
GAZİ ÜNİVERSİTESİ
Gazi
1926

SİBER SAVUNMA ALANINDA YAPAY ZEKÂ TABANLI SALDIRI TESPİTİ VE ANALİZİ

Burcu AYTAN

YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

TEMMUZ 2019

Burcu AYTAN tarafından hazırlanan "SİBER SAVUNMA ALANINDA YAPAY ZEKÂ TABANLI SALDIRI TESPİTİ VE ANALİZİ" adlı tez çalışması aşağıdaki jüri tarafından OY BİRLİĞİ ile Gazi Üniversitesi Bilgisayar Mühendisliği Ana Bilim Dalında YÜKSEK LİSANS TEZİ olarak kabul edilmiştir.

Danışman: Doç. Dr. Necaattin BARIŞÇI

Bilgisayar Mühendisliği Ana Bilim Dalı, Gazi Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum. ...……

Başkan: Doç. Dr. Ümit Sami SAKALLI

Endüstri Mühendisliği Ana Bilim Dalı, Kırıkkale Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum. ……...

Üye: Dr. Öğr. Üyesi Hüseyin POLAT

Bilgisayar Mühendisliği Ana Bilim Dalı, Gazi Üniversitesi

Bu tezin, kapsam ve kalite olarak Yüksek Lisans Tezi olduğunu onaylıyorum. ……...

Tez Savunma Tarihi: 18/07/2019

Jüri tarafından kabul edilen bu tezin Yüksek Lisans Tezi olması için gerekli şartları yerine getirdiğini onaylıyorum.

……..……

Prof. Dr. Sena YAŞYERLİ

Fen Bilimleri Enstitüsü Müdürü

ETİK BEYAN

Gazi Üniversitesi Fen Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmasında;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmasında yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu,

bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

Burcu AYTAN

18/07/2019

SİBER SAVUNMA ALANINDA YAPAY ZEKÂ TABANLI SALDIRI TESPİTİ VE ANALİZİ

(Yüksek Lisans Tezi)

Burcu AYTAN

GAZİ ÜNİVERSİTESİ

FEN BİLİMLERİ ENSTİTÜSÜ

Temmuz 2019

ÖZET

Bilgi ve bilgisayar teknolojilerinin hızlı gelişimi, hız ve verimlilik artışı ve kolaylık sağlaması nedeniyle birçok bilgi elektronik ortamlara aktarılmıştır. Elektronik ortamların yaygınlaşması; kişisel veya kurumsal açıdan önemli bir bilginin başkalarının eline geçmesi, maddi ve manevi zararlara yol açması gibi sorunları da ortaya çıkarmıştır. Korunacak bilginin değerine göre farklılık gösterebilecek olan koruma sistemlerinin aslında tek amacı, saldırganlara ve saldırılara karşı önlem alarak, bilginin mahremiyetinin korunmasıdır. Özellikle kamu kurum ve kuruluşlarının itibarını kaybetmesine neden olabilecek bilginin gizliliği, bütünlüğü ve erişilebilirliğinin bozulması yönünde yapılan saldırılara karşı sistem güvenlik uzmanları önlem almaya çalışmakta ve sistemin kaynaklarını yetkisiz erişimden korumaktadır. Bilgiyi korurken, var olan sistemlerin sürekliliğinin sağlanması, yapılan saldırılara karşı alınan önlemlerin güncelliğinin korunması, değişen saldırı ve yöntemlerin bilinmesi ve var olan sisteme adapte edilmesi gerekmektedir. Günümüzde bilgi ve bilgisayar güvenliğinin öneminin kavranmasıyla geliştirilen araçlardan biri olan Saldırı Tespit Sistemleri ile sistemlere yapılan yetkisiz erişimler ve kötüye kullanımlar tespit edilerek, bunların yol açabileceği zararlar engellenmeye çalışılmaktadır. Bu çalışmada saldırı tespit sistemleriyle ilgili uygulamalarda en çok kullanılan veri setlerinden biri olan "KDD Cup'99" veri seti kullanılarak hizmet dışı bırakma saldırıları ve bilgi tarama saldırıları Weka aracında yer alan makine öğrenme algoritmaları ile tespit edilmeye çalışılmış ve Geri Yayılma Algoritması ile %99.88, Karar Ağacı Algoritması ile %99.82, Ripper Kuralı Algoritması ile %99.74 ve Rasgele Orman Algoritması ile de %99.94 oranında başarı sağlanmıştır.

Bilim Kodu	:	92432
Anahtar Kelimeler	:	Saldırı Tespit Sistemleri, Bilgisayar Ağlarında Anormallik Tespiti, Siber Saldırılar, Makine Öğrenmesi, Yapay Zekâ
Sayfa Adedi	:	61
Danışman	:	Doç. Dr. Necaattin BARIŞÇI

ARTIFICIAL INTELLIGENCE BASED INTRUSION DETECTION AND ANALYSIS IN THE CYBER DEFENCE AREA

(M. Sc. Thesis)

Burcu AYTAN

GAZİ UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

July 2019

ABSTRACT

Vast amount of information has been transferred into electronic media due to the rapid development of computer technologies, increase of speed and efficiency as well as ease of use. The fact that the electronic media become widespread has also resulted in various problems such as taking the possession of information which have personal and corporate importance to third persons and causing pecuniary losses and intangible damages. The sole purpose of the protection systems which may differ according to value of the information to be protected is to protect the confidentiality of the information by taking measures against the attacks and attackers. In order to prevent public bodies and institutions from losing their reputation in particular, system security experts are trying to take measures against the attacks that aim to harm confidentiality, integrity and accessibility of the information and they protect the system resources from unauthorized access. While protecting the information, it is necessary to maintain the continuity of the existing systems, update the measures against the potential attacks, have knowledge of changing attacks and methods as well as adapting all these actions to the existing system. Today, intrusion detection systems which are one of the tools developed in parallel to the understanding the importance of information and computer security, are utilized to detect unauthorized access and abuse of systems thereby preventing potential damages to be caused by them. In this study, it has been tried to identify attacks for rendering inoperable and information retrieval thanks to the machine learning algorithms in Weka tool by using the "KDD Cup 99" data set that is one of the commonly used data sets in the applications for intrusion detection systems and so with the Back Propagation Algorithm 99.88%, the Decision Tree Algorithm 99.82%, the Ripper Rule Algorithm 99.74% and the Random Forest Algorithm 99.94% success has been achieved at the end of the study.

Science Code	:	92432
Key Words	:	Intrusion Detection Systems, Anomaly Detection in Computer Networks, Cyber Attacks, Machine Learning, Artificial Intelligence
Page Number	:	61
Supervisor	:	Assoc. Prof. Dr. Necaattin BARIŞÇI

TEŞEKKÜR

Tez çalışmamın gerçekleştirilmesinde, değerli bilgilerini benimle paylaşan, kendisine ne zaman danışsam bana kıymetli zamanını ayırıp sabırla ve büyük bir ilgiyle bana faydalı olabilmek için elinden geleni sunan, her sorun yaşadığımda yanına çekinmeden gidebildiğim, güler yüzünü ve samimiyetini benden esirgemeyen, kıymetli bilgi, birikim ve tecrübeleri ile bana yol gösterici ve destek olan, ilgisini ve önerilerini göstermekten kaçınmayan değerli danışman hocam sayın Doç. Dr. Necaattin BARIŞÇI'ya sonsuz teşekkür ve saygılarımı sunarım.

Yüksek lisans eğitimim boyunca yardım, bilgi ve tecrübeleri ile bana sürekli destek olan Gazi Üniversitesi Teknoloji Fakültesi Bilgisayar Mühendisliği bölümündeki tüm hocalarıma teşekkür ederim.

Bu çalışmayı hazırlarken geçirdiğim süreçte benden yardımını esirgemeyen Şube Müdürüm Sayın Serdar GÜLBAĞ'a ve çalışmalarım boyunca maddi ve manevi destekleriyle beni hiçbir zaman yalnız bırakmayan ve desteklerini hiçbir zaman eksik etmeyen değerli aileme, sevgili eşim Veysel AYTAN'a ve biricik kızım Ayris'e sonsuz teşekkürler ederim.

İÇİNDEKİLER

Sayfa

Sayfa

ÇİZELGELERİN LİSTESİ

Çizelge **Sayfa**

ŞEKİLLERİN LİSTESİ

Şekil **Sayfa**

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Simgeler	Açıklamalar
B	Basic – Temel
C	Content – İçerik
DN	Doğru Negatif
DoS	Denial of Service – Servis Dışı Bırakma
DP	Doğru Pozitif
DR	Tespit Oranı – Detection Rate
FAR	Yanlış Alarm Oranı – False Alarm Rate
FN	False Negative
FP	False Positive
FTP	Dosya Aktarım Protokolü
GA	Genetik Algoritma
H	Host
IDS	Intrusion Detection System – Saldırı Tespit Sistemi
IPS	Intrusion Prevention System – Saldırı Önleme Sistemi
NAO	Normal Algılama Oranı
RF	Random Forest – Rasgele Orman
R2L	Remote to Local Attack
STO	Saldırı Tespit Oranı
T	Traffic – Trafik
TAO	Toplam Algılama Oranı
TN	True Negative
TP	True Positive
U2R	User to Root Attack
YP	Yanlış Pozitif
YN	Yanlış Negatif

1. GİRİŞ

Siber Güvenlik, siber ortamda; kurum, kuruluş ve kullanıcıların varlıklarını korumak amacıyla kullanılan araçlar, politikalar, güvenlik kavramları, güvenlik teminatları, kılavuzlar, risk yönetimi yaklaşımları, faaliyetler, eğitimler, en iyi uygulamalar ve teknolojiler bütünüdür [1]. Siber saldırı girişimlerinde düşman olarak belirlenen hedefe saldırıda bulunmak, karşı savunma yapmak, hedefteki siber uzayda istihbarat verileri toplamak siber savaş faaliyetlerini oluşturmaktadır.

Siber saldırı; dünyanın herhangi bir yerindeki bilgisayar kontrolü altındaki sistemlere internet ortamından izinsiz erişip kritik alt yapının yönetimini ele geçirmeye çalışmaktır. Siber saldırının silahları ise internet ortamına bağlı bir bilgisayarın tuşları, bu tuşlara dokunan parmaklar ve yazılımlardır [2]. Etkin ağlarda siber saldırı ile tüm alt yapılar bir anda yerle bir edilebilir, en güçlü ülke bile hareket edemez hale getirilebilir. Etkin ağlarda dinamik olarak gelişen saldırılara karşı savunmak için geleneksel sabit algoritmalar ile yazılım geliştirmek güçtür. Bu durum, yazılım esnekliği ve öğrenme yeteneği sağlayan yapay zekâ yöntemleri uygulanarak ele alınabilir [3].

