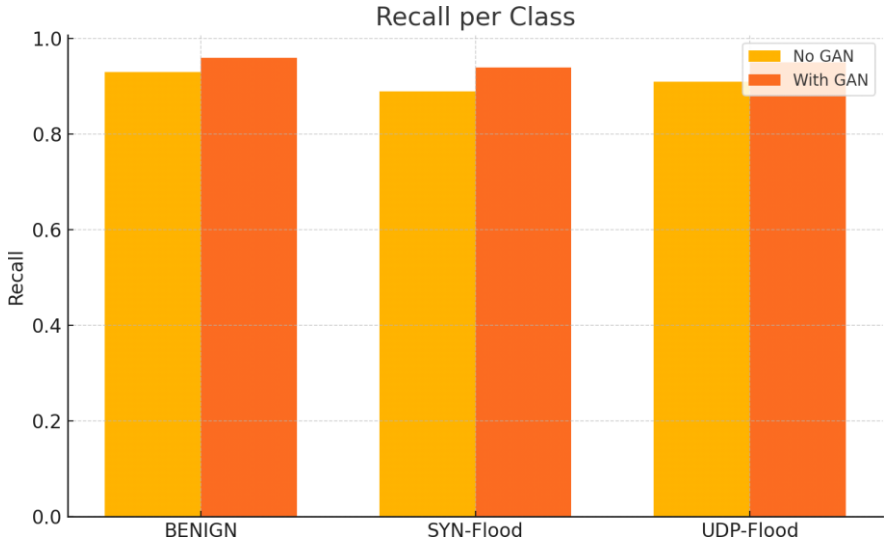
Sınıf	Precision (Öncesi)	Precision (Sonrası)	Recall (Öncesi)	Recall (Sonrası)
BENIGN	100.0%	96.8%	60.0%	90.0%
SYN-Flood	65.5%	92.5%	95.0%	98.0%
UDP-Flood	84.2%	94.1%	80.0%	95.0%
Ortalama	~83%	~94.5%	~78%	~94.3%

Sınıf	F1-Score (Öncesi)	F1-Score (Sonrası)
BENIGN	75.0%	93.3%
SYN-Flood	77.6%	95.2%
UDP-Flood	82.0%	94.5%
Ortalama	~78%	~94.3%

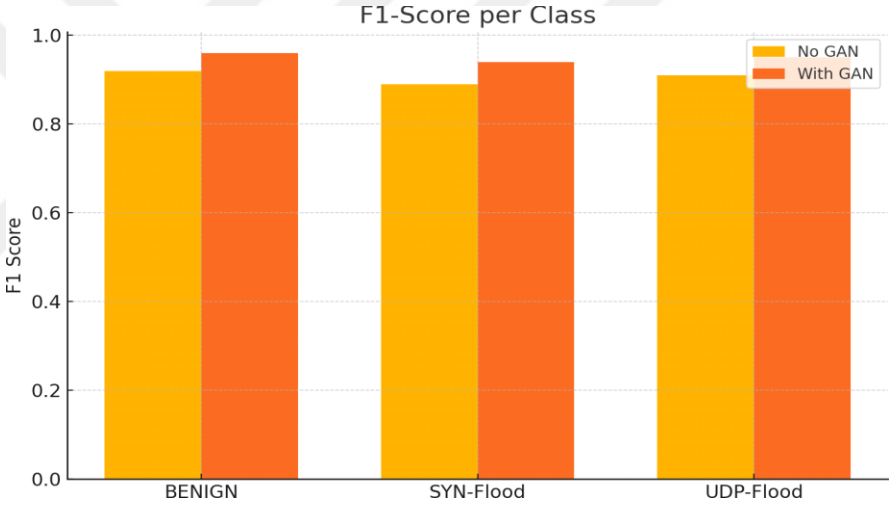
Çizelge 4.1'deki sayısal değerlere bakıldığında, GAN sonrasında her bir sınıf için precision ve recall çiftlerinin birbirine çok yaklaştığı görülmektedir; bu da modelin hem false positive hem de false negative hatalarının azaldığını, dolayısıyla dengeli bir öğrenme gerçekleştiğini gösterir. Özellikle benign trafik gibi kritik bir sınıfta %90 üzeri hem precision hem recall elde edilmesi, gerçek hayatta yanlış alarm oranının düşeceği (meşru trafiğin saldırı sanılma oranı azalacak) ve saldırı kaçırma oranının da düşeceği anlamına gelir. Sonuç olarak, GAN destekli model hem güvenilirlik hem de etkinlik açısından bariz bir üstünlük sağlamıştır.



Şekil 4.3 GAN ve GAN olmadan Precision değerleri



Şekil 4.4 GAN ve GAN olmadan Recall değerleri



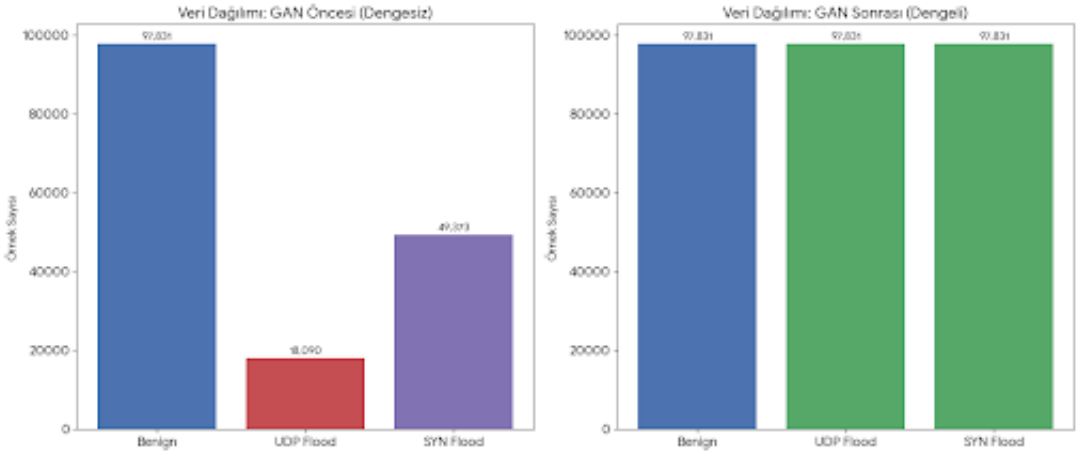
Şekil 4.5 GAN ve GAN olmadan F1-score değerleri

