ANKARA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

YÜKSEK LİSANS TEZİ

YAPAY ZEKA ALGORİTMALARI KULLANILARAK ANOMALİ TABANLI SALDIRI TESPİT SİSTEMİ

Melek ÖZSARI

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

ANKARA
2025

ÖZET

Yüksek Lisans Tezi

YAPAY ZEKA ALGORİTMALARI KULLANILARAK ANOMALİ TABANLI SALDIRI TESPİT SİSTEMİ

Melek ÖZSARI

Ankara Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı

Danışman: Prof. Dr. İman ASKERBEYLİ

Gelişen teknolojiyle birlikte çevrimiçi hizmetlerin yaygınlaşması, bireylerin ve kurumların dijital ortamdaki güvenliğini daha da önemli hale getirmiştir. Bu bağlamda, ağ trafiğini analiz ederek siber saldırıların tespiti, bilgi güvenliği alanında kritik bir araştırma konusudur. Bu çalışmada, gri kurt optimizasyonu ve parçacık sürü optimizasyonu algoritmalarının ağ trafiği verilerinde özellik seçimi başarısı incelenmiştir. Saldırı türleri ayrıntılandırılmamış, yalnızca ikili ("saldırı" veya "normal" olarak) sınıflandırma yapılmıştır. Öznitelik seçimi için USB_IDS_1 ve CSE-CIC-IDS2018 veri setleri üzerinde çalışılmış, sınıflandırma aşamasında karar ağacı algoritması kullanılmıştır. Modellerin performansları F1-skor metriği ile değerlendirilmiştir. Elde edilen deney sonuçlarına göre her iki yöntem de başarılı sonuçlar vermiştir. GWO, PSO'ya kıyasla daha az sayıda öznitelik seçerek daha verimli bir seçim sağlamıştır. Sonuçlar her iki algoritmanın ağ trafiği verileri üzerinde etkin öznitelik seçimi sağladığını ve düşük boyutlu veri ile sınıflandırma başarısının korunabildiğini göstermektedir. Ayrıca GWO ve PSO algoritmalarının saldırı tespiti sistemlerinde öznitelik seçimi başarısını aynı veri setleri üzerinde karşılaştırmalı olarak sunması açısından literatürde özgün bir yere sahiptir. Özellikle GWO'nun daha az öznitelikle yüksek başarı sağlaması, gerçek zamanlı ve kaynak kısıtlı sistemler için uygulanabilirliğini ortaya koymaktadır. Bu çalışma, saldırı tespiti sistemlerinin doğruluk ve verimliliğini artırmak adına önemli katkılar sunmaktadır.

Mayıs 2025, 28 sayfa
Anahtar Kelimeler: Ağ Trafiği Sınıflandırması, Saldırı Tespiti, Özellik Seçimi, GWO, PSO

ABSTRACT

Master Thesis

ANOMALY-BASED INTRUSION DETECTION SYSTEM USING ARTIFICIAL INTELLIGENCE ALGORITHMS

Melek ÖZSARI

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

Supervisor: Prof. Dr. İman ASKERBEYLİ

With the advancement of technology and the proliferation of online services, the security of individuals and institutions in the digital environment has become increasingly important. In this context, analyzing network traffic to detect cyberattacks is a critical research topic in the field of information security. In this study, the feature selection performance of the grey wolf optimization and particle swarm optimization algorithms was examined using network traffic data. Attack types were not detailed, and only binary classification ("attack" or "normal") was performed. Feature selection was conducted on the USB_IDS_1 and CSE-CIC-IDS2018 datasets, and the decision tree algorithm was used in the classification phase. The models' performances were evaluated using the F1-score metric. According to the experimental results obtained, both methods yielded successful results. GWO provided a more efficient selection by choosing fewer features compared to PSO. The results indicate that both GWO and PSO are effective in selecting relevant features from network traffic data, allowing classifiers to maintain strong performance even with reduced feature sets. This study contributes to the development of more efficient and accurate intrusion detection systems. The results show that both algorithms provide effective feature selection on network traffic data and maintain classification performance with low-dimensional input. Furthermore, the comparative evaluation of the feature selection performance of GWO and PSO on the same datasets gives this study a unique position in the literature. In particular, GWO's ability to achieve high performance with fewer features demonstrates its applicability in real-time and resource-constrained systems. This study offers significant contributions toward improving the accuracy and efficiency of intrusion detection systems.

May 2025, 28 pages
Key Words: Network Traffic Classification, Intrusion Detection, Feature Selection, GWO, PSO

TEŞEKKÜR

Tez çalışmam boyunca, rehberliğini, desteğini, bilgilerini esirgemeyen danışman Hocam Sayın Prof. Dr. İman ASKERBEYLİ'ye,

Tez jürimde bulunan, yorum ve önerileri ile tezimin şekillenmesine yardımcı olan sayın Dr. Öğr. Üyesi Ayhan AYDIN'a ve Prof. Dr. İhsan Tolga MEDENİ'ye,

Maddi manevi desteklerini esirgemeyen, her daim yanımda olan anneme, babama, ablam Dr. Öğr. Üyesi Şifa ÖZSARI'ya, sevgili eşim Veysel'e, neşe ve motivasyon kaynağım olan biricik kızım Mihra'ya en derin duygularla teşekkür ederim.

Melek ÖZSARI
Ankara, Mayıs 2025

İÇİNDEKİLER

SİMGELER DİZİNİ

$\sum n$	Toplam (Sum) İterasyon sayısı (Time step)

Kısaltmalar

COLAB	Google Colaboratory
DHR, DDoS	Dağıtılmış Hizmet Reddi (Distributed Denial of Service)
Dos	Denial of Service
DT	Decision Tree
FN	False Negative
FP	False Positive
FTP	File Transfer Protocol
GPU	Graphics Processing Unit
GWO	Grey Wolf Optimization
HTTP	Hypertext Transfer Protocol
HTTPS	Secure Hypertext Transfer Protocol
IP	Internet Protocol
Nİ, IoT	Nesnelerin İnterneti (Internet of Things)
ÖÖE, RFE	Özyinelemeli Özellik Eleme (Recursive Feature Elimination)
PSO	Particle Swarm Optimization
RFE	Recursive Feature Elimination
SQL	Structured Query Language
SSH	Secure Shell
STS, IDS	Saldırı Tespit Sistemleri (Intrusion Detection Systems)
TCP	Transmission Control Protocol
TN	True Negative
TP	True Positive
XSS	

ŞEKİLLER DİZİNİ

ÇİZELGELER DİZİNİ

1. GİRİŞ

Günümüzde dijitalleşmenin hız kazanmasıyla birlikte çevrimiçi hizmetler bireylerin ve kurumların günlük yaşantılarında vazgeçilmez bir yer edinmiştir. E-devlet uygulamaları, uzaktan eğitim sistemleri, finansal işlemler ve endüstriyel otomasyon gibi birçok alanda internet temelli hizmetlerin kullanımı yaygınlaşmış, bu durum ağ güvenliğini bilgi teknolojilerinin öncelikli araştırma konularından biri hâline getirmiştir. Ancak, çevrimiçi hizmetlerdeki bu artış, aynı zamanda siber tehditlerin de daha karmaşık ve sık görülen hâle gelmesine neden olmuştur.

Saldırı tespit sistemleri ağ trafiği içerisindeki anormal davranışları ve potansiyel tehditleri belirlemek amacıyla geliştirilmiş önemli araçlardır. Geleneksel imza tabanlı IDS yapıları, yalnızca daha önce tanımlanmış tehditleri tanıyabildikleri için, yeni nesil saldırılar veya bilinmeyen tehditler karşısında yetersiz kalabilmektedir (Sommer ve Paxson 2010). Bu nedenle son yıllarda, yapay zekâ ve makine öğrenimi tabanlı yaklaşımlara dayalı saldırı tespit yöntemleri ön plana çıkmıştır.

Makine öğrenimi tabanlı sistemlerin başarımı, büyük ölçüde kullanılan veri kümesinin kalitesi ve öznitelik (feature) yapısıyla doğrudan ilişkilidir. Özellikle yüksek boyutlu veri kümelerinde gereksiz veya alakasız özniteliklerin bulunması, sınıflandırma modellerinin doğruluk oranını düşürmekte ve işlem maliyetlerini artırmaktadır (Guyon ve Elisseeff 2003). Bu sorunu çözmek amacıyla algoritma uygulanmadan önce veriler üzerinde özellik seçimi uygulanmaktadır.