Bilişim teknolojilerinin hızlı gelişimi kullanıcılar için birçok alanda kolaylık sağlarken bu sistemlerin kötüye kullanılmasından kaynaklanan birçok tehlikeyi de beraberinde getirmiştir. Bilginin gizliliği, bütünlüğü ve erişilebilirliği siber güvenlik açısından önemli birer faktördür. Bir veriye sadece istenen kişiler tarafından erişilmesi veya görüntülenebilmesi bu verinin gizliliğini, verinin herhangi bir değişime uğramadan istenen hedefe iletilmesi verinin bütünlüğünü, saklanan verinin istenildiğinde yetkili kişiler tarafından ulaşılabilir olması ise verinin erişilebilirliğini ifade eder.

Siber saldırıların bir kısmı verinin gizliliği, bütünlüğü ve erişilebilirliğini bozmaya yönelik yapılan saldırılardır. Ancak daha gelişmiş siber saldırılar da mevcuttur. Örneğin günümüzde yapılan siber saldırılar yalnızca bir bilgisayar sistemini etkisiz hale getirmek ya da sistem üzerinde yer alan veriler üzerinde değişiklik yapmakla sınırlı değildir. Ülkelerin kritik altyapılarını hedef almaktadır. Bu kritik altyapılar güvenlik, sağlık, enerji ve ulaşım ağları, haberleşme sistemleri, bankacılık, kamu hizmetleri, askeri komuta kontrol sistemleri gibi kritik sektörlerin bilgi sistem altyapılarıdır. Bu kritik altyapıların siber dünyada hedef haline

gelmesi tüm dünyada önemli bir unsur haline gelmekte ve daha etkin savunma sistemlerinin geliştirilmesi için zemin hazırlamaktadır. Siber saldırıların insan hayatını olumsuz yönde etkileyecek unsurlar içermesi ulusal siber güvenlik politikalarının belirlenmesine neden olmuştur. Siber saldırıların tespit edilmesi, tespit edilen saldırılara karşı taarruz yapılması, sistemin zarar görmeden eski haline geri döndürülmesi, saldırıların kimin tarafından yapıldığının belirlenmesi ve bu konuda gerekli hukuksal yaptırımların sağlanması gibi stratejiler geliştirilmektedir.

Bilgisayar sistemlerinde karşılaşılan birçok saldırı türü mevcuttur. Sistem kesintisi, erişim kısıtlama, verileri bozma veya çalma, içerik değiştirme gibi amaçlar edinilerek saldırılar düzenlenmektedir. Hizmet kesintisi saldırıları, sistemin yoğun istekler sonucu cevap veremez hale gelmesini sağlayan saldırılardır. Yetkisiz veya izinsiz erişim ile yapılan saldırılar, şifre kırıcılar ve klavye dinleme sistemi (keylogger) gibi araçlar ile yapılan ve bilgi ifşa eden saldırılardır.

Sistemlere yapılan siber saldırılar, herhangi bir veri üzerinde değişiklik yapılması, yanlış veri girişi yapılması veya bazı verilerin silinmesi gibi veri aldatmacası saldırıları olabilir. Ayrıca sistem üzerinde yer alan kullanıcı adı ve şifrelerinin keşfedilmesini, sistem açıklıklarından yararlanarak sisteme zarar verilmesini veya özel programlar ile sistemde saldırgan tarafından istenen komutların çalışmasını sağlayacak casus yazılımlar da bilgisayar sistemlerine yapılan siber saldırı çeşitlerindendir.

Kritik altyapılar üzerine yapılan siber saldırılar ise daha çok belirli bir amaca yönelik olarak hazırlanmış saldırılardır. Bir programın içerisine zararlı bir kod parçasının yerleştirilmesi ve hedef alınan sistem üzerindeki tüm bilgilerin yok edilmesi veya kullanılamaz hale gelmesini sağlar.

Son kullanıcıları hedef alan siber saldırılar genellikle sosyal ağlar üzerinden kişi ile ilgili belli başlı bilgiler edindikten sonra onlara yönelik olarak hazırlanmış bir e-posta ile yapılan oltalama saldırılarını veya sosyal mühendislik saldırılarını içermektedir.

Kamu kurum ve kuruluşlarının itibarını kaybetmesine neden olabilecek siber saldırılar ise kuruma ait web sayfasının hizmet kesintisine uğramasına neden olabilecek Servis Dışı Bırakma (Denial of Service - DoS) saldırıları veya web sitesi içerisinde farklı içerikler

sunulmasına neden olabilecek web sayfası hırsızlığı veya kuruma ait web sayfasının başka bir sayfaya yönlendirilmesine neden olan yönlendirme saldırılarından oluşmaktadır.

Bilgisayar veya ağ sistemlerine yapılan saldırıları tespit ederek güvenliğin sağlanması için geliştirilen sistemler, her ne kadar yapılan saldırıların büyük bir çoğunluğunu tespit edebilseler de daha önce hiç karşılaşılmamış olan saldırıların büyük çoğunluğunu tespit edememektedir [4]. Bu saldırıların sistemlerde büyük zararlara yol açması, yeni saldırı çeşitlerinin tespit edilebilmesi, hızla değişen saldırı tiplerinin karşısında, bilgi ve bilgisayar güvenliğinin sağlanması amacıyla bu sistemlerin geliştirilmesinde yapay zekâ yöntemleri kullanılması hedeflenmektedir.

Siber güvenlik ile ilgili araştırma raporlarının günlük olarak kamuya duyuruluyor olması ve dijital dünyanın gittikçe daha açık kaynaklı hale gelmesi ileri güvenlik araştırmaları ve teknolojideki gelişmeler açısından olumlu bir etmen olsa da ne yazık ki güvenlik analistleri tarafından kullanıldığı kadar siber suçlular tarafından da kullanılabilmektedir.

Giderek dijitalleşen dünyada siber saldırıların artış göstermesi ile güçlü ve gelişmiş güvenlik teknolojilerine duyulan ihtiyaçlar da artmaktadır. Dijital dünyada güvenlik, internet ve ağ teknolojilerindeki gelişmeler ile daha da karmaşık hale gelmekte ve bu sebeple en önemli endişe kaynağı olmaktadır. Bu nedenle makine öğrenmesi uygulamalarının hem siber savunma tarafında hem de siber taarruz faaliyetlerinde kullanımına ihtiyaç duyulmuştur.

Siber güvenlik alanında makine öğrenimi teknikleri tespit edilemeyen veya gözden kaçabilen tehditleri belirlemek ve kararlı bir şekilde ele almakta yardımcı olmaktadır. Ancak her ne kadar önlem alınmaya çalışılsa da teknolojinin gelişmesi ve yapay zekanın kullanım alanlarının artış göstermesi aynı zamanda siber suçlular için daha büyük ve daha sofistike saldırılar hazırlamak ve başlatmak için zemin hazırlamaktadır. Bu nedenle güvenlik için istenen başarı oranının %100 olması gerekmektedir.

Günümüzde pek çok kamu kurum ve kuruluşlarının, özel sektör ve bireylerin siber suçlular tarafından mağdur edildiği bilinmektedir. Çalınan kişisel ve finansal bilgiler, istihbarat verileri ve bu tür bilgilerin yasa dışı kullanımı istenmeyen sonuçlar doğurmakta ve geri kurtarımı imkânsız hale gelmektedir. Son yıllarda, siber suçlardan dolayı çok büyük zararlara yol açıldığı ve siber güvenlik konusunda uzman personel sayısının az olmasından

dolayı bu suçlarla mücadelenin eksik kaldığı gözlemlenmektedir. Bu nedenle de insan gibi düşünen ve insan gibi olaylara müdahale edebilen yapay zekâ tabanlı saldırı tespit sistemlerine ihtiyaç duyulmaktadır. Siber saldırıların karmaşıklığının hızla artış göstermesi ve savunma araçlarının bu değişim karşısında daha eski kalması herhangi bir sisteme yetkisiz girişin tespit edilme süresini arttırmaktadır. Sürekli yeni tehditlerin ve sistem zafiyetlerinin ve açıklıkların ortaya çıkması bunlara müdahale etmeyi zorlaştırmaktadır.

Artan siber güvenlik tehditlerine karşı siber saldırıları tespit etmek, durdurmak ve yanıtlamak için yapılan çalışmalar makine öğrenme tekniklerine odaklanmaktadır. Çeşitli siber güvenlik alanlarında saldırı tespiti ve müdahale için makine öğreniminden yararlanılabilir. Temel olarak yapılan görevler makine öğrenme algoritmaları sayesinde otomatikleştirilerek güvenlik süreçleri de hızlandırılabilir ve siber güvenlik alanında çalışan analistlerin diğer saldırıları tespit edebilmesi kolaylaştırılabilir.

Bu tez çalışması beş bölümden oluşmaktadır. İlk bölümde bu tez çalışmasında ele alınan konunun ne olduğuna dair kısa bir bilgi verilmiş, araştırmanın amacı ve önemi vurgulanmış ve bazı tanımlamalar yapılmıştır.

İkinci bölümde literatür araştırması yapılmış ve yapay zekâ tabanlı saldırı tespiti ile ilgili çalışmalardan bahsedilmiştir. Kullanılan algoritmalar, verileri tanımlayıcı özellikler ve daha hızlı sürede saldırı tespit edilebilmesi için gerçekleştirilen yöntemler incelenmiştir.

Üçüncü bölümde yapay zekâ teknikleri incelenmiş ve makine öğrenme algoritmalarından bahsedilmiştir. Bilgi ve bilgisayar güvenliği için geliştirilen saldırı tespit sistemleri araştırılmış ve saldırı tespit sistemlerinin türleri açıklanmıştır. Saldırı tespit sistemlerinde en çok kullanılan veri seti olan KDD 99 veri seti incelenmiş ve veri setinin özellikleri, barındırdığı saldırı türleri ayrıntılı olarak incelenmiştir. Ayrıca çalışmada kullanılan Weka aracının özellikleri, içerdiği algoritmaların yanı sıra çalışmamızda kullanılan makine öğrenme algoritmaları da ayrıntılı olarak anlatılmıştır.

Dördüncü bölümde araştırma bulguları ele alınmıştır. Kullanılan veri setinde yer alan her bir verinin ayırt edilebilmesi için kullanılan özellikler ayrıntılı bir şekilde açıklanmıştır. Eğitim ve test veri setlerinden bahsedilmiş kullanılan algoritmalardaki başarı oranları gösterilmiştir.

Elde edilen sonuçlar daha önce saldırı tespit sistemlerinde kullanılan diğer yöntemlerle karşılaştırılmış ve yapılan çalışmanın önemi vurgulanmıştır.

Son bölümde ise yapılan çalışmanın önemi, elde edilen çıkarımlar ve değerlendirmelerden bahsedilmiştir.