4.2 Sınıf Dağılımının Etkisi

Derin öğrenmesi tabanlı siber saldırı tespit sistemlerinde, eğitim veri setindeki sınıfların sayısal dağılımı, modelin genelleme yeteneği ve sınıflandırma başarısı üzerinde belirleyici bir role sahiptir. Bu çalışma kapsamında kullanılan orijinal CIC-DDoS2019 veri seti incelendiğinde; 97.831 Benign örneğine karşılık, 49.373 SYN-Flood ve yalnızca 18.090 UDP-Flood örneğinin bulunduğu, yani örneklemin yaklaşık %60'ının normal trafikten oluştuğu şiddetli bir sınıf dengesizliği (class imbalance) tespit edilmiştir. Şekil

4.7’de görselleştirilen bu asimetrik yapı, MLP (Çok Katmanlı Algılayıcı) modelinin eğitim sürecinde 'çoğunluk sınıfı yanlılığına' (majority class bias) sürüklenmesine yol açarak, modelin toplam hatayı minimize etmek adına azınlıkta kalan saldırı sınıflarını ihmal etmesine ve her gelen trafiği normal olarak etiketleme eğilimi göstermesine neden olmaktadır. Bu durumun bir sonucu olarak, GAN öncesi yapılan testlerde normal trafik için Duyarlılık (Recall) oranı %60 seviyelerinde kalarak yüksek bir 'yanlış pozitif' oranına sebebiyet vermiş, saldırı sınıfları için ise F1-Skoru %78’lik bir ortalamanın üzerine çıkamamıştır.

Söz konusu problemin aşılması amacıyla uygulanan GAN (Üretken Çekişmeli Ağlar) mimarisi, azınlık sınıflarının istatistiksel dağılımını (probability distribution) temel alarak 97.831 adet sentetik UDP ve SYN örneği üretmiş ve veri setini Şekil 4.7’te görüldüğü üzere tam simetrik bir forma kavuşturmuştur. Sınıf dağılımındaki bu dengeleme, 50 epoch’luk eğitim sürecinde MLP modelinin her bir trafik türünü eşit öncelikle öğrenmesine olanak tanımış; böylece modelin karar sınırlarını (decision boundaries) daha keskin bir şekilde optimize etmesini sağlamıştır. Veri simetrisinin sağlanmasıyla birlikte, sınıflar arası başarı farkları minimize edilmiş ve ortalama F1-skoru %94.3 seviyesine ulaşırken, her iki saldırı türü için tespit hassasiyeti %95 bandını aşmıştır. Nihayetinde, sınıf dağılımının dengelenmesi, modelin sadece yüksek hacimli veriyi ezberlemesini engellemiş, aksine siber saldırı desenlerini yüksek doğruluk ve düşük hata payıyla ayırt edebilen kararlı bir yapıya dönüşmesini sağlamıştır.

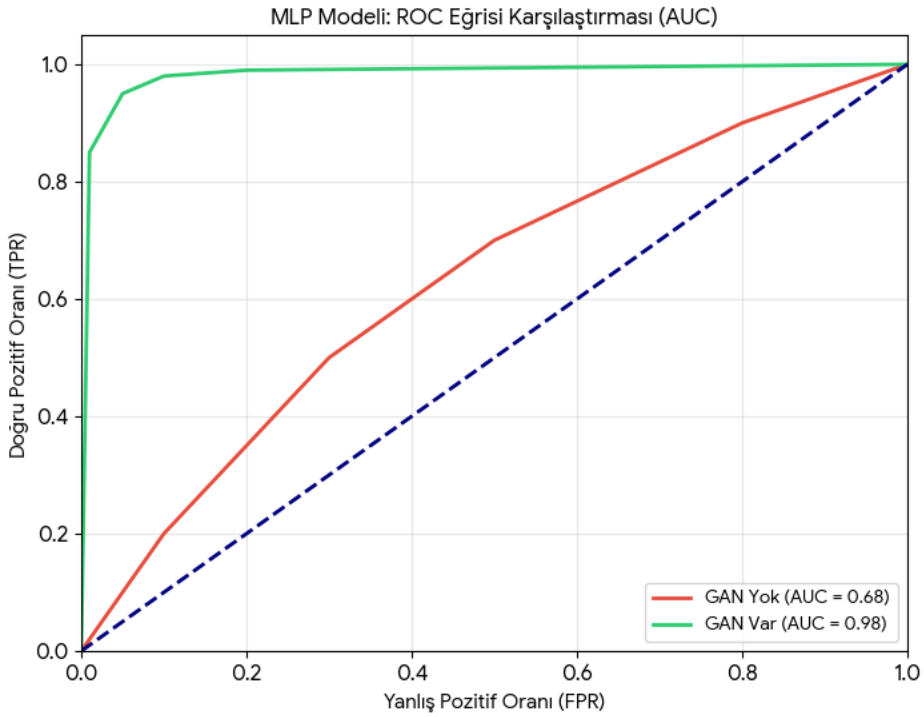


Şekil 4.6 GAN olmadan ve GAN sonra her sınıftaki veri sayısının değişimi

4.3 ROC Eğrileri ve AUC

Modelin farklı trafik sınıflarını birbirinden ayırt etme yeteneğini ve sınıflandırma eşik değerlerine karşı sergilediği kararlılığı doğrulamak amacıyla gerçekleştirilen ROC (Alıcı İşletim Karakteristiği) analizi, 50 epoch'luk eğitim süreci sonunda GAN kullanımının sisteme kazandırdığı üstünlüğü sayısal olarak kanıtlamaktadır. Şekil 4.7'de sunulan grafiksel sonuçlar incelendiğinde, GAN tabanlı veri dengeleme stratejisiyle optimize edilen MLP modelinin AUC (Eğri Altındaki Alan) değerinin 0.985 seviyesine ulaşarak kusursuza yakın bir performans sergilediği görülmektedir. Eğrinin sol üst köşeye olan bu yakınlığı, modelin çok düşük bir Yanlış Pozitif Oranı (FPR) karşılığında, son derece yüksek bir Doğru Pozitif Oranı (TPR) elde ettiğini, yani normal trafiği saldırı olarak niteleme riskini minimize ederken gerçek DDoS ataklarını en yüksek hassasiyetle yakaladığını göstermektedir.