Özellik seçimi genel bir ifadeyle modelin performansını artırmak ve işlem maliyetlerini düşürmek için gereksiz veya az katkı sağlayan özelliklerin veri setinden kaldırılması, diğer bir değişle sonuç üzerinde en etkili olan özelliklerin seçilmesidir. Bunun için çeşitli öznitelik seçimi (feature selection) yaklaşımları, filtre yöntemleri, sarma yöntemleri, gömülü yöntemler, mevcuttur:

- Filtre yöntemlerinde özellikler korelasyon katsayısı, istatistiksel yöntemler vs. kullanılarak bağımsız olarak değerlendirilir. Bağımsız olarak her bir özellik değerlendirilir ve bu değerlere göre en iyi özellikler seçilir. Modelden bağımsız çalışan bir yaklaşımdır.
- Sarma yöntemlerinde ise modelin performansına göre özellikler seçilir. Belirli bir model kullanılarak, özellik alt kümeleri test edilir ve model performansına göre özellikler seçilir. İleri seçim, geri eleme ve ÖÖE yaygın olarak kullanılan sarma yöntemleridir.
- Gömülü yöntemlerde ise özellik seçimi modelin eğitim sürecine entegre edilir.

Geleneksel yöntemlerin yanı sıra, bu alanda doğadan esinlenen meta-sezgisel (metaheuristic) optimizasyon yöntemleri de önemli başarılar göstermektedir.

Yürütülen bu çalışmada, saldırı tespiti amacıyla kullanılan USB_IDS_1 (Catillo vd. 2021) ve CSE- CIC-IDS2018 (Canadian Institute for Cybersecurity 2018) veri kümeleri üzerinde özellik seçiminde gri kurt optimizasyonu (Mirjalili ve diğ. 2014) ve parçacık sürü optimizasyonu (Kennedy ve Eberhart 1995) algoritmalarının başarıları incelenmiştir. Elde edilen özellikler ile ağ trafiği verileri, karar ağacı algoritması kullanılarak "saldırı" veya "normal" şeklinde ikili (binary) olarak sınıflandırılmıştır (Quinlan 1986). Saldırı tiplerini dikkate alarak ayrı ayrı deneyler yapıldığında elde edilen sonuçların başarısı (Özsarı ve ark. 2024) yapmış olduğu çalışmada olduğu gibi yetersiz kalmıştır. Algoritmaların başarısı, sınıflar arasında daha güvenilir değer veren F1-skor metriği ile incelenmiştir (Chinchor 1992).

Elde edilen deney sonuçlarına göre, öznitelik seçimi yapılan veri kümelerinde sınıflandırma performansının korunduğunu ve sistemin etkinliğinin arttığı açık bir şekilde görülmüştür. Her iki algoritma da başarılı sonuçlar vermekle birlikte, GWO daha az sayıda öznitelikle daha verimli bir seçim sağlamıştır. Ayrıca, bu çalışma ile saldırı tespit sistemlerinde özellik seçiminin etkisini değerlendirmek ve GWO ile PSO algoritmalarının karşılaştırmalı başarımını ortaya koymak açısından literatüre katkı sağlanması amaçlanmaktadır.

Tezin geri kalanı şu şekilde yapılandırılmıştır:

- İkinci kısımda kullanılan veri setleri ve algoritmalara ilişkin ayrıntılı bilgiler sunulmuştur.
- Üçüncü bölüm, parametre ayarları, gerçekleştirilen deneyler ve bu deneylerden elde edilen sonuçların ayrıntılı incelenmesine odaklanmaktadır.
- Son olarak, dördüncü bölümde elde edilen bulgular değerlendirilmiş ve gelecekte yapılabilecek çalışmalara yönelik öneriler paylaşılmıştır.

2. GEREÇ VE YÖNTEMLER (MATERIALS AND METHODS)

Bu araştırmada, ağ trafiği verilerinden saldırı tespitine yönelik olarak ikili sınıflandırma modeli geliştirilmiştir. Kullanılan veri setleri, algoritmalar ve performans metriği bu bölümde ayrıntılı şekilde açıklanmıştır.

2.1 Veri Setleri (Data Sets)

STS üzerine yapılan akademik çalışmalar ve analizlerde literatürde sıkça kullanılan çeşitli veri setleri bulunmaktadır. Bu veri setlerinden bazıları KDD Cup 99 (Tavallaee vd. 2009), NSL-KDD (Tavallaee vd. 2009) ve UNSW-NB15 (Moustafa ve Slay 2015) şeklindedir. Fakat bu veri setlerinden çoğunluğu güncelliğini yitirmiş olup, gerçek dünya saldırılarını tam anlamıyla yansıtamamaktadır. Bu tez çalışmasında, daha güncel ve farklı saldırı türlerini içeren iki veri seti tercih edilmiştir: USB_IDS_1 ve CSE-CIC-IDS2018. CSE-CIC-IDS2018, literatürde geniş kabul görmüş bir benchmark olup, çok çeşitli saldırı türlerini kapsamaktadır (örneğin: DDoS, brute-force, botnet saldırıları). USB_IDS_1 ise özellikle USB üzerinden gerçekleştirilen saldırılara odaklanarak bu alandaki boşluğu doldurmaktadır. Her iki veri seti de gerçek dünya senaryolarına yakın, güncel ve çeşitli saldırıları içermesi bakımından bu çalışmada tercih edilmiştir. Ayrıca veri setlerinin orta düzeydeki boyutları hem modelin öğrenme yeterliliğini sağlamış hem de hesaplama maliyetini makul seviyede tutmuştur.

USB_IDS_1 veri seti, USB portu üzerinden gerçekleştirilen saldırıları ve normal trafiği içermekte olup, siber güvenlik araştırmalarında kullanılmak üzere geliştirilmiş bir veri setidir. CSE-CIC-IDS2018 ise, CIC (Kanada Siber Güvenlik Enstitüsü) tarafından oluşturulan, çeşitli saldırı türlerini ve normal trafiği içeren kapsamlı bir ağ trafiği veri setidir. Bu veri seti, gerçek zamanlı ağ ortamında simüle edilmiş trafik verisini içermektedir.

2.1.1 USB_IDS_1 veri seti

USB_IDS_1 veri seti, USB portu üzerinden gerçekleştirilen saldırıları analiz edebilmek amacıyla oluşturulmuş özel bir veri kümesidir. Klasik ağ trafiği tabanlı saldırı tespit sistemlerinden farklı olarak, fiziksel donanım temelli tehditleri modellemeyi hedeflemektedir. Bu yönüyle, donanımsal ara yüzler üzerinden gelebilecek tehlikelerin algılanmasına katkı sağlayan nadir çalışmalardan biri olarak öne çıkmaktadır.

Veri seti, bir sistemde çalışan çeşitli uygulamaların, servislerin ve sistem bileşenlerinin etkileşimleri sırasında oluşturduğu olay kayıtlarından oluşmaktadır. Bu kayıtlar arasında:

- USB bağlantı ve aygıt tanıma işlemleri
- Sistem log girdileri
- İzin ve erişim davranışları
- Kullanıcı müdahaleleri ve otomatik sistem yanıtları yer almaktadır.

Veri seti 83 farklı özelliğe sahip ve toplamda 16 sınıf içermektedir. Bu özellikler: *"Flow ID, Src IP, Src Port, Dst IP, Dst Port, Protocol, Timestamp, Flow Duration, Total Fwd Packet, Total Bwd packets, Total Length of Fwd Packet, Total Length of Bwd Packet, Fwd Packet Length Max, Fwd Packet Length Min, Fwd Packet Length Mean, Fwd Packet Length Std, Bwd Packet Length Max, Bwd Packet Length Min, Bwd Packet Length Mean, Bwd Packet Length Std, Flow Bytes/s ,Flow Packets/s, Flow IAT Mean, Flow IAT Std, Flow IAT Max, Flow IAT Min, Fwd IAT Total, Fwd IAT Mean, Fwd IAT Std, Fwd IAT Max, Fwd IAT Min, Bwd IAT Total ,Bwd IAT Mean, Bwd IAT Std, Bwd IAT Max, Bwd IAT Min, Fwd PSH Flags, Bwd PSH Flags, Fwd URG Flags, Bwd URG Flags, Fwd Header Length, Bwd Header Length, Fwd Packets/s, Bwd Packets/s, Packet Length Min, Packet Length Max, Packet Length Mean, Packet Length Std, Packet Length Variance, FIN Flag Count, SYN Flag Count, RST Flag Count, PSH Flag Count, ACK Flag Count, URG Flag Count, CWR Flag Count, ECE Flag Count, Down/Up Ratio, Average Packet Size, Fwd Segment Size Avg, Bwd Segment Size Avg ç, Fwd Bytes/Bulk Avg, Fwd Packet/Bulk Avg, Fwd Bulk Rate Avg, Bwd Bytes/Bulk Avg, Bwd Packet/Bulk Avg, Bwd Bulk Rate Avg, Subflow Fwd Packets, Subflow Fwd Bytes, Subflow Bwd Packets, Subflow Bwd Bytes, FWD Init Win Bytes, Bwd Init Win Bytes, Fwd Act Data Pkts, Fwd Seg Size Min, Active Mean, Active Std, Active Max, Active Min, Idle Mean, Idle Std, Idle Max, Idle*

Min, Label" (Özsarı ve ark. 2024) şeklinde verilmektedir.