2. LİTERATÜR ARAŞTIRMASI

Literatürde yapay zekâ tabanlı saldırı tespiti ile ilgili birçok başarılı çalışma mevcuttur. Birbirinden farklı algoritmalar kullanılarak saldırılar tespit edilmiş, en tanımlayıcı özelliklerin seçilip diğer özelliklerin saf dışı bırakılması ile daha hızlı sürede saldırı teşhisi yapılmış, birden fazla algoritmanın bir arada kullanılması ile de daha doğru sonuçların elde edilmesi amaçlanmıştır. Ancak, her ne kadar yüksek oranda başarılar elde edilse de tam anlamıyla başarı sağlanamamış yeni saldırı tipleri için yanlış sınıflandırmalarla karşılaşılmıştır.

Hanifi ve arkadaşları [5] ağ saldırılarının tespiti için KDD 99 veri setini kullanılarak k-ortalama algoritması ile yarı eğitmenli yeni bir saldırı tespit sistemi tasarlamışlardır. Eğitim aşamasında, normal örnekler k-ortalama algoritması uygulanarak kümelere ayrılmıştır. Verileri normal ve saldırı verisi olarak gruplayabilmek için örneklerin kümenin merkezinden uzaklıklarına göre, bir eşik değer hesaplaması yapılmıştır. Hesaplanan eşik değere yakın olan veriler normal veri, uzak olan veriler ise saldırı verisi olarak nitelendirilmiştir.

Bu çalışmada, yeni bir yarı-eğitmenli anormal durum tespit yöntemi önerilmiştir. Sistem sadece normal veri tipleri ile eğitilmiştir. İlk olarak normal örneklerin özellikleri tanımlanmış ve bu tanımlamaya uygun veriler benzerliklerine göre sınıflandırılmıştır. Daha sonra bilinmeyen bir örneğin normal örneklere ait kümelerden birine belirlenen bir eşik seviyesi kadar benzer olup olmadığına bakılmıştır. Herhangi bir kümeye benzer olmayan örnekler anormal durum olarak tespit edilmiştir.

Kullanılan algoritmadaki K değeri büyüdükçe verilerin birbirine daha çok benzediği kümeler elde edilmiştir. Böylece normal örnekler benzerliklerine göre gruplandırılmıştır. Bilinmeyen bir örnek eğer mevcut kümelerden birine yeterince benziyorsa normal, hiçbir kümeye benzemiyorsa anormal olarak değerlendirilmiştir.

Tasarlanan sistemde saldırı tipine önem vermeksizin saldırının yakalanma işlemine öncelik verilmiştir. Bu yüzden sistemin sınıflandırma başarısı ölçülürken tüm saldırı tipleri anormal saldırı sınıfının altında değerlendirilmiştir. Genel olarak ağ trafiği verilerinde normal kayıtların fazla olması nedeniyle normal veriler kıyas alınmış ve yarı-eğitmenli yaklaşım

tercih edilmiştir. Sistem sadece normal veri ve saldırı verisi ayrımı yapabildiği için yeni gelen saldırı tiplerinin yakalanmasında da başarılı olmuştur. Tasarlanan bu sistemin daha önce farklı yöntemler ile yapılan benzer çalışmalardaki başarıya yakın bir başarı elde ettiği gözlemlenmiştir.

Bir saldırı tespit sistemi olası güvenlik tehditlerini tanımlamak için bilgisayar veya ağ içindeki farklı alanlardan bilgileri toplar ve analiz eder. Alakasız ve gereksiz özellikler içeren büyük miktarda verilerin doğru tanımlanabilmesi için özellik seçimi önemlidir. Büyük veri kümelerinde veri madenciliği yaparken genel sistem performansını önemli ölçüde artırmak için ön işlem adımının yapılması gerekmektedir. Mevcut Saldırı Tespit Sistemleri, izinsiz giriş veya kötüye kullanım gibi anormal durumları saptamak için tüm veri özelliklerini inceler. Bir saldırı tespit sisteminin incelemesi gereken denetim verisi miktarı küçük bir ağ için bile çok büyük olduğundan analizleri zordur. Bu nedenle saldırı tespit sistemleri veri filtreleme veya veri kümeleme yoluyla işlenecek veri miktarını azaltmalıdır.

Hasan ve arkadaşları [6] KDD 99 veri setindeki 41 özelliği 25'e indirmiş ve daha iyi sonuçlar elde etmişlerdir. Bu çalışmada rasgele orman algoritmasına dayanan iki aşamalı bir özellik seçimi yaklaşımına yer verilmiştir. İlk adım, önem derecesi daha yüksek olan özelliklerin seçilmesi ve arama sürecinin başlatılmasıdır. İkinci adım ise sınıflandırma aşamasıdır. KDD 99 veri setinde çok fazla miktarda veri bulunduğu için RRE-KDD veri seti şeklinde tanımlanan ve KDD 99 veri setindeki kayıtlardan oluşan yeni bir veri kümesi üzerinde çalışma gerçekleştirilmiştir. Sistem tarafından doğrudan işlenen veri miktarını azaltmak için özellik seçimi tercih edilmiştir.

KDD 99 veri setinde 4 tip saldırı mevcuttur ve her saldırı tipi için belirlenmiş farklı özellikler vardır. Saldırı tiplerinin belirlenmesinde etkin rol alan 25 temel özellik baz alınarak sistem geliştirilmiştir. Rasgele orman algoritması uygulanan sistemde 25 özellik ile yapılan tespit oranının 41 özellik ile yapılan tespit oranına göre daha iyi olduğu gözlemlenmiştir. Özellik sayısının azaltılması ile zaman tasarrufunun da sağlandığı ve 25 değişkenli uygulamanın 41 değişkenli uygulamadan daha az zamanda tespit işlemini gerçekleştirdiği belirtilmiştir.

Bu çalışmada, giriş özelliklerini azaltarak rasgele orman algoritması ile saldırı tespiti için performans iyileştirmeye odaklanan bir sistem tasarlanmıştır. Daha az sayıda özellik, hem veri yönetimi hem de işlem süresi açısından her zaman avantajlıdır. Elde edilen sonuçlar,

rasgele orman sınıflandırmasının azaltılmış özelliklerle (25 özellik) kabiliyetinin, tüm özelliklerle (41 özellik) sınıflandırılmasından elde edilen sonuçtan daha doğru sonuç verdiğini göstermektedir. Ayrıca, rasgele orman ile 25 özelliği işlemek için gereken süre, 41 özellikli işlem süresinden daha küçüktür. 41 özellikli sınıflandırma ile 10.62 dakikalık eğitim süresinde %91.41 oranında doğruluk elde edilirken, 25 özellikli sınıflandırmada 7.98 dakikalık eğitim süresinde %91.90 oranında doğruluk elde edilmiştir.

Özgür ve Erdem çalışmalarında [7] sınıflandırıcı başarısını artırmak için, tek sınıflandırıcının yerine sınıflandırıcı füzyonu kullanımını önermektedir. Bu çalışmada; nitelik seçme ve sınıflandırıcı füzyonu ile ağırlık belirleme işlemlerinin, Genetik Algoritma (GA) kullanılarak yapılması önerilmektedir. Nitelik seçme ve makine öğrenme algoritmalarının birleştirilmesinin birlikte kullanımı başarı oranını arttırmış, hatalı tespit oranını azaltmıştır.

Büyük veri setlerinde sınıflandırma yaparken özellik seçimine dikkat etmek gerekmektedir. Fazla miktarda özellik bulundurulması makine öğrenme algoritmalarındaki eğitim ve test süresini arttırmaktadır. Doğru tespitlerde bulunmak için doğru özelliklerin seçilmesi gerekmektedir. Örneğin bir verinin saldırı verisi mi normal veri mi olduğunu anlayabilmek için en önemli özellikler neler ilk olarak bunların tespit edilmesi gerekmektedir. Bu çalışmada seçilen niteliklerin hem makine öğrenme algoritmasındaki performans artışına hem de eğitim ve test aşamalarındaki hıza dikkat edilmiştir.

KDD99 veri setinde çok fazla örneklem olması eğitim ve test aşamalarını zorladığı için ilk olarak bu veri setindeki örneklem sayısı azaltılmıştır. Daha sonra da KDD99 veri setindeki 41 özellik yerine daha az özellik kullanılarak nitelik sınıflandırma yapılabileceği ve bu şekilde daha kısa sürede saldırıların tespit edilebileceği gösterilmiştir.

Sınıflandırma işleminde ise başarıyı artırmak için, birden fazla makine öğrenme algoritması birlikte kullanılmıştır. Kullanılan sınıflandırıcılar: Adaboost, Karar Ağacı, Lojistik Regresyon, Naive Bayes, Rastgele Orman, Gradient Boosting, En Yakın K Komşu ve Yapay Sinir Ağları (Çok Katmanlı Algılayıcı) olmuştur. En başarılı sonuçların 3 veya 4 algoritmanın birleşmesinden elde edilen sonuçlar olduğu ve sınıflandırıcı füzyonlarında en çok Adaboost, Karar Ağacı, Lojistik Regresyon ve en yakın komşu algoritmalarının kullanıldığı ifade edilmiştir.

Genetik Algoritma-Nitelik Sınıflandırma-Ağırlık Belirlemenin birlikte kullanımının tek sınıflandırıcı sonuçlarından daha başarılı olduğu gösterilmiştir. Bu yöntem ile eğitim ve test süresi azaltılarak, doğruluk oranı değerleri daha yüksek bir sınıflandırıcı füzyonu elde edilmiştir. Füzyon yöntemlerinin nitelik sayısına göre sınıflandırma başarıları karşılaştırıldığında 41 nitelik için doğruluk oranı 0.79 iken GA ile seçimde 0.91 olmuştur. Tek sınıflandırıcı kullanıldığı durumda nitelik sayısı azalırken başarı oranının arttığı tespit edilmiştir. Tek bir algoritma kullanmak yerine birden fazla algoritmanın birleştirilerek kullanılmasında ise doğruluk oranının daha da arttırılabileceği belirtilmiştir.

Phyu Thi Htun ve Kyaw Thet Khaing [8] çalışmalarında sınıflandırma başarısını artırmak ve saldırıların doğru sınıflandırılmasını sağlamak için özellik seçimini başarıyla kullanmıştır. Özellik seçimi Hizmet Reddi (DoS) saldırılarının belirlenmesi için etkili olmuştur. Yani bir saldırının hizmet reddi saldırısı olarak nitelendirilebilmesi için hangi özelliklerin kullanılması gerektiği belirlenmiş ve buna göre KDD99 veri setindeki 41 özellik azaltılmıştır. Sınıflandırma yöntemleri olarak da Rasgele Orman Algoritması ve k-En Yakın Komşu Algoritması kullanılmıştır. Saldırıların tespit edilebilmesi için en yaygın kullanılan sınıflandırma algoritması k-En Yakın Komşu algoritmasıdır. En verimli ve en etkili algoritma ise Random Forest algoritmasıdır. Bu nedenle çalışmada bu iki algoritma tercih edilmiştir.