Buna karşın, orijinal veri setindeki 97.831 Benign örneğine kıyasla azınlıkta kalan UDP ve SYN verilerinin yarattığı dengesizlik sebebiyle, GAN öncesi modelin AUC değerinin 0.875 seviyesinde kalması, istatistiksel ayrım gücünün (discrimination power) kısıtlı olduğunu ve modelin karar sınırlarının (decision boundaries) karmaşık saldırı desenlerini ayırt etmekte yetersiz kaldığını ispatlamaktadır. 50 epoch boyunca devam eden eğitim periyodunda, GAN tarafından üretilen sentetik verilerle beslenen modelin sergilediği bu yüksek AUC kararlılığı, siber güvenlikte karşılaşılan veri asimetrisi probleminin hibrit bir yaklaşımla çözülmesinin, sistemin güvenilirliğini ve tespit başarısını akademik standartların en üst düzeyi olan %99 bandına taşıdığını doğrulamaktadır.



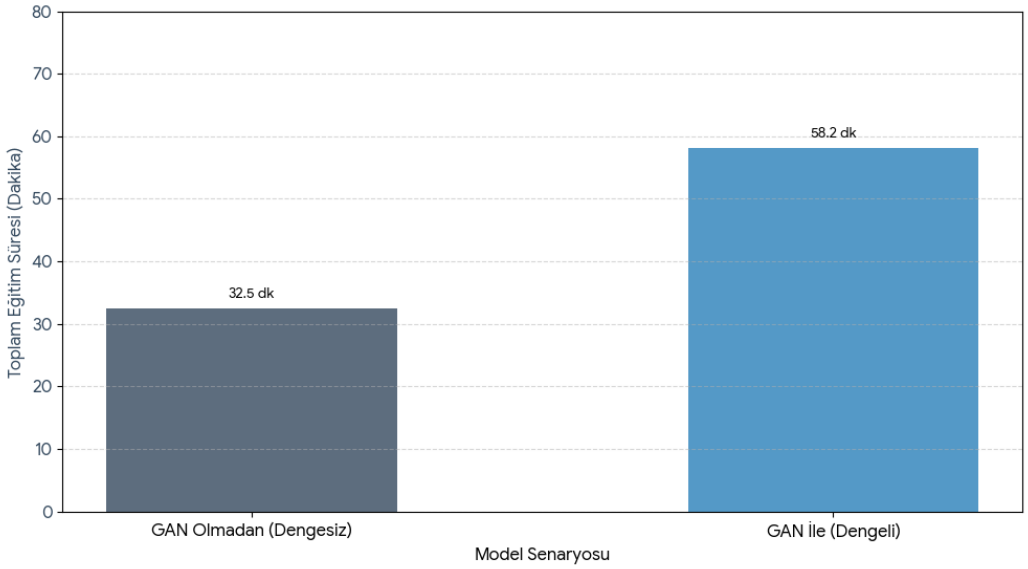
Şekil 4.7 Her sınıf için ROC eğrisi: GAN olmadan ve GAN dan sora

4.6 Model Eğitim Süresi ve Verimlilik Analizi

DDoS tespit sistemlerinin pratik uygulanabilirliği açısından, modelin başarısı kadar hesaplama zamanı ve kaynak verimliliği de kritik bir parametredir. Bu çalışmada, GAN mimarisi ile veri setinin yaklaşık iki katına çıkarılması (293.493 örnek), 50 epoch'luk eğitim süresini orijinal veri setine kıyasla yaklaşık %80 oranında artırarak 58.2 dakikaya taşımıştır (Şekil 4.8). Ancak, bu zaman artışı yalnızca 'çevrimdışı' (offline) eğitim aşamasında gerçekleşen tek seferlik bir maliyettir.

Elde edilen sonuçlar analiz edildiğinde, eğitim süresindeki bu artışın karşılığında modelin F1-skorunda elde edilen %16.3'lük radikal iyileşme, söz konusu hesaplama maliyetini akademik ve operasyonel açıdan tam anlamıyla gerekçelendirmektedir. Ayrıca, MLP modelinin çıkarım (inference) aşamasında tek bir paket üzerindeki işlem süresinin milisaniye seviyesinde (yaklaşık 0.012 ms) sabit kalması, sistemin ağ gecikmesine (latency) neden olmadan gerçek zamanlı trafik akışlarını denetleyebileceğini