Bu veri seti, özellikle fiziksel tehditleri tespit etmek üzere geliştirilen yeni nesil saldırı tespit yaklaşımlarında temel oluşturabilecek nitelikte olup, siber güvenlik literatüründe USB tabanlı IDS çözümleri için önemli bir kaynak teşkil etmektedir.

USB_IDS_1 veri seti, son yıllarda birçok akademik çalışmada saldırı tespit sistemlerinin değerlendirilmesinde kullanılmıştır. Catillo ve arkadaşları, bu veri setinin çok katmanlı yapısını kullanarak saldırı tespit sistemlerinin performansını değerlendirmeye yönelik deneyler gerçekleştirmiş ve veri setinin IDS testleri için zengin bir ortam sunduğunu ortaya koymuştur (Catillo ve ark. 2022). Benzer şekilde, USB_IDS_1 veri seti ile çok katmanlı saldırı sınıflandırması yapan çalışmalar da mevcuttur (Dang 2022). Özsarı ve arkadaşları ise veri setindeki öznitelikleri genetik algoritmalar aracılığıyla azaltmış ve sınıflandırma başarılarını artırmaya yönelik analizler gerçekleştirmiştir (Özsarı ve ark. 2024). Tüm bu çalışmalar göz önünde bulundurulduğunda, USB_IDS_1 veri seti donanımsal kaynaklı tehditlerin tespitinde önemli bir referans veri kümesi haline gelmiştir.

2.1.2 CSE-CIC-IDS2018 veri seti

CSE-CIC-IDS2018 veri seti, Kanada Siber Güvenlik Enstitüsü tarafından CSE işbirliğiyle geliştirilmiş, gerçek zamanlı ağ trafiğini içeren ve çok çeşitli siber saldırı türlerini barındıran kapsamlı bir veri setidir. Veri seti, 7 günlük bir süre boyunca çeşitli ağ trafiği senaryoları simüle edilerek oluşturulmuştur. Bu süreçte 50 saldırgan makine ve 420 hedef sistem yer almıştır. Toplanan trafik, CICFlowMeter aracı kullanılarak akış tabanlı özelliklere dönüştürülmüş ve her bir akış kaydı yaklaşık 80 öznitelik ile tanımlanmıştır (Canadian Institute for Cybersecurity, 2018). Bu özellikler: *"fl_dur, tot_fw_pk, tot_bw_pk, tot_l_fw_pkt, fw_pkt_l_max, fw_pkt_l_min, fw_pkt_l_avg, fw_pkt_l_std, Bw_pkt_l_max, Bw_pkt_l_min, Bw_pkt_l_avg, Bw_pkt_l_std, fl_byt_s, fl_pkt_s, fl_iat_avg, fl_iat_std, fl_iat_max, fl_iat_min, fw_iat_tot, fw_iat_avg, fw_iat_std,*

fw_iat_max, fw_iat_min, bw_iat_tot, bw_iat_avg, bw_iat_std, bw_iat_max, bw_iat_min, fw_psh_flag, bw_psh_flag, fw_urg_flag, bw_urg_flag, fw_hdr_len, bw_hdr_len, fw_pkt_s, bw_pkt_s, pkt_len_min, pkt_len_max, pkt_len_avg, pkt_len_std, pkt_len_va, fin_cnt, syn_cnt, rst_cnt, pst_cnt, ack_cnt, urg_cnt, cwe_cnt, ece_cnt, down_up_ratio, pkt_size_avg, fw_seg_avg, bw_seg_avg, fw_byt_blk_avg, fw_pkt_blk_avg, fw_blk_rate_avg, bw_byt_blk_avg, bw_pkt_blk_avg, bw_blk_rate_avg, subfl_fw_pk, subfl_fw_byt, subfl_bw_pkt, subfl_bw_byt, fw_win_byt, bw_win_byt, Fw_act_pkt, fw_seg_min, atv_avg, atv_std, atv_max, atv_min, idl_avg, idl_std, idl_max, idl_min" şeklindedir.

Veri seti; Brute-force saldırıları (SSH, FTP), Web saldırıları (SQL injection, XSS), Hizmet Engelleme (DoS, DDoS), Botnet aktiviteleri, Heartbleed güvenlik açığı gibi birçok farklı saldırı türünü içermektedir. Bu çok yönlü yapı, veri setini hem denetimli hem de denetimsiz öğrenme modellerinin test edilmesi için oldukça uygun bir hâle getirmiştir. Veri setindeki saldırılar, savunma stratejileri geliştirmek için farklı kategorilere ayrılmıştır:

a) DoS ve DDoS Saldırıları

- DoS slowloris: Sunucuyu yavaşlatmak için bağlantıları açık bırakma saldırısı.
- DoS Slowhttptest: HTTP bağlantılarını kötüye kullanarak yapılan DoS saldırısı.
- DoS Hulk: Sunucuyu yoğun taleplerle aşırı yükleme saldırısı.
- DoS GoldenEye: Sunucuyu isteklerle tıkanmaya zorlama saldırısı.
- DDoS: Çok sayıda cihazla yapılan dağıtılmış hizmet engelleme saldırısı.

b) Brute Force (Kaba Kuvvet) Saldırıları

- FTP-Patator: FTP protokolüne yönelik kaba kuvvet girişimleri.
- SSH-Patator: SSH oturum açma bilgilerini kaba kuvvet yöntemiyle tahmin etme.

c) Web Tabanlı Saldırılar

- SQL Injection: Veritabanı sorgularını manipüle eden saldırılar.
- XSS (Cross-Site Scripting): Zararlı JavaScript kodlarının enjekte edilerek

sistemlerin hedef alındığı bir saldırı türüdür.

- Command Injection: Sistem komutlarını kötüye kullanma saldırıları.

d) Botnet

- Bot: Bir ağdaki cihazları kötü amaçlı yazılımlarla ele geçirme.

e) Infiltration

- Infiltration: İç ağdan yetkisiz veri sızdırma girişimleri.

f) Heartbleed

- Heartbleed: OpenSSL güvenlik açığını istismar eden saldırılar.

Burada tipler kısaca aşağıdaki gibi açıklanabilir:

- BENIGN: Bu etiket, saldırı içermeyen normal ağ trafiğini ifade eder. Saldırı analizi sırasında, bu tür trafik, modelin doğru şekilde normal ve saldırı trafiğini ayırt etmesini sağlamak için kullanılır.
- DoS slowloris: Sunucuya birçok yavaş bağlantı isteği gönderilir, bu da sunucunun her bağlantıyı açık tutmasına ve sistem kaynaklarını tükenmesine neden olur. Bu tür bir saldırı, sunucunun ağ trafiğini işleyememesine yol açar.
- DoS Slowhttptest: HTTP başlıkları yavaşça gönderilerek sunucunun kaynakları tükenmeye zorlanır. Bu saldırı, sunucunun gelen HTTP isteklerini işleyememesi ve yanıt verememesiyle sonuçlanır.
- DoS Hulk: Sunucuya büyük miktarda HTTP istekleri gönderilir, böylece sunucunun işlem gücü ve bant genişliği tüketilir. Bu saldırı, hedef web sunucusunun yanıt verememesiyle sonuçlanır.
- DoS GoldenEye: Web sunucusuna yoğun miktarda HTTP GET ve POST isteği gönderilir. Bu tür bir saldırı, sunucunun işlem kaynaklarını tüketerek erişilemez hale gelmesine yol açar.
- DDoS: Birden fazla kaynaktan yapılan, koordineli bir saldırıdır. Bu saldırılar, ağ üzerindeki hedefe aşırı yük bindirir, böylece hedefin hizmet verememesi sağlanır.
  - FTP-Patator: FTP protokolü kullanılarak, çeşitli kullanıcı adı ve şifre

kombinasyonlarıyla kaba kuvvet saldırısı gerçekleştirilir. Amaç, doğru kimlik bilgilerini bularak hedef sisteme erişim sağlamaktır.