Bir ağdaki DoS saldırılarının sınıflandırılması için Rastgele Orman ve k-En Yakın Komşu algoritmalarına dayalı karma bir sınıflandırıcı ile bilinen ve bilinmeyen saldırıların WEKA aracını kullanarak tespit edilmesinde yüksek doğruluk sağlayabilme amaç edinilmiştir. Eğitim ve Test verilerinin aynı olmamasına dikkat edilmiş ve ezberci yaklaşımdan kaçınılmıştır. İlk olarak özellik azaltma işlemi Rasgele Orman Algoritması kullanılarak yapılmış daha sonra önerilen algoritmalar birlikte kullanılmıştır.

KDD99 veri setindeki DoS (Hizmet Reddi) saldırılarının (smurf, land, pod, teardrop, neptune, back) Random Forest algoritması ile tespit edilmesi sağlanmış ve %99.87 oranında doğruluk tespit edilmiştir. Bunu takip eden diğer sınıflandırıcı algoritması, k-en yakın komşu algoritması ile %99.85 doğruluk oranına ulaşılmıştır. Önerilen sistem ile iki algoritmanın birlikte kullanılması sonucunda %99.97 oranında doğruluk elde edilmiştir.

Sağıroğlu, Yolaçan ve Yavanoğlu [4] çalışmalarında saldırı tespit sistemlerini incelemişler ve yapay zekâ tabanlı zeki bir saldırı tespit sistemi tasarımı yapmışlardır. Geliştirilen bu tasarımın tespit ettiği saldırıların ve başarı oranlarının önceki çalışmalar ile karşılaştırılması yapılarak bu çalışmada uygulanan yaklaşımın başarı oranları değerlendirilmiştir. Elde edilen sonuçlardan, zeki yaklaşımın bu tip problemlere uygulanabilir olduğu gösterilmiştir.

Saldırı veri kümesi için istatiksel olarak en çok karşılaşılan saldırıların yer almasına dikkat etmişlerdir. KDD'99 veri kümesinin azaltılmış hali olan %10'luk veri kümesinden rasgele elde edilen 3 farklı eğitim kümesi ile çalışılmıştır. Eğitim veri kümelerinden farklı olan beş adet test veri kümesi oluşturulmuş ve sistem test edilmiştir. Bu farklı eğitim ve test kümelerinden elde edilen bilgiler ışığında; eğitim ve test verilerinin sayısının fazla olduğu durumlarda yüksek hesaplama zamanına ihtiyaç duyulduğu ve yapılan testler sonucunda az sayıda örneğin olduğu veri kümeleri kullanıldığında ise saldırı tespitinde başarının düştüğü gözlemlenmiştir.

Bu çalışma kapsamında gerçekleştirilen Zeki Saldırı Tespit Sistemi uygulamasında, farklı eğitim ve test kümeleriyle yapılan testlerden elde edilen en iyi sonucun %97,92 başarılı olduğu, en düşük sonucun ise %81,93 olduğu tespit edilmiştir.

Aggarwala ve Sharma çalışmalarında [9] KDD veri setinin 41 niteliğini kategorize ederek Temel, İçerik, Trafik ve Host olarak dört gruba ayırmıştır. Eğitim ve test verileri Weka aracındaki Rastgele Ağaç algoritması kullanılarak çalıştırılmış ve sonuçlar Tespit Oranı - Detection Rate (DR) ve Yanlış Alarm Oranı - False Alarm Rate (FAR) bazında değerlendirilmiştir. Bu çalışmada amaç saptama oranında iyileştirmeler yaparken yanlış alarm oranında en aza indirmeyi sağlayacak nitelik sınıfını tespit ederek veri setinin uygunluğunu arttırmaktır.

Veri seti üzerinde yapılan bu deneysel analiz sırasında, tespit oranı ve yanlış alarm oranı için dört nitelik sınıfının her birinin katkısı incelenmiş, minimum FAR ile maksimum DR elde etmek için veri setinin uygunluğunu geliştirmeye yardımcı olabilecek nitelikler gözlemlenmiştir. Saldırı Tespit Sistemlerinin değerlendirilmesi için kullanılan birçok ölçüm tekniğinin mevcut olduğu ancak en belirgin ve en sık kullanılan iki ölçüm tekniğinin Yanlış Alarm Oranı - FAR ve Tespit Oranı - DR olduğu kanıtlanmaya çalışılmıştır.

Dört öznitelik sınıfı: Temel, İçerik, Trafik ve Host şu şekilde tanımlanmıştır:

Temel (B): TCP bağlantılarının özellikleridir.
İçerik (C): Etki alanı bilgisi tarafından önerilen bir bağlantı içindeki özelliklerdir.
Trafik (T): İki saniyelik bir zaman penceresi kullanarak hesaplanan özniteliklerdir.
Host (H): İki saniyeden uzun süren saldırıları değerlendirmek için tasarlanmış özelliklerdir.

Temel sınıf özelliklerinin daha yüksek Tespit Oranı - DR gösterirken, Host özelliklerinin daha düşük Tespit Oranı - DR gösterdiği tespit edilmiştir. Aynı zamanda Trafik özellikleri daha düşük Yanlış Alarm Oranı - FAR gösterirken, İçerik özellikleri daha yüksek Yanlış Alarm Oranı - FAR göstermektedir. Temel sınıf özellikleri kaldırıldığında algılama oranının düştüğü, eklendiğinde ise yükseldiği tespit edilmiştir. En Kötü Yanlış Alarm Oranı – FAR İçerik sınıfı özellikleri için en yüksek değere ulaşmıştır. En Kötü Algılama veya Tespit Oranı – DP Trafik sınıfı özelliklerinde minimum seviyede gözlemlenmiştir. Her bir nitelik sınıfının iki ölçümü nasıl etkilediği tek tek incelenmiştir.

Sabhnani ve Serpen [10] çalışmalarında otomatik tekniklerle imza analizini birleştirerek kural geliştirmekte ve önerilen kural oluşturma tekniği sayesinde, KDD veri setini kullanarak Yönetici Hesabı ile Yerel Oturum Açma (Remote to Local Attack – R2L) saldırılarını tespit edebilmektedir. KDD veri setindeki iki R2L saldırısı, warezmaster ve warezclient saldırıları için, düşük yanlış alarm oranları ve yüksek doğru algılama oranları gözlemlenmiştir.

Bu çalışmada KDD veri setindeki eğitim ve test veri setleri birleştirilerek, seçilen R2L saldırılarının analizi yapılmıştır. Ayrıca önerilen kurallar ile algılama oranını artırmak ve yanlış alarmları azaltmak hedeflenmiştir. KDD veri kümesi çeşitli R2L saldırılarından oluşur. Bu deneyde sadece warezmaster ve warezclient saldırıları ele alınmıştır. Bu saldırılar ftp sunucularından veri yüklemeyi ve indirmeyi içeren ve buna yönelik imzalar barındıran herhangi bir ftp sunucusuna karşı yapılan tipik saldırıları temsil eder.

İlk olarak ele alınan belirli bir saldırı için KDD eğitim ve test veri setleri birleştirilmiştir. Birleştirilmiş KDD veri setindeki her bir kayıt, bu saldırıya ait olan özellikleri taşıyıp taşımadığına göre etiketlenmiştir. Daha sonra bu etiketlenmiş veri setine C4.5 karar ağacı algoritması uygulanmıştır.

Warezmaster saldırısı; dosya aktarma protokolü (File Transfer Protocol - FTP) sunucusuyla ilişkilendirilmiş sistem hatasını kullanan bir saldırı türüdür. Normalde misafir kullanıcılar hiçbir zaman FTP sunucusunda yazma izinlerine sahip değillerdir. Dolayısıyla sunucuya hiçbir zaman dosya yükleyemezler. Bu saldırı, bir FTP sunucusunun yanlışlıkla sistemdeki kullanıcılara yazma izinleri vermesi durumunda gerçekleşir. Bu saldırıyı önlemenin en basit ve açık yolu, FTP sunucusundaki kullanıcılara doğru izinler vermektir.

Warezmaster saldırısının gerçekleştirilebilmesi için, bir FTP bağlantısı sırasında dosya yükleme işleminin yapılması gerektiğinden, gözlemlenebilecek ilgili özellikler; devam etmekte olan bir FTP oturumu, yüklenen dosyalar ve gizli dizinlerin oluşturulmasıdır. Veri aktarımı sırasında kaynak ve hedef arasında büyük miktarda veri transferi gerçekleşir ve bu veri transferi belirli bir süre analiz edilerek dolaylı olarak gözlemlenirse veri yükleme işlemi için saldırı olup olmadığı tespit edilebilir.

Bu çalışmada Warezmaster saldırısını tespit etmek için oluşturulan kurallardan birincisi; bir FTP oturumu sırasında, kaynaktan gönderilen büyük miktarda verinin, hedefle karşılaştırılmasıdır. Diğer bir kural ise; bir misafir kullanıcının FTP bağlantısı üzerinden giriş yapması ve gizli dizinler oluşturmasıdır.

Warezclient Saldırısı; genellikle dosya indirme işlemi sırasında yapılan saldırı türüdür. KDD eğitim ve test veri setlerinde toplam 893 warezclient saldırı kaydı mevcuttur. Çok sayıda kaçırılan alarmın nedeni, bu saldırının özelliklerinin, FTP dosya indirme işleminin normal davranışına çok benzemesidir.

Tespit oranları ve yanlış alarm oranları, KDD eğitim seti ve değerlendirilen her saldırı için bir araya getirilen veri setleri üzerinden hesaplanmıştır. R2L kategorisindeki iki özel saldırı için sadece % 0,005 yanlış alarm oranı ve % 53,08 tespit oranı elde edilmiştir. Her iki R2L saldırısı üzerinde gerçekleştirilen imza analizinin, KDD veri kümelerindeki ve gerçek ana bilgisayar ve ağ trafiğinden elde edilen diğer veri türlerindeki diğer saldırılar için de tahmin edilebilecek derecede iyi bir tespit seviyesi sergilediği gözlenmiştir.

Bizim üzerinde çalıştığımız bu tez çalışmasında ise KDD 99 veri setinde yer alan örneklemlerden elde ettiğimiz yeni bir veri seti üzerinde Hizmet Engelleme ve Bilgi Tarama saldırıları keşfedilmeye çalışılmıştır. Çalışmamızda eğitim ve test veri setlerinin birebir aynı

olmamasına özen gösterilmiştir. Ayrıca veriler sadece saldırı verisi ve normal veri olarak değil aynı zamanda Hizmet Engelleme ve Bilgi Tarama Saldırısı olarak da sınıflandırılmakta ve başarı hesaplaması da saldırı türünün belirlenmesi şeklinde yapılmaktadır.