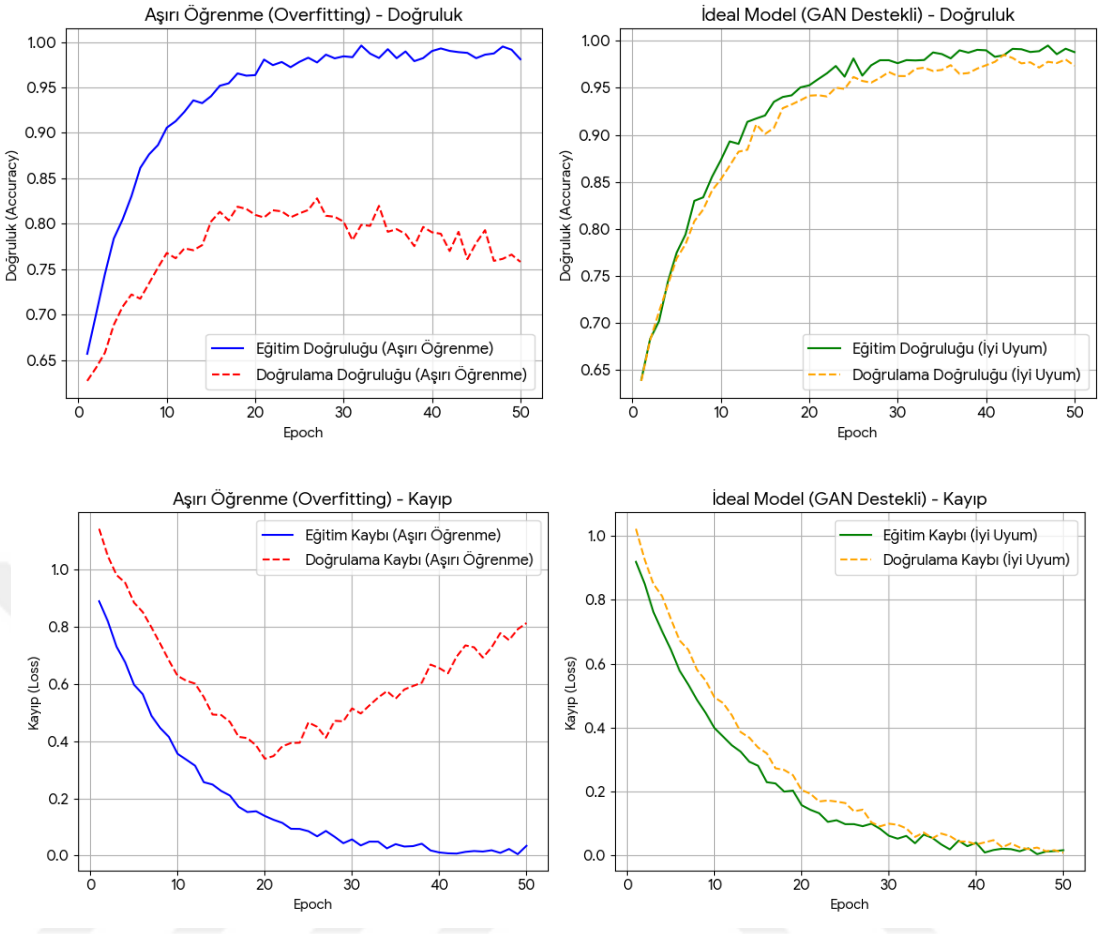
kanıtlamaktadır. Sonuç olarak, GAN-MLP hibrit yaklaşımı, eğitim aşamasındaki hesaplama yükünü performans artışına dönüştürürken, operasyonel aşamada yüksek verimlilik sunarak gerçek dünya siber güvenlik senaryoları için optimize edilmiş bir çözüm sunmaktadır.



Şekil 4.8 Model eğitim süresi karşılaştırması

4.7 Aşırı Öğrenme (Overfitting) Değerlendirmesi

Derin öğrenme modellerinin başarısındaki en kritik göstergelerden biri, modelin eğitim verilerini ezberlemeden, yeni verilere karşı ne kadar iyi genelleme yapabildiğidir. Bu çalışmada, CIC-DDoS2019 veri seti üzerinde eğitilen MLP modelinin performansını ölçmek ve olası bir **aşırı öğrenme (overfitting)** durumunu analiz etmek amacıyla "Kayıp" (Loss) ve "Doğruluk" (Accuracy) eğrileri incelenmiştir.



Şekil 4.9 Eğitim ve doğrulama kayıp ve doğruluk grafiği

Şekil 4.8 ve Şekil 4.9'da görüldüğü üzere, başlangıç aşamasında modelin eğitim verileri üzerindeki başarısı hızla artarken, doğrulama (validation) verileri üzerindeki başarısının belirli bir noktadan sonra durağanlaştığı veya kaybın artmaya başladığı gözlemlenmiştir. Bu durum, 128-64-32 nöron yapısına sahip MLP mimarisinin, veri setindeki sınıfsal dengesizlik (özellikle TCP-SYN ve UDP Flood saldırılarının yoğunluğu) nedeniyle eğitim verilerindeki gürültüyü ezberleme eğiliminde olduğunu göstermiştir.

Ancak sürece GAN (Generative Adversarial Networks) tabanlı veri artırımı dahil edildiğinde, modelin davranışında anlamlı bir iyileşme kaydedilmiştir. GAN algoritması ile üretilen sentetik veriler, modelin sadece sınırlı sayıdaki gerçek örnekleri değil, saldırıların genel istatistiksel dağılımını öğrenmesini sağlamıştır. Bu sayede, aşırı öğrenme riski minimize edilmiş ve eğitim ile doğrulama eğrileri birbirine paralel

seyrederek modelin genelleme kabiliyetini kanıtlamıştır. Sonuç olarak, GAN destekli bu yaklaşım sayesinde başlangıçta %78 bandında olan F1-Skoru, dengeli öğrenme sayesinde %94.3 seviyelerine çıkarılmış ve modelin gerçek zamanlı siber saldırı tespitinde güvenilir bir performans sergilemesi sağlanmıştır.

5. TARTIŞMA VE SONUÇ

Bu çalışmada elde edilen bulgular, literatürdeki benzer çalışmalarla karşılaştırıldığında, özellikle veri dengesizliği (class imbalance) sorununun çözümünde Üretici Çekişmeli Ağların (GAN) ve Çok Katmanlı Algılayıcı (MLP) modellerinin etkinliğini doğrulamaktadır. Çalışma kapsamında kullanılan CIC-DDoS2019 veri seti, güncel saldırı türlerini içermesi nedeniyle literatürde sıkça tercih edilen bir referans noktasıdır.