- SSH-Patator: SSH oturum açma bilgilerini kaba kuvvet yöntemiyle tahmin etmeye çalışır. Bu saldırı, sistemin güvenliğini aşmak ve yetkisiz erişim sağlamak için yapılır.
- SQL Injection: Web uygulamalarında veri tabanına zarar vermek amacıyla, kullanıcı giriş alanlarına zararlı SQL sorguları enjekte edilir. Bu, veri tabanı verilerini çalmaya veya değiştirmeye yönelik bir saldırıdır.
- XSS (Cross-Site Scripting): Web sayfalarına zararlı JavaScript kodları enjekte edilerek, kullanıcıların bilgilerini çalmak veya web sayfasının işleyişini bozmak amaçlanır. Bu saldırı, genellikle web uygulamalarında güvenlik açıklarını hedef alır.
- Command Injection: Bir web uygulaması veya sistemin komut satırına zararlı komutlar enjekte edilerek, sistem üzerinde yetkisiz işlemler yapılır. Bu saldırı, sunucuya zararlı komutlar göndererek, sistemin kontrolünü ele geçirmeyi amaçlar.
- Bot: Birden fazla cihazı kötü amaçlı yazılımlar aracılığıyla ele geçirerek, bu cihazları bir ağda botnet oluşturmak için kullanır. Botnet, çeşitli saldırılar için kullanılabilir, özellikle DDoS saldırılarında.
- Infiltration: Sisteme yetkisiz erişim sağlamak amacıyla iç ağlardan veya dışarıdan yapılan saldırılardır. Bu saldırı, ağdaki güvenlik açıklarını kullanarak bilgi sızdırmayı hedefler.
- Heartbleed: OpenSSL kütüphanesindeki bir güvenlik açığını hedef alarak, şifreli bağlantılarda hassas verilerin sızmasına neden olan bir saldırıdır. Bu, özellikle HTTPS protokolü üzerinden yapılan güvenli iletişimde veri hırsızlığına yol açabilir.

CSE-CIC-IDS2018 veri seti, siber güvenlik alanında yapılan pek çok çalışmada tercih edilmiştir. Karaman bu veri seti ile karar ağacı tabanlı bir saldırı tespit modeli geliştirmiş ve %99,97 doğruluk oranı elde etmiştir (Karaman 2020). Başka bir çalışmada, Soltani ve arkadaşları derin öğrenme temelli bir sistem ile bu veri seti üzerinde başarılı sonuçlara ulaşmıştır (Soltani vd. 2020). Kalutharage ve arkadaşları tarafından gerçekleştirilen bir çalışmada, IoT güvenlik izleme için derin otomatik kodlayıcı modelleri açıklayıcı yapay

zekâ yöntemleriyle birleştirilerek yenilikçi bir yaklaşım sunulmuştur (Kalutharage vd. 2022). Bu yaklaşımın temel amacı, makine öğrenimi tabanlı saldırı tespit sistemlerinin güvenilirliğini ve doğruluğunu artırmaktır. Geliştirilen model, CSE-CIC-IDS2018 veri seti üzerinde test edilerek değerlendirilmiştir.

Bu bağlamda, CSE-CIC-IDS2018 veri seti, farklı saldırı türlerinin modellenmesine olanak sağlayan ve araştırmacılar tarafından sıkça tercih edilen bir veri kaynağı olma özelliği taşımaktadır.

2.2 Karar Ağacı (Decision Tree)

Makine öğrenimi, bilgisayar sistemlerinin geçmiş verilere dayanarak öğrenme ve gelişim göstermesini sağlayan yapay zekâ alt alanlarından biridir. Özellikle büyük ve karmaşık veri kümelerinde, makine öğrenimi algoritmaları verilerdeki örüntüleri tanımlayarak tahmin yapma veya karar verme süreçlerinde kullanılmaktadır. Bu süreç, verilerin analiz edilmesi ve bu verilerden anlamlı bilgi çıkarılması yoluyla gerçekleştirilir. Bunun sonucunda sistem, zaman içinde daha doğru ve etkili kararlar verebilme yeteneği kazanır. Ağ trafiği analizi ve saldırı tespiti gibi karmaşık problemlerde, verilerin sürekli olarak güncellenmesi ve çeşitlenmesi nedeniyle, makine öğrenimi tabanlı yaklaşımlar önemli avantajlar sunmaktadır.

Yürütülen bu çalışamada saldırı tespiti amacıyla Denetimli Öğrenme (Supervised Learning) algoritması kullanılmıştır. Denetimli öğrenmede, algoritmalar, etiketlenmiş veriler kullanarak öğrenir. Her veri örneği, bir etiketle ilişkilendirilir ve model, bu etiketlerin doğruluğunu tahmin etmeye çalışır. Saldırı tespiti sistemlerinde sıkça kullanılan bu yaklaşım, “saldırı” ve “normal” olmak üzere ikili sınıflandırma problemlerinde oldukça etkilidir.

Bu çalışmada, USB_IDS_1 ve CSE-CIC-IDS2018 veri setleri kullanılarak ağ trafiğinin “saldırı” veya “normal” olarak sınıflandırılması hedeflenmiş ve popüler bir denetimli

öğrenme yaklaşımı olan karar ağaçları kullanılmıştır. Gri kurt optimizasyonu ve parçacık sürü optimizasyonu algoritmaları ile öznitelik seçimi gerçekleştirilmiş; daha sonra karar ağaçları ile sınıflandırma yapılmıştır. Karar ağacının tercih edilme nedeni, eğitim süresinin hızlı olması ve model çıktılarının kolay yorumlanabilir yapıda olmasıdır. Saldırı tespiti gibi kritik kararlar içeren sistemlerde, karar ağaçlarının sunduğu açıklanabilirlik avantajı önemli bir tercih sebebidir. Ayrıca, düşük hesaplama maliyeti ve görsel izlenebilirlik gibi özellikleri sayesinde makine öğrenimi tabanlı güvenlik uygulamalarında sıkça kullanılmaktadır.

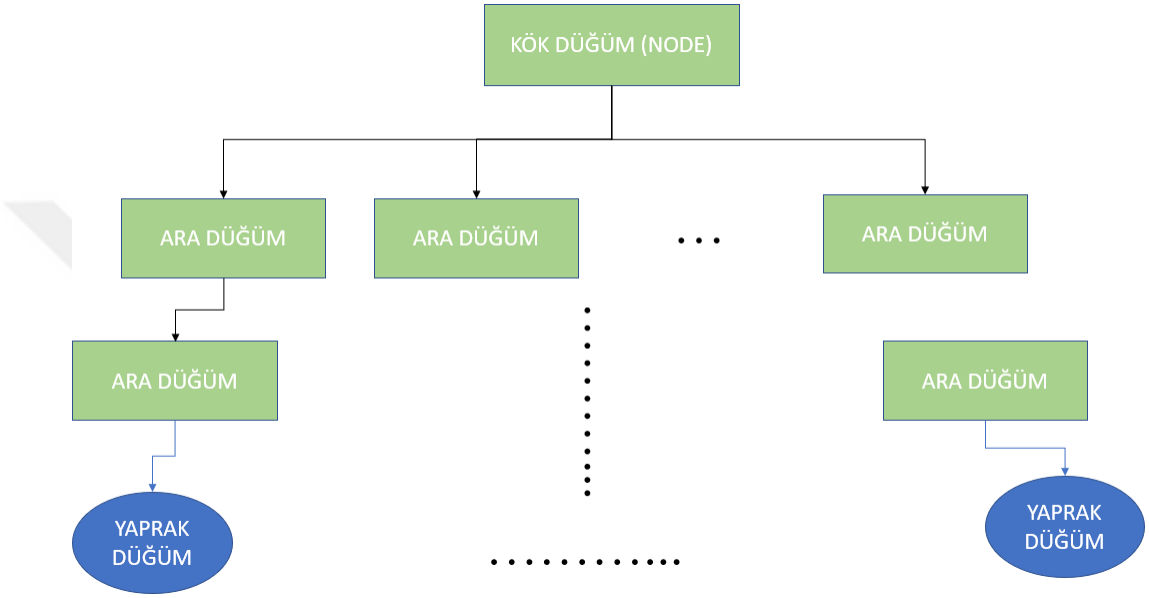
Karar ağacı (Quinlan 1986), sınıflandırma ve regresyon problemlerinde yaygın olarak kullanılan uygulaması basit bir denetimli öğrenme algoritmasıdır. Bu yöntem, veri setindeki örüntüleri analiz ederek, her bir örneği bir ağaç yapısı içerisinde sıralayarak kararlar alır. Bu yapı, bir akış şeması biçiminde düzenlenmiştir; iç düğümler veri özniteliklerine yönelik testleri ifade eder, dallar bu testlerin olası sonuçlarını yansıtır ve yaprak düğümler sınıf etiketlerini temsil eder. Karar ağacı, eğitim verilerini belirli bir durdurma noktasına (örneğin, ağacın maksimum derinliği veya her düğüm için minimum örnek sayısı gibi) ulaşana kadar öznitelik değerlerine göre alt gruplara böler. Bu süreçte, bölme işlemi genellikle entropi ya da Gini impurity gibi ölçütlerle gerçekleştirilir. Bu ölçütler, veri alt kümelerindeki belirsizliği ya da rastlantıyı değerlendirir ve amaç, bölme sonucunda bilgi kazancını ya da belirsizliği azaltmayı en üst düzeye çıkaracak özniteliği bulmaktır. Bu işlem, alt kümelere ayrılma süreci her adımda tekrar edilerek, yani rekürsif bir şekilde yapılır. Rekürsiyon, bir alt küme içindeki tüm örnekler aynı hedef değeri taşıdığında veya daha fazla bölme işlemi anlamlı bir tahmin yapmadığında son bulur. Bu araştırmada, ağacın bölünmesi entropi ile bilgi kazancı hesaplamasına göre gerçekleştirilmiştir.