İlk olarak verinin saldırı verisi mi normal bir veri mi olduğuna karar verilmekte daha sonra bu veri bir saldırı verisi ise Hizmet Engelleme mi Bilgi Tarama mı şeklinde sınıflandırma yapılmakta son olarak da eğer bir Hizmet Engelleme saldırısı ise hangi tip Hizmet Engelleme saldırısı olduğu bir Bilgi Tarama saldırısı ise hangi tip Bilgi Tarama saldırısı olduğu tespit edilmektedir.

Geri Yayılım, Karar Ağacı, Ripper Kuralı ve Rasgele Orman Algoritması gibi 4 farklı algoritmanın kullanıldığı çalışmada en iyi performansı Rasgele Orman Algoritması sergilemiştir.

3. MATERYAL VE METOD

Bilgi ve iletişim sistemleri her geçen gün daha fazla kullanılmaları ile birlikte, bu sistemlerin güvenliğinin sağlanması hem ulusal güvenliğin, hem de rekabet gücünün önemli bir boyutu haline gelmiştir. Bilgi ve iletişim sistemlerinde bulunan güvenlik zafiyetleri, bu sistemlerin hizmet dışı kalmasına veya kötüye kullanılmasına, can kaybına, büyük ölçekli ekonomik zarara, kamu düzeninin bozulmasına ve/veya ulusal güvenliğin ihlaline neden olmaktadır [11].

Aktif siber savunma konsepti ele alındığında birden fazla, yöntem, teknik veya yaklaşım ortaya çıkmaktadır. Güvenlik uzmanları gelişmiş siber saldırıları tamamen durdurmanın çok mümkün olmadığının farkındadırlar. Bundan dolayı da zararı aza indirecek yöntemler geliştirmeye çalışmaktadırlar. Bu çabalar bazen aldatma, bazen yavaşlatma, bazen şaşırtmaca bazen karşı atak yaparak düşmanı engelleme ya da tamamen yok etme olarak ortaya çıkmaktadır. Bundan dolayı da savunmayı güçlü kılmak adına farklı metotlar öne sürülmektedir.

Saldırıların tanınması ve önceden tespit edilmesi, ayrıca gelebilecek saldırıların tahmin edilerek uyarı verilmesi veri madenciliği ve yapay zekâ gibi yöntemler ile sağlanabilir. Tekrar edilen işlerin otomatikleştirilmesi, gelen saldırıların anlamlandırılması ve az olan insan gücünün doğru yerde kullanılması daha güvenli bir sistem oluşturmak için önemlidir.

Zeki yaklaşımların kullanılmaya başlaması ile birlikte anormallik tespiti yaklaşımı biraz daha ön plana çıkmış ve bu sayede anormallik tespitinin yeni saldırıları tespit edebilme yeteneği arttırılmıştır. Yapay zekâ araştırmacılarının baştan beri ulaşmak istediği ideal, insan gibi düşünen ve davranan sistemler geliştirmek olmuştur.

3.1. Yapay Zekâ Teknikleri

Gelişen dünyanın bilişim teknolojisinde siber saldırılar en hızlı büyüyen tehditlerden birisi haline gelmiştir. Bilgiyi internet üzerinden erişilebilir kılmak için her gün yeni araçlar ve teknikler ortaya çıkmakta bu da güvenlik açıklarına neden olmaktadır. Bilgi güvenliği ve güvenilirliği için siber savunma kaçınılmaz hale gelmektedir.

Son yıllarda ağ teknolojilerinde yaşanan baş döndürücü gelişmeler hemen her işin bilgisayar ağları üzerinden yapılmasını mümkün hale getirmiştir. Bilgisayar sistemleri ve bilgisayar ağlarının gelişmesi ile aya insan göndermek, uluslararası ticaret yapmak, pilotsuz uçakları savaştırmak gibi işler yapılabilir hale gelmiştir. Fakat insanlar bunlara rağmen bilgisayar sistemleri ve bilgisayar ağlarına tam olarak güvenememektedirler. Güvensizliğin temelinde ise bilgi-işlemin saldırılar karşısında tam başarılı olamaması vardır. Güvenlik, sağlam bir mühendislik ve sosyal altyapı gerektirmektedir [12].

Siber tehditlerin ve saldırıların üstesinden gelmek için umut verici bir araç olan yapay zekâ teknikleri aynı zamanda kötü niyetli saldırganlar tarafından da kullanılabilir. Örneğin, makine öğreniminden yararlanılarak akıllı kötü amaçlı yazılımlar oluşturulmaktadır. Dolayısıyla, makine öğrenimi siber savunma için ümit verici çözümler sunarken aynı zamanda siber suçluların hedefledikleri saldırıları gerçekleştirmelerine de olanak sağlamaktadır. Siber savunma uzmanlarının, saldırı verilerini daha iyi anlamak için aktif olarak analiz ettiği sistemler üzerinde kullandıkları yapay zekâ teknikleri saldırganların da kullanıcılar hakkındaki verileri çalabilmeleri ve saldırılarını daha iyi hale getirmek için kullandıkları birer silah haline gelmiştir. Örneğin yasadışı erişim sağlayarak sisteme erişen bir saldırgan gelişen bu araçlar sayesinde hedeflenen kullanıcıların e-postalarını analiz edebilmekte, e-posta kalıplarını daha iyi anlayabilmekte ve daha iyi kimlik avı e-postaları oluşturmak için bu araçlardan yararlanabilmektedir.

Siber suçlular sistem ve uygulamaya yönelik güvenlik açıklıklarından faydalanmak ve kötü niyetli faaliyetlerde bulunmak için yapay zekâ teknolojilerinden yararlanmaktadırlar.

Yapay zekâ, sonuç çıkarma, tasarım yapma gibi faaliyetler ile düşünme arasında bağlantı kurarak, insan beynini taklit etmek için makinelerin nasıl kullanılacağı üzerine çalışır. İnsan gibi düşünen ve insan gibi olaylardan çıkarım yaparak cevap veren sistemler yapay zekâ ile çalışan sistemleri oluşturmaktadır.

Yapay zekânın amacı, normal olarak insan zekâsını gerektiren görevleri yapabilecek makineler yapmaktır. Yapay zekâ araştırmalarının amacı ise, insan varlığında gözlemlediğimiz ve "akıllı davranış" olarak adlandırdığımız davranışları gösterebilen bilgisayarlar yapmaktır.

3.1.1. Uzman sistemler (Expert systems)

Ancak bir uzman insanın çözebileceği karmaşık problemlerin bilgisayar ile çözümüne olanak sağlayan sistemlere uzman sistemler denir. Uzman sistemler bir yapay zekâ programlamadır. Genel olarak bir algoritma kullanmazlar, önemli olan veri tabanından çok bilgiye sahip olmaktır.

Uzman sistemler simülasyon, hesap yapma gibi işlevlere sahip olabilir. Bu sistemlerde birçok farklı model vardır ve en yaygını kural tabanlı (rule base) olanıdır. Kural tabanlı sistemler, sistem trafiğini inceleyip kurallar oluşturur ve saldırı tespiti sırasında belirlenen kurallara göre davranışlar sınıflandırır.

Bu sistemler, güvenlik derecesi seçimini önemli ölçüde kolaylaştırır ve sınırlı kaynakların üst düzeyde kullanımı için yol gösterir. Uzman sistemler kullanılarak saldırı tespiti yapılabilir [3].

Uzman sistemler, bilgisayarlaştırılmış danışma programlarıdır. Bunlar belirli bir problemi çözmek için uzmanın muhakeme sürecini ve bilgisini kullanırlar. Yapay zekâ teknolojisinin en fazla kullanılan uygulamalarından biridir. Bu sistemler özellikle organizasyonlar tarafından verimliliği artırmak ve uzman bulmanın gittikçe zorlaştığı yerlerde kullanılmak üzere talep edilmektedirler.

3.1.2. Akıllı ajanlar (Intelligent agents)

Akıllı ajanlar, etkileşimli olarak kullanıcıdan gelen emirler ile çalışan "davranışçı" ajanlardır. Bunlar kullanıcı ile birebir etkileşim içindedirler ve ondan gelecek komutlara göre davranışlarını düzenlerler.

Akıllı ajan, önceden var olan ve elindeki bilgilere dayanarak başarım ölçütünü yaparak karar veren bir mekanizmaya sahiptir. Akıllı ajan kullanan savunma sistemleri ile simülasyonla işbirliği yaparak Dağıtık Hizmet Dışı Bırakma (Distributed Denial of Service – DDoS) saldırılarına karşı savunma gösterebilir.

3.1.3. Genetik algoritmalar

Genetik algoritmalar siber savunmada, trafik verilerine basit kurallar uygulamak için kullanılabilir. Bu kurallar, anormal trafik verilerinden normal verileri ayırmak içindir. Genetik algoritmalar öncelikle küçük boyutlu rastgele üretilmiş kural kümesleri ile başlar, daha sonra bu kural kümesi genişletilir [3].

3.2. Makine Öğrenme

Makine öğrenmesi, matematiksel ve istatistiksel yöntemler kullanarak mevcut verilerden çıkarımlar yapan ve bu çıkarımlarla bilinmeyene dair tahminlerde bulunan yöntemleri konu edinen bir bilim dalıdır [13].

Makine öğrenmesi, mevcut verilerden öğrenme sağlayarak yeni veriler üzerinde tahminler yapan algoritmalar içermektedir. Makine öğrenme algoritmalarında başarı hesaplama, algoritmanın eğitimi sırasında kullanılan verilere bağlıdır. Eğitim aşamasında ne kadar kapsamlı ve düzgün veriler kullanılırsa, test aşamasında o kadar sağlıklı sonuçlar elde edilir.

Günümüzde en sık karşılaşılan problemlerden biri haline gelen siber saldırıların tespit edilebilmesi için makine öğrenme algoritmalarından faydalanılmaktadır. Bir saldırıyı tespit edebilmek için saldırı verisini iyi tanımlamak gerekmektedir. Bunun için sistemde gerçekleşen her olay büyük veri kümeleri halinde depolanarak analiz edilir ve kötü niyetli faaliyetler yani saldırı verileri için model tanımı yapılır. Bu model tanımına bağlı olarak sisteme gelen diğer veri akışları benzer aktiviteler ile ilişkilendirilerek normal bir veri trafiği mi yoksa saldırı verisi mi olduğunun çıkarımı yapılır.