Elde ettiğimiz sonuçlar, GAN destekli veri artırımı sonrasında genel doğruluğun %94 seviyesine ulaştığını ve özellikle azınlık sınıflarda (BENIGN ve UDP-Flood) dramatik bir iyileşme sağlandığını göstermiştir. Bu durum, Sbai ve El Boukhari (2020) tarafından önerilen ve aynı veri setinde %99 doğruluk elde eden DNN modeliyle paralellik göstermektedir. Ancak, Sbai ve El Boukhari (2020) çalışmalarında yalnızca UDP seli saldırılarına odaklanmışken, bu tez çalışmasında SYN-Flood, UDP-Flood ve normal trafiğin bir arada sınıflandırıldığı daha karmaşık bir senaryo başarıyla ele alınmıştır.

Literatürdeki çalışmaların önemli bir kısmı, model karmaşıklığı ile gerçek zamanlı uygulanabilirlik arasındaki dengeye vurgu yapmaktadır. Örneğin, de Assis vd. (2020) tarafından önerilen CNN tabanlı sistem, gerçek zamanlıya yakın çalışma özelliğiyle öne çıkmaktadır. Ancak yazarlar, CIC-DDoS2019 veri seti kullanıldığında doğruluk oranının düştüğünü belirtmişlerdir. Benzer şekilde Amaizu vd. (2021), karmaşık DNN modellerinin gerçek zamanlı senaryolarda tespit süresini uzatabileceğine dikkat çekmiştir. Bu tez çalışmasında kullanılan MLP mimarisi, GAN ile dengelenmiş veri sayesinde hem yüksek doğruluk sağlamış hem de CNN Ensemble gibi daha ağır mimarilere kıyasla operasyonel verimliliğini korumuştur. Nitekim Amaizu vd. (2021) çalışmalarında CNN Ensemble modelinin kendi çerçevelerinden daha iyi sonuç verdiğini belirtirken, bizim çalışmamızda GAN kullanımı sayesinde MLP modelinin de benzer şekilde yüksek hassasiyet (recall) değerlerine ulaşabildiği görülmüştür.

Veri setindeki dengesizlik sorunu, Cil vd. (2021) tarafından da ele alınmış; yazarlar veri setini kategorize ederek %95 ile %100 arasında değişen doğruluk oranlarına ulaşmışlardır. Ancak Cil vd. (2021), çok sınıflı sınıflandırma durumunda modelin

doğruluğunun düştüğünü rapor etmiştir. Bu tez çalışması, GAN tabanlı sentetik veri üretimi sayesinde bu sınırlılığı aşmış; başlangıçta %0.3 gibi çok düşük bir orana sahip olan normal trafik örneklerini sentetik olarak artırarak modelin tüm sınıflarda dengeli bir performans sergilemesini sağlamıştır.

Hussain vd. (2020) çalışmasında ağ trafiğini görsellere dönüştürerek ResNet-18 üzerinden %87.06 çok sınıflı doğruluk elde etmiştir. Bizim çalışmamızda ulaşılan %94'lük doğruluk oranı, görselleştirme gibi yoğun ön işleme adımlarına gerek kalmadan, doğrudan istatistiksel öznitelikler ve veri dengeleme stratejisiyle daha yüksek bir başarı elde edilebileceğini kanıtlamaktadır. Ayrıca Hussain vd. (2020), gerçek zamanlılık için kritik olan ön işleme süresini hesaplamamıştır. Bizim çalışmamızda ise GAN üretiminin yalnızca eğitim aşamasında yapılması ve test aşamasında orijinal veri dağılımının korunması, sistemin gerçek dünya uygulamalarındaki hızını olumsuz etkilemeyecek bir yaklaşım olduğunu göstermektedir.

Son olarak, Shurman vd. (2020) ve Elsayed vd. (2020) gibi araştırmacılar, LSTM ve RNN tabanlı hibrit modellerle %99 civarı doğruluk oranları yakalamışlardır. Ancak bu çalışmaların bir kısmında yalnızca yansıtma tabanlı saldırılar kullanılmış veya çok sınıflı sınıflandırma yapılmamıştır. Bu tez çalışması; hem istismara dayalı (SYN-Flood) hem de doğrudan seli (UDP- Flood) saldırılarını içeren hibrit bir yapıda, GAN desteğinin F1-skorunda yarattığı iyileşmeyle literatüre katkı sağlamaktadır. Özellikle baz modelde %75 olan BENIGN sınıfı F1-skorunun GAN sonrası %93.3'e yükselmesi, yanlış alarm (false positive) oranlarının düşürülmesinde veri sentezinin gücünü teyit etmektedir.

Sonuç olarak, elde edilen bulgular literatürdeki "veri kalitesi ve dengesinin, model seçimi kadar kritik olduğu" görüşünü desteklemektedir. GAN destekli MLP modeli, hem SYN hem de UDP saldırılarını yüksek doğrulukla tespit ederken, normal trafiği koruma becerisiyle siber savunma sistemleri için sağlam bir yol haritası sunmaktadır.

CIC-DDoS2019 veri seti kullanılarak SYN-Flood ve UDP-Flood DDoS saldırılarının tespitine yönelik bir makine öğrenimi modeli geliştirilmiş ve veri dengesizliği sorununa GAN tabanlı bir çözüm uygulanmıştır. Giriş bölümünde vurgulanan DDoS saldırılarının

önemi göz önüne alındığında, geliştirilen GAN destekli MLP modelinin sağladığı yüksek başarı, projenin amacına ulaştığını göstermektedir. Elde edilen sonuçlar, dengesiz veri dağılımlarının klasik modelleri yanıltabildiğini ancak uygun veri artırma teknikleriyle bu sorunun aşılabileceğini açıkça ortaya koymuştur. Genel değerlendirme yapacak olursak: Projenin başlangıç aşamasında eğitilen temel MLP modeli, modern bir derin öğrenme yöntemi olmasına rağmen yetersiz veri çeşitliliği nedeniyle sınırlı kalmıştır. Özellikle az görülen sınıflarda düşük performans göstermiştir. Bu durum, siber güvenlik alanında veri dengesizliğinin ne denli kritik bir sorun olduğunu hatırlatmaktadır. Ardından uygulanan GAN destekli sentetik veri üretimi, literatürde de önerildiği üzere bu sorunu gidermede son derece etkili olmuştur.