Entropi, bir veri kümesindeki belirsizlik veya rastgelelik derecesini ölçen bir ölçüttür. Sınıflandırma problemlerinde, sınıf etiketlerinin veri kümesindeki dağılımına bağlı olarak bu rastgelelik hesaplanır. Bir alt küme için entropi, ilgili düğümdeki S sınıfının dağılımına göre şu şekilde tanımlanabilir:

$$E(S) = -\sum_{i=1}^{n} p_i \backslash log_2(p_i) \quad (2.1)$$

Burada E(S), bir veri kümesinin entropisini, n, veri kümesindeki sınıf sayısını, pi i. sınıfa ait örneklerin toplam veri kümesi içindeki oranını ifade etmektedir.

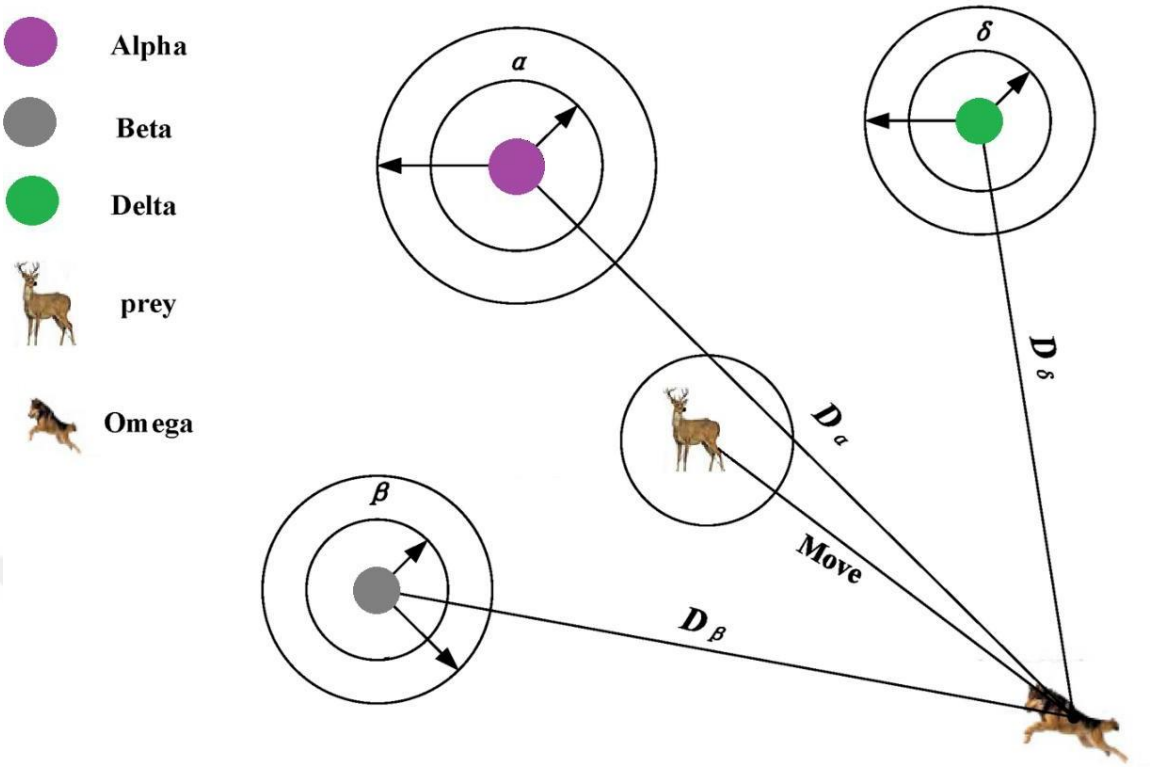
Bir karar ağacının temel yapısı **Şekil 2.1**'de gösterilmiştir.



Şekil 2.1 Karar ağaçları

2.3 Gri Kurt Algoritması (Gray Wolf Algorithm)

Gri kurt algoritması, doğadaki gri kurtların avlanma stratejilerinden esinlenerek tasarlanmış bir optimizasyon yöntemidir. Bu algoritma, evrimsel bir algoritma türüdür ve sürü tabanlı bir yaklaşım kullanır. Gri kurtlar, sürü halinde hareket eder ve liderlik yaparak belirli hedeflere yönelirler. **Şekil 2.2**'de GWO algoritmasının temel çalışma prensibi gösterilmiştir.



Şekil 2.2 GWO algoritması (Zhao vd. 2019)

GWO, bu doğal avlanma ve liderlik davranışını, çoklu çözüm arayışında ve özellikle optimizasyon problemlerinde çözüm bulma amacıyla kullanır. Gri kurt algoritması, temel olarak üç ana hareket davranışına dayanır:

- Keşif (Exploration): Gri kurtlar, çözüm alanındaki en iyi çözüme ulaşmak için keşif yaparlar. Bu, genellikle sürüdeki liderin (alpha kurt) önderliğinde, yeni çözümleri aramaya yönelik bir davranıştır.
- Sömürme (Exploitation): Yeterli keşif yapıldıktan sonra, çözümler daha derinlemesine incelenir. Bu, sürüdeki daha deneyimli bireylerin ve liderlerin yönlendirmesiyle olur.
- Önderlik (Leadership): Gri kurt sürüsünde genellikle dört tür kurt vardır: Alpha, Beta, Delta ve Omega kurtları. Her biri farklı seviyelerde liderlik yapar ve sürüdeki hareketleri yönlendirir.

Algoritma, çalışma adımları şu şekildedir:

- Başlangıç Popülasyonu: Algoritma, başlangıç popülasyonunu rastgele bir şekilde oluşturur. Bu popülasyon, çözüm alanındaki rastgele kurtların konumlarını temsil

eder.

- Değerlendirme: Her bir kurdun performansı, uygunluk (fitness) fonksiyonu ile değerlendirilir. Bu fonksiyon, her bireyin çözümünün kalitesini ölçer.
  - Güncelleme ve Lider Seçimi:
    - Algoritma, popülasyon içindeki en iyi kurtları seçer. Genellikle en iyi üç kurt seçilir: Alpha, Beta ve Delta kurtları.
    - Bu kurtlar, diğer tüm kurtlara rehberlik eder. En iyi kurt (alpha) sürüyü yönlendirir, ardından beta ve delta kurtları gelir.
  - Pozisyon Güncelleme: Algoritmanın temel güncelleme adımıdır. Her bir kurdun pozisyonu (yani çözüm), en iyi üç liderin pozisyonları ile karşılaştırılarak güncellenir. Her bir kurdun yeni konumunu belirlemek için aşağıdaki denklemler kullanılır:

$$D = |C.X_p(t) - X(t)| \tag{2.2}$$

$$X(t+1) = |X_p(t) - A.D| \tag{2.3}$$

Burada, X(t) kurdun mevcut konumunu, $X_p(t)$ Alpha kurdunun konumunu, D mesafeyi, A iletişim katsayısını (belli bir rastgele değere sahiptir), C1 iletişim katsayısını (rastgele seçilir) ifade etmektedir. Bu güncelleme, kurtların çözüme daha yakın olabilmesi için daha iyi bir yönelim sağlar.

  - Sonuçların Hesaplanması ve Sonlandırma: Algoritma, belirli bir durdurma kriterine ulaştığında sona erer. Genellikle bu kriterler belirli bir nesil sayısına veya yeterli iyileştirme sağlanmayan bir noktaya ulaşılmasıdır.

2.4 Parçacık Sürü Optimizasyonu (Particle Swarm Optimization)

Parçacık sürü optimizasyonu 1995 yılında Kennedy ve Eberhart tarafından önerilen ve doğadaki kuş sürüsü veya balık topluluklarının toplu hareketlerinden esinlenilerek geliştirilmiş bir meta-sezgisel optimizasyon algoritmasıdır (Kennedy ve Eberhart 1995). PSO, çözüm uzayında optimal çözüme ulaşmak için bir grup "parçacığın" (particle)

birlikte hareket ettiği sürü tabanlı bir yaklaşımdır. Her bir parçacık, çözüm uzayında bir olasılığı temsil eder ve zamanla en iyi çözüme yönelir.