Modellerin eğitim veri kümesi genellikle önceden tanımlanmış ve kaydedilmiş veri setlerinden oluşur. Bunlar daha sonra tehdit sayılabilecek verilerin izlenmesi, tanımlanması ve tehditlere karşı cevap verilebilmesi için kullanılabilir. Örneğin, kötü amaçlı yazılımların çeşitli davranışlarının belirlenmesi ile veri setlerinde buna göre sınıflandırma yapılabilir. Zararlı yazılımların davranışları analiz edilerek modellemeler oluşturulur ve yeni gelen zararlı bir yazılımın tespit edilmesinde ve sınıflandırmasında öğretilen sisteme uygun olarak verilen tepkiler otomatikleştirilir. Bu sayede benzer saldırıların tespit ve müdahale sürecinin

otomatikleştirilmesine olanak sağlanmış olur ve yeni bir tehdit türü hızlı bir şekilde tanımlanıp sınıflandırılabilir.

Daha önce gerçekleşen siber saldırıların analiz edilmesi ve hangi ağ türlerinin hangi saldırılar için kullanıldığı veya hangi portlar üzerinden hangi saldırıların gerçekleştirildiği gibi veriler kullanılarak çıkarımlar yapılabilir. Bu sayede belirli bir ağ alanına göre bir saldırının olasılığının ve etkisinin ölçülmesi sağlanabilir ve böylece kuruluşların siber saldırılara mağdur olma riski azaltılmış olur.

Makine öğrenimi, güvenlik analistleri tarafından gerçekleştirilen ve rutin olarak tekrarlayan siber güvenlik faaliyetlerinin otomatikleştirilmesi için de kullanılabilir. Örneğin, bazı siber saldırıları başarılı bir şekilde tespit etmek ve bunlara cevap vermek için güvenlik analistleri tarafından geçmiş saldırı kayıtlarının ve raporlarının analiz edilmesi gerekmektedir. Bu bilgiler sayesinde benzer saldırıların tanımlanması ve insan müdahalesi olmadan, tanımlanmış modele göre yanıt verebilecek bir yapı oluşturulması sağlanmış olur.

Tam olarak güvenlik süreçlerinin otomatikleştirilmesi mümkün olmasa da; kötü amaçlı yazılımların algılanması, güvenlik açıklıklarının değerlendirilmesi gibi durumlarda makine öğrenmesi teknikleri ile sistemlerin otomatikleştirilebileceği söylenebilir.

Makine öğrenme modeli, verilerden bazı bilgiler edinir ve bu bilgileri eğitim verisi olarak kaydeder. Daha önce hiç karşılaşmadığı veya öğrenmediği bir bilgi ile karşılaştığında ise eğitim verilerinden edindiği bilgiye dayanarak bir çıkarım yapar ve tahminde bulunur. Bu nedenle eğitim verileri makine öğrenme algoritmaları için çok önemlidir. Gereksiz veya işe yaramayan veriler ile öğrenme yapılırsa çıkarım yapmak ve tahminde bulunmak zor olacak ve yanlış pozitif veya yanlış negatif oranları artacaktır. Eğitim verimiz ne kadar kaliteli olursa öğrenme o kadar başarılı olacak ve daha az hata oranı ile yüzde yüze yakın doğru sınıflandırmalar yapılacaktır.

Son yıllarda ortaya çıkan sorunlardan birisi de siber suçluların bir makine öğrenme modelindeki eğitim setine erişmesi ve verileri değiştirmesidir. Sistem yanlış veriler ile eğitildiğinde yanlış sonuçlar üretmekte ve yanlış sınıflandırmalar yapmaktadır. Bu nedenle makine öğrenimine dayalı projeler için geniş güvenlik önlemleri alınmalı ve veri kümeleri belli periyotlar ile takip edilmelidir.

Bilişim sistemlerinin gelişmesiyle, saldırı tespit sistemlerinin kullanımı önem kazanmıştır. Bu sistemlerin çalışması, genellikle sınıflandırma problemi çerçevesinde değerlendirilebilir. Günümüzde makine öğrenmesi yöntemleri bilgisayarların daha doğru eylemler gerçekleştirmesi amacıyla birçok farklı şekilde kullanılmaktadır. Makine öğrenmesi, bilgisayarların gerçekleştirdikleri eylemleri daha doğru bir hale getirmek üzere değiştirmesi veya uyarlaması olarak tanımlanabilir.

Makine öğrenmesini bilgisayarların bir başarım ölçütünü, örnek veri veya geçmiş deneyim kullanarak eniyilemesi olarak da tanımlayabiliriz [14].

3.3. Saldırı Tespit Sistemleri

Günümüzde ağların kompleks bir yapıya sahip olması, internet başta olmak üzere diğer ağlara birçok erişim noktası ile bağlı olunması, siber saldırıların her geçen gün çeşitlenmesi ve artması, aynı zamanda bu karmaşık ağ sistemlerinin artık sadece şifreleme veya güvenlik duvarı ile korunamayacağı gibi gerçeklikler ağ trafiğinin devamlı izlenip saldırı girişimlerinin gerçek zamanlı olarak tespitini kaçınılmaz kılmıştır [15].

Gelişen teknoloji ile birlikte internet ortamında yer alan uygulamaların tümü saldırıya açık halde bulunmaktadır. Bu uygulamaların güvenliklerinin test edilmesi için kullanılan virüs programları sadece tanımlanmış kötü niyetli yazılımları tespit edebilmekte, güvenlik duvarı ürünleri ise zararlı olarak tanımlanan IP adreslerinden gelen istekleri engelleyebilmektedir. Ancak bu yazılımlar olası saldırıların çeşitliğinin artması ile yeterli miktarda güvenlik sağlayamamaktadır. Ağ saldırılarının etkin bir şekilde saptanması için gerekli verilerin toplanması ve doğru bir şekilde analiz edilmesi önemlidir. Herhangi bir saldırı veya tehdit durumunda ağ üzerindeki tüm aktivitelerin izlenmesi ve bu aktivitelerin analiz edilmesi oldukça zordur. Bu ürünlere ek olarak geliştirilen Saldırı Tespit Sistemleri (Intrusion Detection System – IDS), bir ağın ya da sistemin yaptığı aktiviteleri kontrol ederek saldırıları tespit etmeye çalışmaktadır.

Saldırı Tespit Sistemi (Intrusion Detection System - IDS) ve Saldırı Önleme Sistemi (Intrusion Prevention System - IPS), siber alanda en önemli savunma teknolojileridir. Yıllardır izinsiz girişlerin tespit edilebilmesi için muazzam bir çaba sarf edilmiştir ancak tam anlamıyla güvenlik sağlanamamıştır.

Saldırı Tespit Sistemleri ağ üzerinde gerçekleşen tüm trafiği inceleyen, analiz eden ve anlamlı veriler elde etmemize yarayan sistemlerdir. Saldırı Engelleme Sistemleri ise saldırıların saptanıp önlenmesi için çalışan sistemlerdir.

Saldırı Tespit Sistemleri ağ üzerindeki tüm verileri kontrol ederken olası tehditleri tanımlayabilmek için verilerin özelliklerini incelemekte ve bunlarla ilgili çıkarım yapabilmektedir. Normal ağ trafiğinde hangi verinin hangi Internet Protokol (IP) adreslerinden gelebileceği, hangi bağlantı noktası (port) üzerinden iletişim sağlamak isteyebileceği veya hangi protokolü kullanacağı sistemler üzerinde kaydedilir. Aynı şekilde saldırı verilerini analiz edebilmek için de ön tanımlamalar yapılır. Örneğin daha önceden tespit edilmiş saldırılara ait özellikler sisteme tanıtılarak buna benzer bir trafiğin oluşması engellenebilir veya normal bir trafik ile karşılaştırma yapılarak daha önce karşılaşılmamış bir saldırı olasılığı engellenebilir.

Kısacası bir saldırı tespit sistemi, hem kurum dışından hem de kurum içinden gelen tehditleri içeren olası güvenlik tehditlerini tanımlamak için bir bilgisayardaki veya bir ağdaki farklı alanlardaki bilgileri toplar ve analiz eder. Uygulamaların güvenliklerinin test edildiği, kötü yazılım içerenlerinin tespit edilerek silindiği, kötü niyet içeren bağlantıların tespit edilerek reddedildiği güvenlik duvarı ve virüs programları gibi yazılımlar saldırıları büyük ölçüde engellemektedir; ancak olası saldırıların çeşitliliği ve büyüklüğü göz önüne alındığında bu yazılımlar yeterli olmamaktadır. Çözüm olarak öne çıkan Saldırı Tespit Sistemleri, bir ağın ya da sistemin yaptığı aktiviteleri kontrol ederek saldırıları tespit etmeye çalışan, engellemek için karşı girişimde bulunmayan gerçek zamanlı çalışan yazılım ürünleridir.

Saldırı Tespit Sistemlerinde karşılaşılabilecek en büyük problem hatalı onaylama (False Positive) ve hatalı reddedilme (False Negative) oranlarıdır. Bu oranların minimum düzeye indirilebilmesi için eğitim verilerinin geniş tutulması ve iyi seçilmesi gerekmektedir.

Siber güvenlik analistleri genellikle birden fazla olaya cevap vermek için zaman harcamak zorundadırlar. Ancak yapay zekâ ve insan zekâsı kombinasyonunun, her birinin kendi üretebileceğinden çok daha iyi sonuçlar ürettiği görülmektedir. Bu birleşim sonucu oluşacak mimari ile siber savunma alanında uzman düzeyinde bir karar verme yeteneğine sahip mekanizma oluşturulabilir. Oluşturulan sistem bazen yanlış pozitifler içerebilir. Makine öğrenimi sınıflandırıcıları yanlış pozitifler ile doğru pozitifler arasında ayrım yapmak için

uyarı verisi konusunda eğitilebilmektedir. Böylece yalnızca doğru pozitifleri içeren senaryolarda analistleri uyaracak otomatik bir sistem oluşturulabilir.

Saldırı Tespit Sistemlerinde doğru sınıflandırma yapmak için kullanılan bazı durumlar vardır ve bu durumlara göre algoritmanın başarısı ölçümlenebilir. Başarı ölçümünde ilk adım verinin normal veri mi yoksa saldırı verisi mi olduğuna karar vermektir. Ancak bazen saldırı verileri normal bir veri gibi, normal veriler de saldırı verisi gibi sınıflandırılabilir.

Başarı ölçümlerinde karşımıza çıkabilecek bazı durumlar şu şekilde sınıflandırılabilir:

1. Durum => Normal bir giriş verisinin normal olarak sınıflandırılması: Doğru Pozitif (DP);
2. Durum => Saldırı verisinin normal bir veri olarak sınıflandırılması: Yanlış Pozitif (YP);
3. Durum => Normal bir verinin saldırı verisi olarak sınıflandırılması: Yanlış Negatif (YN);
4. Durum => Bir saldırı verisinin saldırı olarak sınıflandırılması: Doğru Negatif (DN)'tir.