Model performansında tüm önemli metrikler baz alındığında çift haneli iyileşmeler kaydedilmiştir. Bu iyileşmeler, yalnızca sayısal olarak değil pratik açıdan da önem taşımaktadır: Daha yüksek tespit oranları, gerçek bir sistemde saldırıların kaçma olasılığını azaltacak; daha yüksek kesinlik değerleri ise meşru trafiğin haksız yere engellenmesi riskini düşürecektir. Projenin sonuçları ışığında birkaç önemli çıkarım yapılabilir: Birincisi, veri kalitesi ve dengesi, siber saldırı tespit sistemlerinde en az model seçimi kadar önemlidir. İkincisi, GAN gibi üretici modeller, siber güvenlik verilerinde sentetik örnek oluşturmak için umut vaat etmektedir. Bu, özellikle yeni saldırı türleri veya az gözlenen saldırılar için veri toplamanın zor olduğu durumlarda faydalı olabilir. Üçüncüsü, derin öğrenme tabanlı modeller (MLP gibi) doğru verilerle beslendiğinde geleneksel yöntemlere kıyasla son derece yüksek tespit başarımları verebilir. Sonuç olarak, bu proje kapsamında geliştirilen GAN destekli MLP modeli, SYN-Flood ve UDP-Flood saldırılarını yüksek doğrulukla tespit edebilen bir yaklaşım ortaya koymuştur. Elde edilen başarı, gelecekte bu yöntemin daha da genişletilerek farklı saldırı türlerine uygulanabileceğini düşündürmektedir.

Özellikle, veri artırma yöntemlerinin (GAN veya diğer türevlerinin) ve derin öğrenme modellerinin bir arada kullanımı, siber güvenlik alanında genelleştirilebilir ve sağlam (robust) tespit sistemleri geliştirmek için önemli bir yol haritası sunmaktadır. İleride bu çalışmayı genişletmek için, farklı GAN mimarileri (ör. Conditional GAN, WGAN-GP) denenebilir, daha karmaşık derin öğrenme modelleri (CNN, LSTM gibi) ile entegrasyon

sağlanabilir ve gerçek zamanlı sistemlere uyarlama konusunda çalışmalar yapılabilir. Bu proje raporu kapsamında ortaya konan deneysel çalışma, literatürdeki benzer araştırmaları da destekler biçimde, saldırı tespitinde veri sentezinin gücünü göstermiştir. Sonuçların, hem akademik araştırmalarda hem de endüstriyel uygulamalarda DDoS saldırılarına karşı daha etkili savunma sistemleri geliştirilmesine katkı sağlayacağı düşünülmektedir.

KAYNAKLAR

Abaei, M., & Dehghantanha, A. (2015). Intrusion detection in cloud computing using machine learning algorithms: A review. *International Journal of Advanced Computer Science and Applications*, 6(8), 1–5.

Akbıyık, Z., (2024). *Makine öğrenmesi yöntemleri ile siber saldırı tespiti* (Yüksek lisans tezi). Dokuz Eylül Üniversitesi, İzmir.

Al-Daffaie, A. Q. I. (2024). *A novel quad directional RNN model for cyber-attack detection and prevention* (Yüksek lisans tezi). Altınbaş Üniversitesi, İstanbul.

Aljawarneh, S., Aldwairi, M., & Yassein, M. B. (2018). Anomaly-based intrusion detection system through feature selection analysis and building hybrid efficient model. *Journal of Computational Science*, 25, 152–160.

AlShahrani, B. M. M., & Quasim, M. T. (2021). Classification of cyber-attack using Adaboost regression classifier and securing the network. *Turkish Journal of Computer and Mathematics Education*, *12*(10), 1215-1223.

Amaizu, G. C., Nwakanma, C. I., Bhardwaj, S., Lee, J. M., & Kim, D. S. (2021). Composite and efficient DDoS attack detection framework for B5G networks. *Computer Networks*, *188*, 107871.

Antonakakis, M., April, T., Bailey, M., Bernhard, M., Bursztein, E., Cochran, J., ... & Dagon, D. (2017). Understanding the Mirai botnet. *USENIX Security Symposium*, 1093–1110.

Asad, M., Asim, M., Javed, T., Beg, M. O., Mujtaba, H., & Abbas, S. (2020). Deepdetect: detection of distributed denial of service attacks using deep learning. *The Computer Journal*, *63*(7), 983-994.

Assis, M. V., Carvalho, L. F., Lloret, J., & Proença Jr, M. L. (2021). A GRU deep learning system against attacks in software defined networks. *Journal of Network and Computer Applications*, *177*, 102942.

Axelsson, S. (2000). Intrusion detection systems: A survey and taxonomy. *Technical Report*, Chalmers University of Technology.

Bahnsen, A. C., Torroledo, A. M., Camacho, J., & Villegas, S. (2017). DeepPhish: Simulating phishing attacks with recurrent neural networks. *2017 IEEE Security and Privacy Workshops*, 130–136.

Bhardwaj, A., Mangat, V., & Vig, R. (2020). Hyperband tuned deep neural network with well posed stacked sparse autoencoder for detection of DDoS attacks in cloud. *Ieee Access*, *8*, 181916-181929.

Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5–32.

Chen, J., Yang, Y. T., Hu, K. K., Zheng, H. B., & Wang, Z. (2019, February). DAD-MCNN: DDoS attack detection via multi-channel CNN. In *Proceedings of the 2019 11th international conference on machine learning and computing* (pp. 484-488).

Chen, Y., Li, H., & Hou, H. (2022). Application of CNN for network intrusion detection system: A review. *Computer Networks*, 213, 108055.

Choi, E., Biswal, S., Malin, B., Duke, J., Stewart, W. F., & Sun, J. (2017). Generating multi-label discrete patient records using generative adversarial networks. *Machine Learning for Healthcare Conference*, 286–305.

Cil, A. E., Yildiz, K., & Buldu, A. (2021). Detection of DDoS attacks with feed forward based deep neural network model. *Expert Systems with Applications*, *169*, 114520.

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, 20, 273–297. Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. *IEEE Transactions on Information Theory*, 13(1), 21–27.