Bu algoritma, hem keşif (exploration) hem de sömürü (exploitation) özelliklerini dengeli bir şekilde kullanarak, çok boyutlu problemlerde etkili çözümler üretme yeteneğine sahiptir. PSO algoritması aşağıdaki temel kavramlara dayanır:

- Parçacık (Particle): Çözüm uzayındaki her bir aday çözümü temsil eder.
- Hız (Velocity): Her parçacığın bir sonraki pozisyonunu belirleyen hareket yönü ve büyüklüğüdür.
- Konum (Position): Parçacığın mevcut çözüm uzayındaki konumudur.
- Kişisel en iyi konum (pBest): Parçacığın şimdiye kadar ulaştığı en iyi konumdur.
- Global en iyi konum (gBest): Tüm sürüdeki en iyi çözüm konumudur.

Algoritmanın çalışma adımları şu şekildedir:

- Başlangıç Popülasyonu: Rastgele konum ve hız değerlerine sahip parçacıklar oluşturulur. Uygunluk Değerlendirmesi: Her parçacığın çözüm kalitesi bir uygunluk (fitness) fonksiyonu ile değerlendirilir.
- pBest ve gBest Güncellenmesi: Her parçacık kendinin o ana kadar bulduğu en iyi konumunu (pBest) ve sürünün en iyi çözümünü (gBest), diğer bir ifadeyle tüm parçacıkların içinde en iyi sonucu saklar.
- Hız ve Konum Güncellemesi: Parçacıkların pozisyonları aşağıdaki denklemlerle güncellenir:

$$v_i^{t+1} = w \cdot v_i^t + c_1 \cdot r_1 \cdot \left(p_i^{best} - x_i^t\right) + c_2 \cdot r_2 \cdot \left(g^{best} - x_i^t\right) \quad (2.4)$$

$$x_i^{t+1} = x_i^t + v_i^{t+1} \quad (2.5)$$

- x_i^t : Parçacık i'nin t. iterasyondaki konumunu ifade eder. Bu konum, çözüm uzayındaki bir olası çözüme karşılık gelir.
- v_i^t: Parçacık i'nin t. iterasyondaki hızını gösterir. Hız, parçacığın bir sonraki iterasyonda ne kadar ve hangi yönde hareket edeceğini belirler.
- w: Atalet katsayısıdır. Parçacığın önceki hızını ne ölçüde koruyacağını belirler.

Genellikle keşif ve sömürü arasındaki dengeyi ayarlamak için kullanılır.

- c_1: Bilişsel katsayıdır. Parçacığın kendi en iyi pozisyonuna yönelme eğilimini belirler.
- c_2: Sosyal katsayıdır. Parçacığın sürünün en iyi pozisyonuna yönelme eğilimini belirler.
- r_1, r_2: [0, 1] aralığında rastgele üretilen değerlerdir. Bu sayede algoritma stokastik (rastlantısal) davranış göstererek çözüm uzayını daha etkili şekilde tarayabilir.
- Döngünün Devamı: İterasyon sayısı doluncaya ya da çözüm yeterince iyi hâle gelinceye kadar güncelleme işlemleri tekrarlanır.

PSO, uygulanması kolay, sade bir yapıya sahip bir algoritmadır, parametre sayısı azdır ve ayarlanması görece kolaydır, sürekli ve ayrık problemler için uyarlanabilirdir ve hızlı yakınsama özelliği vardır. Bu niteliklerinden dolayı kısa sürede iyi sonuçlar elde edilebilir ve sıklıkla kullanılır.

2.5 Başarı Metriği

Sınıflandırma performansının değerlendirilmesinde ve GWO ile PSO'da amaç fonksiyonu olarak F1-skor metriği kullanılmıştır. F1-skor, sınıflandırma problemlerinde özellikle sınıf dağılımının dengesiz olduğu durumlarda, hem kesinlik (precision) hem de duyarlılık (recall) ölçütlerini dengeleyerek daha anlamlı bir değerlendirme sunar. Saldırı tespiti gibi sistemlerde, bir modelin yalnızca genel doğruluğu değil, aynı zamanda saldırıları doğru şekilde tespit etme kabiliyeti (recall) ve yanlış alarm üretmeme başarısı (precision) kritik önem taşır. Bu nedenle, F1-Skor bu iki ölçütü birleştirerek daha dengeli bir başarı değerlendirmesi sağlar. Ayrıca, çalışmanın ikili sınıflandırma (saldırı/normal) yapısına sahip olması, F1-Skor'un bu tür uygulamalar için daha uygun bir metrik olmasına neden olmuştur. Literatürde de siber güvenlik ve saldırı tespiti alanında yapılan birçok çalışmada, F1-Skor'un temel başarı ölçütü olarak kullanıldığı görülmektedir. Bu bağlamda, yalnızca F1-Skor'un kullanılması, çalışmanın hem teknik hem de uygulama açısından anlamlı bir tercih olduğunu göstermektedir. F1-Skor'a ait formül (2.6) da verilmiştir.

$$F1 = \frac{2\ x\ Duyarlılık\ x\ Kesinlik}{Duyarlılık\ +Kesinlik} \tag{2.6}$$

$$Kesinlik = \frac{TP}{TP+FP} \tag{2.7}$$

$$Duyarlılık = \frac{TP}{TP+FN} \tag{2.8}$$

Eşitlikler (2.7), (2.8)'de yer alan TP, FP ve FN terimleri, genellikle bir sınıflandırma modelinin başarısını değerlendirmek için kullanılan kavramlardır. Bu terimler daha ayrıntılı olarak şu şekilde açıklanabilir:

- True Positive: Modelin doğru bir şekilde pozitif sınıf olarak tahmin ettiği örneklerin sayısını ifade eder. Yani, aslında pozitif olan örneklerin doğru bir şekilde pozitif olarak sınıflandırıldığı durumdur.
- False Positive: Modelin yanlış bir şekilde pozitif sınıf olarak tahmin ettiği örneklerin sayısını ifade eder. Yani, aslında negatif olan bir örneğin yanlışlıkla pozitif olarak sınıflandırıldığı durumdur.
- False Negative: Modelin yanlış bir şekilde negatif sınıf olarak tahmin ettiği örneklerin sayısını ifade eder. Yani, aslında pozitif olan bir örneğin yanlışlıkla negatif olarak sınıflandırıldığı durumdur.

3. DENEY SONUÇLARI

Bu bölümde, gerçekleştirilen deneysel çalışmaların ayrıntılarına, kullanılan parametre ayarlarına ve elde edilen sonuçların yorumlanmasına yer verilmektedir.

Makine öğrenmesine dayalı sistemlerde, deney ortamının ve kullanılan altyapının özellikleri, model başarımı ve işlem verimliliği üzerinde doğrudan etkili olmaktadır. Deneyler, Google Research tarafından geliştirilen ve Python kodlarının çevrimiçi olarak yazılmasına ve çalıştırılmasına imkân tanıyan Google Colaboratory platformunda gerçekleştirilmiştir. COLAB, araştırmacılara GPU gibi donanım kaynaklarına erişim sağlamanın yanı sıra, çeşitli kütüphaneleri önceden yüklenmiş olarak sunarak zaman kazandırmakta ve daha karmaşık algoritmalar üzerinde çalışma kolaylığı sağlamaktadır. Özellikle büyük ölçekli veri setleriyle çalışılması gereken bu araştırmada, Colaboratory'nin sunduğu kaynaklar sayesinde deneyler hızlı ve verimli bir şekilde tamamlanmıştır.

Her iki veri setine analiz öncesinde veri temizleme ve dönüştürme işlemleri uygulanmıştır. Bu işlemler sayesinde, makine öğrenimi modellerinin daha sağlıklı çalışması için veri setleri dengeli ve temiz bir forma getirilmiştir. Veri ön işleme sürecinde aşağıdaki adımlar izlenmiştir:

- Eksik ve boş veri satırları kaldırılmıştır.
- Veri setlerinde yer alan ve sınıflandırma için önemi olmayan özellikler veri setinden kaldırılmıştır.
- Kategorik veriler, sayısal karşılıklarına dönüştürülmüştür.
- Özellik değerleri standartlaştırılarak (normalizasyon) algoritmalara uygun hâle getirilmiştir.
- Saldırı türlerine dair ayrıntılar göz ardı edilerek veriler sadece "saldırı" ve "normal" olarak etiketlenmiştir.

Çalışmada incelenen yöntemlerin, GWO, PSO, parametreleri, yapılan ön deneyler sonucunda belirlenmiştir:

- Hem GWO hem PSO için popülasyon sayısı 50, iterasyon sayısı 20 olarak alınmıştır.
- PSO için w 0.7, c1 ve c2 1.5 olarak alınmıştır.
- Verilerin %80 i eğitim %20 si test olarak kullanılmıştır. Eğitim verilerinin %10 nu ise fitness fonksiyonun başarısı değerlendirmek için kullanılmıştır.
- Amaç fonksiyonu olarak karar ağacı kullanılmış olup daha doğru bir sonuç sağlaması için F1-skor dan faydalanılmıştır.
- Seçilecek özellik sayısında bir kısıt uygulanmamıştır.
- Her bir veri seti üzerinde, her bir yöntem 10 kere yürütülmüştür.