Saldırı tespit sistemlerinde, saldırı tespit yöntemi olarak anormallik tespiti ve kötüye kullanım tespiti olmak üzere iki farklı yaklaşım kullanılmaktadır. Anormallik tespitine dayanan yaklaşım; sistemdeki kullanıcıların davranışlarını modellerken, kötüye kullanım tespitine dayanan yaklaşım saldırganların davranışlarını modeller [16].

3.3.1. İmza tabanlı saldırı tespit sistemleri

İmza tabanlı saldırı tespit sistemleri, saldırı tespiti için daha önceden oluşturulmuş saldırılara ait imza barındıran bir veri tabanı tutmaktadır. Sisteme gelen her istek, bu veri tabanı ile karşılaştırılmaktadır. Bu karşılaştırmaya göre saldırı verisi olup olmadığı analiz edilmektedir. Bu yöntemdeki en önemli problem imzaların saldırganlar tarafından değiştirilebilme durumudur. Bu nedenle bilinen saldırıları tespit ederken iyi sonuç üretmesine rağmen veri tabanında yer almayan saldırı tiplerini bulmakta yetersiz kalmaktadır. Saldırı, sistemin normal davranışından farklıdır.

İmza tabanlı bu yaklaşımın çalışma mantığı çok basittir ve anti virüs yazılımlarına benzemektedir. Bir saldırının özellikleri analiz edilir ve saldırı imzaları oluşturulur. Bu saldırı imzalarından oluşturulan veri tabanı saldırıları tespit edebilmek için kullanılır. Bilinen saldırıları algılamada yüksek hassasiyet gösterirler. Ancak önceden bilinmeyen yeni

saldırıları tespit etme kabiliyetleri yoktur. Bu tür saldırı tespit sistemlerinin daha doğru sonuç verebilmesi için yeni saldırılar tespit edildiğinde veri tabanı hemen güncellenmelidir. Ancak her ne kadar güncel bir veri tabanına sahip olunsa da saldırı algılama yeteneği var olan imza kuralları ile sınırlıdır.

3.3.2. Anormallik tabanlı saldırı tespit sistemleri

Anormallik tabanlı saldırı tespit sistemleri normal ağ trafiği ile şüpheli trafiği birbirinden ayıracak mekanizmalara sahiptir. Anomali tabanlı sistemlerde normal bir veri için ön tanımlama yapılmaktadır. Gelen her istek bu tanımın özellikleri ile karşılaştırılmakta ve uygunluk durumuna göre saldırı verisi mi yoksa normal bir veri mi olduğuna karar verilmektedir. Bu yaklaşımlar, bilinen saldırılar kadar bilinmeyen saldırıları da tespit edebilirler. Bundan ötürü, ağ saldırılarını tespiti için anormal durumu belirlemeye yönelik makine öğrenmesi tabanlı sistemlerin kullanılması son yıllarda önem kazanmıştır.

3.3.3. Hata algılama tabanlı saldırı tespit sistemleri

Hata algılama tabanlı saldırı tespit sistemleri yalnızca önceden tanımlanmış saldırıların özelliklerine dayanarak bilinen saldırıları tespit etmeye çalışan ve hatalı verileri algılayan sistemleridir. Bir saldırı farklı şekillerde gerçekleşebilir ve bu saldırı tespit sistemlerinin doğruluğu sadece saldırı bilgisinin ne kadar iyi olduğuna bağlıdır. Saldırı verilerine ait bilgiler önceden tanımlanır ve sistem üzerine gelen her istek bu sistemlerin tespitine sunulur.

3.3.4. Kural tabanlı saldırı tespit sistemleri

Kural tabanlı saldırı tespit sistemleri eğer – öyleyse (if-then) mantığı ile çalışırlar. Kullanıcı davranış ve faaliyetlerini denetleyerek bir istatistik oluşturur ve bu davranışların dışında bir durum söz konusu olduğunda alarm üretir. Yani sistem üzerinde normal davranış profili oluşturularak sistem tarafından kullanıcının davranışları izlenir. Beklenenin dışında bir davranış sergileniyorsa izinsiz giriş veya kötüye kullanım olarak algılanır.

3.3.5. Veri madenciliği tabanlı saldırı tespit sistemleri

Bu yaklaşımda izlenen sistem üzerindeki veriler kabul edilebilir veya kabul edilemez olarak sınıflandırılmakta ve buna göre etiketleme yapılmaktadır. Yani sistem kabul edilen verileri normal veriler olarak tanımlar, kabul edilemeyen verileri saldırı verileri olarak tanımlar. Yeni gelen veriler bu kategorilere göre sınıflandırılır.

3.3.6. Makine öğrenme tabanlı saldırı tespit sistemleri

Bu tür saldırı tespit sistemlerinde makine öğrenim teknikleri anomali tespiti için yüksek başarı sergilemektedir. Yeni saldırıların daha iyi tespit edilmesi için veri madenciliği kombinasyonuyla birlikte makine öğrenme teknikleri kullanılmaktadır.

Genel olarak bir saldırı tespit sisteminin başarısı saldırıyı etkin bir şekilde tespit etme yeteneğine bağlıdır. Saldırı tespit sistemleri herhangi bir saldırı durumunda sistem sorumlularını ikaz edebilecek alarmlar üretir. Herhangi bir önleme faaliyetinde bulunmazlar. Saldırı önleme sistemleri ise saldırıyı tespit etme, izole etme ve engelleme yeteneğine sahiptirler.

3.4. KDD99 Veri Kümesi

KDD 99; 1999 yılında DARPA veri kümesinin bazı önişlemlerden geçirilmesi ile elde edilmiş 41 özellikten oluşan bir veri kümesidir. Bu veri kümesinin amacı, son yıllarda farklı tekniklerle gerçekleştirilen Saldırı Tespit Sistemleri için eğitim ve test işlemlerinde kolaylık sağlamaktır. KDD veri kümesi ile eğitim ve test sonuçlarının daha hızlı alınabilmesi yapılan çalışmaların sonuca ulaşmasını kolaylaştıran bir faktör olmuştur.

KDD-99 veri seti saldırı tespiti için yaygın olarak kullanılmaktadır. Bu veri seti 4,8 milyon civarında çok büyük miktarda veri içerdiğinden çalışmamızda Kddcup.data_10_percent.gz'den alınan örnek veriler eğitim amacıyla kullanılmıştır [17]. Bu veri seti KDD-99 veri setinin %10'undan oluşmaktadır. Normal davranış sergileyen kayıtlar normal veri olarak etiketlenmektedir. Bunların dışında kalan kayıtlar ise saldırı verisi olarak etiketlenmektedir.

Veri seti içerisinde 4 temel saldırı sınıfı ve normal veriler bulunmaktadır. Bu 4 temel sınıfta ise toplamda 22 tip saldırı verisi yer almaktadır. Veri setinin içerisinde 41 adet bağımsız 1 adet bağımlı değişken bulunmaktadır. Bu değişkenler 4 temel saldırı sınıfına ait kayıtları içerir. Bağımsız değişkenler; ağa gelen bağlantı ile ilgili protokol türü, kullanılan servisler, bağlantının normal bir şekilde sonlanıp sonlanmadığı gibi bağlantı hakkındaki bilgiler, bir bağlantı içerisinde çalıştırılan komut sayısı ve yanlış giriş işlemi sayısı gibi gerçekleştirilen bağlantılarla ilgili çeşitli değerleri tutmaktadır. Bağımlı değişken ise kaydın normal bir bağlantı mı saldırı mı olduğunu belirtir.

Bu veri setindeki her saldırı 4 ana sınıfa dahildir. Bu saldırı sınıflarını aşağıdaki şekilde ifade edebiliriz:

3.4.1. Hizmet engelleme saldırıları (Denial of service – DoS attacks)

Bu saldırılar genel olarak TCP/IP protokol yapısındaki açıklardan faydalanılarak sistemin tüm kaynaklarını tüketip hizmet veremez hale getirmeye ve bir sunucuya birden çok bağlantı isteği göndererek yasal kullanıcıların hizmet almasını engellemeye yöneliktir. Hizmet engelleme saldırılarına örnek olarak "teardrop - gözyaşı", "ping of death - ölüm pingi - İnternet Kontrol Mesajı Protokolü (Internet Control Message Protocol - ICMP) isteği", "smurf – amplification - büyütme " ve "land - çakışma" verilebilir.

Gözyaşı (Teardrop) saldırıları; bir bilgisayara internet üzerinden gelen paketlerin, bilgisayarda bölünerek aktarılması sırasında gerçekleşen saldırılardır. Paket verilere ayrıştırılırken pakette bulunan nesnenin başlangıcından, aynı nesnede verilen bir eleman veya noktaya kadar olan mesafeyi gösteren birimler (ofsetler) kullanılır. Bu ofset bilgilerinin çakışmaması gerekmektedir.

Bu tür saldırılarda IP paketlerinin parçalara bölünmesi sistemi kullanılır. Parçalanan IP paketi bir paket ile alıcı sisteme paketin parçalarını belirtir. Gözyaşı (Teardrop) saldırılarında, paketi gönderen saldırgan, pakete üst üste gelecek şekilde bu parçalanmayı karıştıracak ofset değerleri göndererek sistemi bozmayı amaçlar [18].

Paketi alan bilgisayar, böyle bir durumu kontrol edebilecek mekanizmaya sahip değilse, sistem çöker [19].

Ölüm isteği - Ping of death saldırılarının amacı büyük boyutlu paketler göndererek sistemi bozmaktır. Saldırgan büyük miktarda Internet Kontrol Mesajı Protokolü (ICMP) isteği göndererek saldırıyı gerçekleştirir. Saldırı Tespit Sistemleri tarafından kolaylıkla fark edilen bir saldırı türüdür.

Büyütme (Smurf – Amplification) saldırılarında, saldırgan hedef bilgisayarın IP adresinden sistemin broadcast adresine Internet Kontrol Mesajı Protokolü (ICMP) isteğinde (ping) bulunur. Ancak ping paketi, hedef makinenin IP adresinden geliyormuş gibi görüneceği için ağ üzerindeki bütün makineler, hedef makineye yanıt göndermeye çalışır. Hedef makine bu trafiği karşılayamaz ve bağlantı kesilir [19].

Çakışma - Land saldırılarında hedef sisteme kaynak IP ve hedef IP adresi aynı olan paketler gönderilerek sistemin çakışması sağlanır. Yani sistemin IP adresi hem kaynak adres hem de hedef adres olarak gösterilir [20].