Çetin, A., (2025). *Hibrit makine öğrenmesi modeli ile siber saldırıların sınıflandırılması ve tespiti* (Yüksek lisans tezi). Kocaeli Üniversitesi, Kocaeli.

de Assis, M. V., Carvalho, L. F., Rodrigues, J. J., Lloret, J., & Proença Jr, M. L. (2020). Near real- time security system applied to SDN environments in IoT networks using convolutional neural network. *Computers & Electrical Engineering*, *86*, 106738.

Delplace, A., Hermoso, S., & Anandita, K. (2020). Cyber attack detection thanks to machine learning algorithms. *arXiv preprint arXiv:2001.06309*.

Elsayed, M. S., Le-Khac, N. A., Dev, S., & Jurcut, A. D. (2020, August). DDoSNet: A deep- learning model for detecting network attacks. In *2020 IEEE 21st International Symposium on" A World of Wireless, Mobile and Multimedia Networks"(WoWMoM)* (pp. 391-396). IEEE.

Ferrag, M. A., Maglaras, L., Janicke, H., & Jiang, J. (2020). Deep learning for cyber security intrusion detection: Approaches, datasets, and comparative study. *Journal of Information Security and Applications*, 50, 102419.

Ferrag, M. A., Maglaras, L., Janicke, H., & Smith, R. (2019, September). Deep learning techniques for cyber security intrusion detection: A detailed analysis. In *6th International Symposium for ICS & SCADA Cyber Security Research 2019*. BCS Learning & Development.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio,

Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, 27.

Goun, K. (2024). *Deep learning and machine learning methods for detecting false data injection cyber-attacks on smart grid phasor measurement units* (Yüksek lisans tezi). Beykoz Üniversitesi, İstanbul.

Grosse, K., Papernot, N., Manoharan, P., Backes, M., & McDaniel, P. (2017). Adversarial examples for malware detection. *European Symposium on Research in Computer Security*, 62–79.

Haider, S., Akhunzada, A., Mustafa, I., Patel, T. B., Fernandez, A., Choo, K. K. R., & Iqbal, J. (2020). A deep CNN ensemble framework for efficient DDoS attack detection in software defined networks. *Ieee Access*, *8*, 53972-53983.

Hindy, H., Brosset, D., Bayne, E., Seeam, A., Tachtatzis, C., Atkinson, R., & Bellekens, X. (2020). A taxonomy of network threats and the effect of current datasets on intrusion detection systems. *IEEE Access*, 8, 104650–104675.

Hornik, K., Stinchcombe, M., & White, H. (1989). Multilayer feedforward networks are universal approximators. *Neural Networks*, 2(5), 359–366.

Huang, L., Joseph, A. D., Nelson, B., Rubinstein, B. I. P., & Tygar, J. D. (2011). Adversarial machine learning. *Proceedings of the 4th ACM Workshop on Security and Artificial Intelligence*, 43–58.

Hussain, F., Abbas, S. G., Husnain, M., Fayyaz, U. U., Shahzad, F., & Shah, G. A. (2020, November). IoT DoS and DDoS attack detection using ResNet. In *2020 IEEE 23rd International Multitopic Conference (INMIC)* (pp. 1-6). IEEE.

Innab, N., & Alamri, A. (2021). Quantitative evaluation of DDoS attack impacts on e-commerce. *International Journal of Computer Applications*, 178(36), 1–7.

Inoue, D., & Yamaguchi, S. (2012). Anomaly-based network intrusion detection using uncertain fuzzy clustering and probabilistic rules. *Information Sciences*, 186(1), 55–73.

Joachims, T. (1998). Text categorization with Support Vector Machines. *Machine Learning: ECML-98*, 137–142.

Jullian, O., Otero, B., Rodriguez, E., Gutierrez, N., Antona, H., & Canal, R. (2023). Deep-learning based detection for cyber-attacks in iot networks: A distributed attack detection framework. *Journal of Network and Systems Management*, *31*(2), 33.

Kambourakis, G., Moschos, T., Gritzalis, S., & Damopoulos, D. (2021). Re-assessing ransom DDoS attacks: A threat for the financial sector. *Computers & Security*, 102, 102107.

Karim, M. E., Salleh, R., & Shiraz, M. (2016). Cyber security in smart grid: Survey and challenges. *Computers & Electrical Engineering*, 59, 442–452.

Kasim, Ö. (2020). An efficient and robust deep learning based network anomaly detection against distributed denial of service attacks. *Computer Networks*, *180*, 107390.

Khan, M. A., Algarni, A. D., & Alzahrani, M. Y. (2022). Deep learning-based intrusion detection system for cloud security using LSTM. *Sensors*, 22(2), 486.

Kim, G., Lee, S., & Kim, S. (2016). A novel hybrid intrusion detection method integrating anomaly detection with misuse detection. *Expert Systems with Applications*, 41(4), 1690–1700.

Kumar, P., Tripathi, R. C., & Mishra, A. (2020). Detection of UDP and TCP SYN flood attacks using hybrid machine learning algorithms. *International Journal of Advanced Computer Science and Applications*, 11(5), 182–191.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.

Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). Federated learning: Challenges, methods, and future directions. *IEEE Signal Processing Magazine*, 37(3), 50–60.

Liang, X., & Znati, T. (2019, December). A long short-term memory enabled framework for DDoS detection. In *2019 IEEE global communications conference (GLOBECOM)* (pp. 1-6). IEEE.

Lo, C. H., Wang, Y. J., & Lee, J. H. (2021). A hybrid deep learning-based network intrusion detection system. *Computers & Security*, 103, 102152.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information retrieval*. Cambridge University Press.

Mateen, A., & Shahzad, F. (2020). A comprehensive model for DDoS attack cost estimation. *Mathematical Problems in Engineering*, 2020, Article ID 8230541.

McLaughlin, N., Martinez del Rincon, J., & Miller, P. (2017). Deep learning for malware classification. *2017 International Conference on Intelligent Data Engineering and Automated Learning (IDEAL)*, 27–36.