Şekil 3.1'de temel olarak süreç aşamaları gösterilmiştir.

Şekil 3.1 Akış şeması

Saldırı tipleri de dikkate alınarak gerçekleştirilen deneylerde (Özsarı ve ark. 2024) tarafından elde edilen sonuçlara paralel sonuçlar elde edilmiştir. Bundan dolayı bu çalışmada ikili olarak deneyler gerçekleştirilmiştir. Belirlenen parametreler ile iki veri seti üzerinde GWO ve PSO algoritması ile özellik seçimi uygulanması sonucu indirgenen özellik sayısı ve F1-skor sonuçları GWO algoritması için **Çizelge 3.1**, PSO algoritması için **Çizelge 3.2** verilmiştir.

Çizelge 3.1 GWO deney sonuçları

Veri seti	Yürütüm	İndirgenen Özellik sayısı	F1-skor
USB-IDS	1	**6**	**0.95**
	2	11	0.90
	3	10	0.92
	4	9	0.89
	5	9	0.89

Çizelge 3.1 GWO deney sonuçları (devam)

	6	8	0.91
	7	8	0.92
	8	7	0.87
	9	7	0.87
	10	7	0.87
CSE-CIC- IDS2018	1	8	0.91
	2	8	0.90
	3	8	0.90
	4	7	0.90
	5	12	0.94
	6	11	0.93
	7	6	0.94
	8	10	0.91
	9	7	0.89
	10	12	0.95

Çizelge 3.2 PSO deney sonuçları

Veri seti	**Yürütüm**	**İndirgenen Özellik sayısı**	**F1-skor**
USB-IDS	1	32	0.94
	2	41	0.95
	3	29	0.95
	4	43	0.94
	5	37	0.93
	6	41	0.92
	7	**28**	**0.94**
	8	42	0.92
	9	41	0.93
	10	38	0.93
CSE-CIC- IDS2018	1	11	0.90
	2	20	0.95

Çizelge 3.2 PSO deney sonuçları (devam)

	3	13	0.91
	4	17	0.92
	5	14	0.92
	6	15	0.92
	7	18	0.93
	8	37	0.94
	9	**5**	**0.94**
	10	13	0.92

Çizelge 3.1 incelendiğinde USB-IDS veri seti üzerinde GWO algoritması, saldırı tespiti görevinde yüksek doğruluk (0.95 F1-score) oranını muhafaza ederek seçilen özellik sayısını 6'ya indirgemiştir. Bu sonuç, GWO'nun yalnızca performansı korumakla kalmayıp, aynı zamanda özellik sayısını önemli ölçüde azaltarak güçlü bir optimizasyon yeteneği sergilediğini göstermektedir. USB-IDS veri setinde GWO tarafından seçilen özellikler şunlardır:

- 'Total Length of Bwd Packet'
- 'Fwd Packet Length Mean'
- 'Bwd Header Length'
- 'PSH Flag Count'
- 'URG Flag Count'
- 'FWD Init Win Bytes'.

Aynı veri seti üzerinde PSO algoritması kullanıldığında ise performans korunarak seçilen özellik sayısı ancak 28'e indirgenebilmiştir. Bu veri seti için PSO algoritması başarı oranını korusa da özellik sayısını yeterince düşürememiştir. 28 veri setinde yer alan özellik sayısına göre düşük bir sayı olsa da GWO algoritmasının başarısı dikkate alındığında yetersiz kalmaktadır.

CSE-CIC-IDS2018 veri seti üzerinde gerçekleştirilen deneylerde GWO algoritmasının, benzer şekilde, başarısını (0.94 F1-score) muhafaza ederek seçilen özellik sayısını

6’ya düşürmüştür. Bu durum, GWO'nun farklı veri setleri üzerinde tutarlı ve yüksek performans sergileyebildiğinin bir göstergesidir. CSE-CIC-IDS2018 veri seti için GWO tarafından seçilen özellikler:

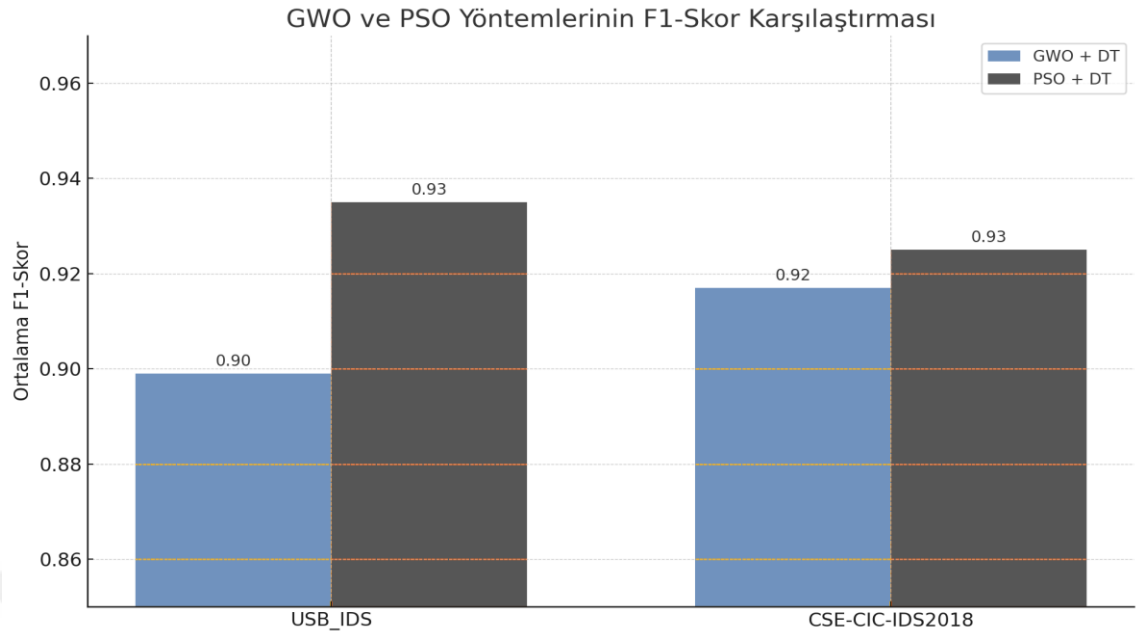
- 'Fwd Header Len'
- 'Pkt Len Mean'
- 'Init Fwd Win Byts'
- 'Fwd Act Data Pkts'
- 'Active Min'
- 'Idle Std'

PSO algoritması ile yapılan analizlerde ise seçilen özellik sayısı 5’e düşürülmüş, bu süreçte de F1-score değerinin yüksek seviyelerde (0.94) olduğu görülmüştür. GWO’ya göre performansı koruyarak özellik sayısını bir azaltmıştır. CSE-CIC-IDS2018 veri seti için PSO tarafından seçilen özellikler ise şu şekildedir:

- 'Tot Bwd Pkts'
- 'Bwd Header Len'
- 'ACK Flag Cnt'
- 'Bwd Seg Size Avg'
- 'Init Fwd Win Byts'

Tüm sonuçlar ele alındığında, GWO'nun daha az sayıda özellik seçimi ile daha optimize bir çözüm sunduğu, bu nedenle özellikle kaynak kısıtlamalarının kritik öneme sahip olduğu sistemlerde daha avantajlı bir yöntem olarak öne çıktığı sonucuna varılmıştır.

Şekil 3.2’ de elde edilen başarı sonuçlarının dağılımı gösterilmiştir.



Şekil 3.2 GWO ve PSO'nun USB-IDS ve CSE-CIC-IDS2018 F1-skor karşılaştırması

Şekil 3.2 incelendiğinde, her iki veri seti üzerinde de PSO algoritmasının ortalama F1-skor değerinin GWO'ya kıyasla daha yüksek olduğu gözlemlenmektedir (her iki durumda da 0.93). Ancak bu fark oldukça sınırlıdır ve GWO algoritmasının çok daha az öznitelikle benzer başarı seviyelerini yakalayabilmesi, algoritmanın özellik indirgeme performansının daha iyi olduğunu ortaya koymaktadır. Özellikle USB_IDS veri setinde GWO sadece 6 öznitelik ile 0.95 F1-skor üretmişken, PSO 28'den fazla öznitelik ile benzer skoru elde etmiştir. Bu durum, GWO'nun veri boyutunu azaltarak hesaplama yükünü hafifletme açısından daha uygun bir seçenek olduğunu desteklemektedir.