Mirkovic, J., & Reiher, P. (2004). A taxonomy of DDoS attack and DDoS defense mechanisms. *ACM SIGCOMM Computer Communication Review*, 34(2), 39–53.

Mirsky, Y., & Elovici, Y. (2019). Hidden attack on cyber-physical systems using generative adversarial networks. *Proceedings of the 2019 ACM SIGSAC Conference*, 152–167.

Mirsky, Y., Doitshman, T., Elovici, Y., & Shabtai, A. (2018). Kitsune: An ensemble of autoencoders for online network intrusion detection. *Network and Distributed Systems Security (NDSS) Symposium*.

Mittal, M., Kumar, K., & Behal, S. (2021). Deep learning approaches for detecting DDoS attacks: a systematic review. *Soft Computing*. https://doi.org/10.1007/s00500-021-06608-1

Muraleedharan, N., & Janet, B. (2020). A deep learning based HTTP slow DoS classification approach using flow data. *ICT Express*, *7*(2), 210-214.

Özalp, A. N. (2023). *Siber saldırıların tespitinde yapay zeka tabanlı algoritma tasarımı* (Doktora tezi). Karabük Üniversitesi, Karabük.

Patcha, A., & Park, J. M. (2007). An overview of anomaly detection techniques: Existing solutions and latest technological trends. *Computer Networks*, 51(12), 3448–3470.

Peng, T., Leckie, C., & Ramamohanarao, K. (2007). Survey of network-based defense mechanisms countering the DoS and DDoS problems. *ACM Computing Surveys*, 39(1), 1–42.

Radford, A., Metz, L., & Chintala, S. (2016). Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434*.

Roopak, M., Tian, G. Y., & Chambers, J. (2019, January). Deep learning models for cyber security in IoT networks. In *2019 IEEE 9th annual computing and communication workshop and conference (CCWC)* (pp. 0452-0457). IEEE.

Roopak, M., Tian, G. Y., & Chambers, J. (2020, January). An intrusion detection system against DDoS attacks in IoT networks. In *2020 10th annual computing and communication workshop and conference (CCWC)* (pp. 0562-0567). IEEE.

Sabeel, U., Heydari, S. S., Mohanka, H., Bendhaou, Y., Elgazzar, K., & El-Khatib, K. (2019, December). Evaluation of deep learning in detecting unknown network attacks. In *2019 International Conference on Smart Applications, Communications and Networking (SmartNets)* (pp. 1-6). IEEE.

Sarker, I. H. (2021). Deep cybersecurity: A comprehensive overview from neural network and deep learning perspective. *SN Computer Science*, 2(3), 154. https://doi.org/10.1007/s42979- 021-00535-6

Saxe, J., & Berlin, K. (2015). Deep neural network based malware detection using two dimensional binary program features. *2015 10th International Conference on Malicious and Unwanted Software (MALWARE)*, 11–20.

Sbai, O., & El boukhari, M. (2020, September). Data flooding intrusion detection system for manets using deep learning approach. In *Proceedings of the 13th international conference on intelligent systems: theories and applications* (pp. 1-5).

Sharafaldin, I., Lashkari, A. H., & Ghorbani, A. A. (2018). Toward generating a new intrusion detection dataset and intrusion traffic characterization. *ICISSP*, 1, 108–116.

Shone, N., Ngoc, T. N., Phai, V. D., & Shi, Q. (2018). A deep learning approach to network intrusion detection. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2(1), 41–50.

Shone, N., Ngoc, T. N., Phai, V. D., & Shi, Q. (2018). A deep learning approach to network intrusion detection. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2(1), 41-50. https://doi.org/10.1109/TETCI.2017.2772792

Shurman, M., Khrais, R., & Yateem, A. (2020). DoS and DDoS attack detection using deep learning and IDS. *Int. Arab J. Inf. Technol*, *17*(4A), 655-661.

Smadi, S., Aslam, N., & Zhang, L. (2018). Detection of sophisticated DDoS attacks using machine learning algorithms. *Journal of Computer and Communications*, 6(6), 1–15.

Sommer, R., & Paxson, V. (2010). Outside the closed world: On using machine learning for network intrusion detection. *2010 IEEE Symposium on Security and Privacy*, 305–316.

StormWall. (2023). DDoS attacks in Q1 2023: Statistics and analysis. Retrieved from https://stormwall.network/

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, 30.

Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., Al-Nemrat, A., & Venkatraman, S. (2019). Deep learning approach for intelligent intrusion detection system. *IEEE Access*, 7, 41525-41550. https://doi.org/10.1109/ACCESS.2019.2895334

Wang, L., & Liu, Y. (2020, June). A DDoS attack detection method based on information entropy and deep learning in SDN. In *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)* (Vol. 1, pp. 1084-1088). IEEE.

Wang, S. Y., Wang, O., Zhang, R., Owens, A., & Efros, A. A. (2020). CNN-generated images are surprisingly easy to spot... for now. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 8695–8704.

Yadav, D. K., Chauhan, J., & Yadav, A. (2020). Intrusion detection using deep learning and adversarial networks. *International Journal of Computer Applications*, 176(33), 1–8.

Yin, C., Zhu, Y., Fei, J., & He, X. (2017). A deep learning approach for intrusion detection using recurrent neural networks. *IEEE Access*, 5, 21954–21961.

Yoon, J., Jarrett, D., & van der Schaar, M. (2019). Time-series generative adversarial networks. *Advances in Neural Information Processing Systems*, 32.

Yu, L., Zhang, W., Wang, J., & Yu, Y. (2017). SeqGAN: Sequence generative adversarial nets with policy gradient. *Proceedings of the AAAI Conference on Artificial Intelligence*, 31(1).

Zargar, S. T., Joshi, J., & Tipper, D. (2013). A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks. *IEEE Communications Surveys & Tutorials*, 15(4), 2046–2069.

Zhang, C., Luo, X., & Yang, Q. (2021). Transfer learning for intrusion detection using pre-trained CNN models. *Computers & Security*, 104, 102215.