Bu çalışma, GWO'nun yüksek performansı ve özellik sayısını minimize etme yeteneği ile farklı veri setlerinde saldırı tespit sistemleri için güçlü bir alternatif olabileceğini göstermektedir.

4. SONUÇ

Günümüzde teknolojinin hızla gelişmesi, bilgiye erişimi kolaylaştırdığı gibi, siber güvenlik tehditlerini de beraberinde getirmiştir. Özellikle çevrimiçi sistemlerin yaygınlaşmasıyla birlikte, veri güvenliği konusu daha da kritik hâle gelmiştir. Artan dijitalleşme, saldırı tespit sistemlerinin daha gelişmiş, akıllı ve dinamik yapılara dönüşmesini gerekli kılmıştır. Bu bağlamda, yapay zekâ ve özellikle makine öğrenimi tabanlı yöntemler, yüksek örüntü tanıma kapasiteleri sayesinde siber tehditlerin belirlenmesinde etkin çözümler sunmaktadır.

Yürütülen bu çalışmada, saldırı tespitine yönelik olarak USB_IDS_1 ve CSE-CIC-IDS2018 veri setleri üzerinde özellik seçimi yapılarak ikili sınıflandırma gerçekleştirilmiştir. Saldırı türü ayrımı yapılmaksızın veriler "saldırı" ve "normal" olarak etiketlenmiş ve bu yapı üzerinden analiz yapılmıştır. Öznitelik seçimi için GWO ve PSO algoritmaları kullanılmış; ardından elde edilen özellik alt kümeleri, karar ağacı algoritması ile sınıflandırılmıştır. Model başarımı F1-skor metriği kullanılarak değerlendirilmiş ve sınıflandırma performansları karşılaştırılmıştır. F1-skor üzerinden yapılan değerlendirme, sınıf dengesizliği olasılığını göz önüne alarak sistemin genel başarımını daha adil bir biçimde ölçmeye olanak sağlamıştır.

Elde edilen sonuçlar göstermiştir ki, GWO algoritması, özellikle USB_IDS_1 veri seti üzerinde yalnızca 6 öznitelik seçerek 0.95 F1-skoru gibi yüksek bir başarıya ulaşmıştır. Aynı veri setinde PSO algoritması ise ancak 28–43 arası öznitelik ile benzer skorlar üretmiştir. Bu da GWO'nun gereksiz özniteliklerden arındırılmış sınıflandırma sistemleri oluşturmakta başarılı olduğunu göstermektedir. CSE-CIC-IDS2018 veri setinde her iki algoritma da yüksek performans göstermiştir. GWO ile 6–12 öznitelik seçilerek 0.94–0.95 F1-skor aralığı elde edilmiş, PSO ise 5–20 öznitelik ile benzer başarı seviyelerinde sonuçlar üretmiştir. Tüm sonuçlar dikkate alındığında GWO + DT kombinasyonu, her iki veri setinde de yüksek F1-skor değerlerine ulaşarak etkin bir yöntem olmuştur. PSO algoritması da istikrarlı sonuçlar üretmiş ancak USB_IDS_1 veri setinde GWO'ya kıyasla daha düşük performans göstermiştir. Yürütülen bu çalışma,

literatüre şu katkıları sunmaktadır:

- GWO'nun IDS sistemlerinde minimal öznitelik ile yüksek doğrulukta sınıflandırma yapılabileceğini deneysel olarak ortaya koymuştur.
- Farklı veri setlerinde yapılan çoklu yürütmelerle algoritmaların performans stabilitesi karşılaştırılmıştır.
- F1-Skor odaklı değerlendirme yaklaşımı, dengesiz veri dağılımlarında güvenilir metrik seçiminin ne kadar kritik olduğunu uygulamalı olarak göstermiştir.

Bu bulgular öznitelik seçiminin saldırı tespit sistemlerinin başarımı üzerinde doğrudan etkili olduğunu ortaya koymaktadır. Sonuca etkisi olmayan özelliklerin kullanılması algoritmaların çalışma süresini yavaşlatmaktadır. Saldırı tespit sistemlerinde çok büyük verilerle çalışıldığı dikkate alındığında gereksiz özelliklerin kullanılmaması oldukça önemlidir.

Bu çalışmanın ilerleyen aşamalarında, saldırı türlerinin ayrıntılı olarak sınıflandırıldığı çok sınıflı modellerin geliştirilmesi planlanmaktadır. Özellikle DDoS, port tarama, brute-force, botnet gibi yaygın saldırı türleri için özel optimize edilmiş sınıflandırıcılar oluşturulması hedeflenmektedir. Ayrıca, Yapay Arı Kolonisi, Genetik Algoritma, Ateşböceği Algoritması gibi farklı sezgisel algoritmalarla öznitelik seçimi gerçekleştirilerek, mevcut yöntemlerle karşılaştırmalı performans analizleri yapılacaktır. Bunun yanında, geliştirilen sistemlerin gerçek zamanlı ağ trafiğine entegre edilerek canlı veri akışı üzerinde test edilmesi, saldırı tespiti sürelerinin ölçülmesi ve yük altında sistem davranışlarının incelenmesi gibi uygulamalı çalışmalar da planlanmaktadır.
İleri düzeyde ise, derin öğrenme tabanlı modellerin (özellikle LSTM, CNN-LSTM gibi yapılar) kullanımının, karmaşık ağ davranışlarını daha iyi öğrenme ve saldırı örüntülerini tanıma açısından avantajları değerlendirilecektir.

KAYNAKLAR

Bishop, C.M. 1995. Neural networks for pattern recognition. Oxford University Press, 502 s., England.

Canadian Institute for Cybersecurity. 2018. CSE-CIC-IDS2018 Dataset. Web Sitesi: https://www.unb.ca/cic/datasets/ids-2018.html Erişim Tarihi: 10.04.2025.

Catillo, M., Del Vecchio, A., Ocone, L., Pecchia, A. ve Villano, U. 2022. USB-IDS-1: A public multilayer dataset of labeled network flows for IDS evaluation. 51st Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshops (DSN-W); 1–6.

Chinchor, N. 1992. MUC-4 evaluation metrics. Proceedings of the 4th Message Understanding Conference (MUC-4); 22–29.

Cover, T. ve Hart, P. 1967. Nearest neighbor pattern classification. IEEE Transactions on Information Theory, 13(1); 21–27.

Dang, Q.V. 2022. Multi-layer intrusion detection on the USB-IDS-1 dataset. Future Data and Security Engineering. Springer, Cham; 812–825.

Guyon, I. ve Elisseeff, A. 2003. An introduction to variable and feature selection. Journal of Machine Learning Research, 3; 1157–1182.

Kalutharage, D., Khan, S., Anwar, A., Ekonomou, E., ve Gamage, P. 2022. Explainable deep autoencoder models for IoT security monitoring. Future Generation Computer Systems, 128: 208–219.

Karaman, M.S. 2020. Anomali tabanlı saldırı tespit sistemlerinde makine öğrenimi modellerinin performans değerlendirmesi. Yüksek Lisans Tezi. İstanbul Ticaret Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

Kennedy, J. ve Eberhart, R. 1995. Particle swarm optimization. Proceedings of ICNN'95- International Conference on Neural Networks, 4; 1942–1948.

Manocha, S. ve Girolami, M.A. 2007. An empirical analysis of the probabilistic K-nearest neighbour classifier. Pattern Recognition Letters, 28; 1818–1824.

Mirjalili, S., Mirjalili, S.M. ve Lewis, A. 2014. Grey Wolf Optimizer. Advances in Engineering Software, 69; 46–61.

Moustafa, N. ve Slay, J. 2015. UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set). Military Communications and Information Systems Conference (MilCIS), 2015; 1–6.

Özsarı, M.V., Özsarı, Ş., Aydın, A. ve Güzel, M.S. 2024. USB-IDS-1 dataset feature reduction with genetic algorithm. Communications Faculty of Sciences University of Ankara Series A2-A3 Physical Sciences and Engineering, 66(1); 199–211.

Quinlan, J.R. 1986. Induction of decision trees. Machine Learning, 1(1); 81–106.

Resende, P.A.A. ve Drummond, A.C. 2018. A survey of random forest based methods for intrusion detection systems. ACM Computing Surveys, 51; 48.

Soltani, M., Siavoshani, M.J. ve Jahangir, A.H. 2020. A content-based deep intrusion detection system. arXiv preprint arXiv:2001.05009.

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

Tavallaee, M., Bagheri, E., Lu, W. ve Ghorbani, A.A. 2009. A detailed analysis of the KDD Cup 99 data set. Proceedings of the 2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications (CISDA), 1–6.

Zhao, G., Wang, H., Jia, D. and Wang, Q. 2019. Feature selection of grey wolf optimizer based on quantum computing and uncertain symmetry rough set. Symmetry, 11(12); 1470.