3.4.2. Kullanıcı hesabının yönetici hesabına yükseltilmesi

Kullanıcı hesabının yönetici hesabına yükseltilmesi saldırıları (User to Root Attack – U2R) saldırganın sistemdeki geçerli bir kullanıcı hesabına erişim sağlayıp mevcut sistemin zayıflıklarına dayanarak yönetici - root yetkisini ele geçirdiği saldırı türleridir. Yani sisteme erişim yetkisi olan fakat yönetici yetkilerine sahip olmayan bir kullanıcının yönetici haklarını elde etmesidir [21]. “Yük modülü”, “Arabellek taşması”, “Kök kitlesi - rootkit”, “perl” gibi, yetkisiz bir kullanıcının algılanmadan bilgisayar sisteminin kontrolünü ele geçirmesini sağlayan ve bir dizi kullanıcı hesabının yönetici hesabına yükseltilmesine ait çeşitli saldırı türleri vardır.

3.4.3. Yönetici hesabı ile yerel oturum açma

Kullanıcı yetkisine sahip olunmadığı durumda, hesabı olmayan bir saldırganın, mevcut makine açıklarına bağlı olarak hedef ağdaki bilgisayara bazı paketler göndererek misafir ya da başka bir kullanıcı olarak izinsiz erişim yapılması ile gerçekleşen saldırılar yönetici hesabı ile yerel oturum açma (Remote to Local Attack – R2L) saldırılarıdır [21].

R2L saldırı tipleri “Yasadışı yollarla sistem kaynaklarına erişim - phf”, “Misafir kullanıcı hesabı ile FTP sunucusuna zararlı yazılım yükleme - warezmaster”, “Normal bir kullanıcı hesabı ile FTP sunucusuna dosya indirme - warezclient”, “Casus - spy”, “Posta sunucusu saldırısı - imap”, “Ftp sunucusu saldırısı - ftp_write”, “Çok atlamalı gizlenme saldırısı - multihop” ve “Misafir kullanıcı şifresini değiştirme saldırısı - guess_passwd” şeklindedir.

3.4.4. Bilgi tarama saldırıları (Probing attacks)

Bu tür saldırılar bir sunucunun ya da herhangi bir makinanın geçerli IP adreslerini, ağdaki bilgisayar sayısını, bilgisayardaki kullanıcı sayısını ve kullanıcı bilgilerini, aktif giriş kapılarını (port) veya işletim sistemini öğrenmek için yapılır [21].

Bilgi tarama saldırısının çeşitleri: “Açık portların ve çalışan servislerin belirlenmesi için yapılan tarama saldırısı - nmap”, “Şeytan – satan”, “Ip süpürme – ipsweep” ve “Port süpürme – portsweep” dir.

KDD 99 veri setindeki saldırı sınıfları ve türleri Çizelge 3.1’de gösterilmektedir.

Çizelge 3.1. Saldırı sınıfları ve türleri

Saldırı Sınıfı	Açıklama	Saldırı Türü
Hizmet Engelleme (Denial of Service)	Servisleri hizmet veremez hale getirme	back, land, neptune, pod, smurf, teardrop
Yönetici Hesabı ile Yerel Oturum Açma (Remote to Local - R2L)	Makinedeki kullanıcı yetkilerini ele geçirme	ftp_write, guess_passwd, imap, multihop, phf, spy, warezclient, warezmaster
Kullanıcı Hesabının Yönetici Hesabına Yükseltilmesi (User to Root – U2R)	Normal kullanıcı hesabından root hesabına geçmeye çalışma	buffer_overflow, perl, loadmodule, rootkit
Bilgi Tarama (Probing)	Bilgisayar veya ağ bilgisini elde etmeye çalışma	ipsweep, nmap, portsweep, satan

KDD 99 veri setinde yer alan saldırı sınıflarının çeşitleri ve açıklamaları Çizelge 3.2’de yer almaktadır.

Çizelge 3.2. Saldırılar ve açıklamaları

Back	Bir istemcinin Apache web sunucusuna karşı birçok ters eğik çizgi içeren URL isteğinde bulunarak Hizmet Reddi Saldırısı yapması
Dict	Telnet bağlantısı üzerinden, hesap adı için basit değişkenler kullanarak geçerli bir kullanıcıya ait şifrelerin tahmin edilmesi
Eject	User –> root girişi başarılı olursa Solaris'te eject programını kullanarak arabellek taşmasına yol açılması
Ffb	Ffbconfig kullanarak arabellek taşması
Format	Fdformat kullanılarak arabellek taşması
ftp-write	Uzak FTP kullanıcısının FTP dizininde yazılabilir .rhost dosyasını oluşturması ve yerel oturum açması
Guest	Misafir hesabı için telnet üzerinden şifrenin tahmin edilmeye çalışılması
Imap	Imap portunu kullanarak arabellek taşması
Ipsweep	Birden fazla ana bilgisayar üzerinde port taraması veya ping işleminin izlenmesi
Land	Aynı kaynak ve hedef ile uzak bir ana makineye bir UDP paketi gönderilmesi ile gerçekleşen Hizmet Reddi Saldırısı
Loadmodule	Gizli olmayan yük modülü saldırısı
Multihop	Bir kullanıcın bir makinayı ele geçirmesi
Neptune	Bir veya birden fazla port üzerinden gerçekleştirilen Syn Seli Hizmet Reddi Saldırısı
Nmap	Nmap aracı kullanarak ağ taraması
Perlmagic	Kullanıcı kimliğini perl betiğinde root olarak ayarlayan ve bir root kabuğu oluşturan Perl Saldırısı
Phf	İstemcinin yanlış yapılandırılmış bir web sunucusuyla bir makinede rasgele komutlar çalıştırmasına izin veren çalıştırılabilir CGI betiği
Pod	Ping ile yapılan Hizmet Reddi Saldırısı
Portsweep	Bir sunucu üzerinde çalışan servisleri belirlemek için yapılan tarama saldırısı
Rootkit	Bir kullanıcının bir veya daha fazla bileşenini kurduğu yer
Satan	İyi bilinen zayıflıkların tespit edildiği ağ tarama saldırısı Üç farklı seviyede çalışır. En basit seviye 0'dır.
Smurf	Icmp echo cevap seli ile yapılan Hizmet Reddi Saldırısı
Spy	Önemli bilgileri bulmak amacıyla bir kullanıcının tespit edilmekten kaçınarak makineye girmesi ve erişim elde etmek için birkaç farklı zafiyet yöntemi kullanması

Çizelge 3.2. (devam) Saldırılar ve açıklamaları

Syslog	syslog servisi için çözülemeyen kaynak ip adresi ile 514 numaralı bağlantı noktasına bağlanarak gerçekleştirilen Hizmet Reddi Saldırısı
Teardrop	Hatalı parçalanmış UDP paketlerinin bazı sistemlerin yeniden başlatılmasına neden olduğu Hizmet Reddi Saldırısı
Warez	Kullanıcının anonim bir FTP sitesine giriş yaparak gizli bir dizin oluşturduğu saldırılar
Warezclient	Daha önce warezmaster tarafından adsız bir FTP aracılığıyla gönderilmiş yasadışı yazılımların kullanıcılar tarafından indirilmesi
Warezmaster	Telif hakkıyla korunan yazılımların yasa dışı kopyalarının FTP sunucusuna yüklenmesi

3.5. WEKA Uygulaması

Yapılan testler WEKA (Waikato Environment for Knowledge Analysis) uygulaması üzerinde gerçekleşmiştir. WEKA, kullanımı ücretsiz, açık kaynak kodlu, içerisinde pek çok sınıflandırma, regresyon, demetleme, bağıntı kuralları, yapay sinir ağları algoritmaları ve önişleme metotları barındıran, yaygın kullanılan bir veri madenciliği aracıdır [22]. WEKA önceden işlenen verilerin makine öğrenmesi teknikleri ile sınıflandırılmasında kullanılmaktadır. Waikato Üniversitesi'nde açık kaynaklı bir Java kodu olarak geliştirilmiştir ve GPL lisansı altında dağıtılmaktadır. Bu uygulamanın bir ekran görüntüsü Şekil 3.1'de gösterilmektedir.

Şekil 3.1. WEKA ekran görüntüsü

Makine öğrenimi ve istatistik için birçok kütüphane WEKA'da mevcuttur. Örneğin, veri ön işleme, veri hazırlama, sıralama, gruplandırma, özellik seçimi ve özellik çıkarma gibi

işlemleri bu kütüphaneler yardımıyla yapabiliriz. Ayrıca WEKA sınıflandırma, kümeleme ve ilişkilendirme kuralları için birçok algoritmayı da desteklemektedir.

Bu çalışmada WEKA'nın 3.8 versiyonu kullanılmıştır. Bu program; arff, arff.gz, names, data, csv, c45, libsvm, dat, bsi, xrff, xrff.gz uzantılı verileri, verilerin depolandığı veri tabanlarını ve URL'lerin kullanımını da desteklemektedir.

Model başarımının ölçülmesinde kullanılan en yaygın ölçüt, modele ait doğruluk oranıdır. Doğru sınıflandırılmış örnek sayısının (TP+TN), toplam örnek sayısına (TP+FP+TN+FN) oranıdır. Hata oranı ise bu değerin 1'e tamlayanıdır. Diğer bir ifadeyle yanlış sınıflandırılmış örnek sayısının (FP+FN), toplam örnek sayısına (TP+FP+TN+FN) oranıdır.

Burada;

• TP: True Positive – Doğru örneklerin doğru olarak sınıflandırılması
• TN: True Negative – Doğru örneklerin yanlış olarak sınıflandırılması
• FP: False Positive – Yanlış örneklerin doğru olarak sınıflandırılması
• FN: False Negative – Yanlış örneklerin yanlış olarak sınıflandırılmasını ifade etmektedir [22].

3.6. Makine Öğrenme Algoritmaları

Veri sınıflandırma için kullanılabilecek birçok makine öğrenme tekniği vardır. Bu tez çalışmasında kullanılan makine öğrenme algoritmaları Geri Yayılma (Back-Propagation) Algoritması, Karar Ağacı Algoritması, Ripper Kuralı Algoritması ve Rasgele Orman Algoritması şeklindedir.

3.6.1. Geri yayılma (Back-propagation) algoritması

Geri yayılma algoritması, basitliği ve uygulamadaki görüş açısı gibi başarılarından dolayı ağ eğitimi için kullanılan en popüler algoritmalardan biridir. Bu algoritma; hataları geriye doğru çıkıştan girişe azaltmaya çalışmasından dolayı geri yayılım ismini almıştır. Geri yayılmalı öğrenme kuralı ağ çıkışındaki mevcut hata düzeyine göre her bir tabakadaki ağırlıkları yeniden hesaplamak için kullanılmaktadır [23]. Back propagation algoritması,

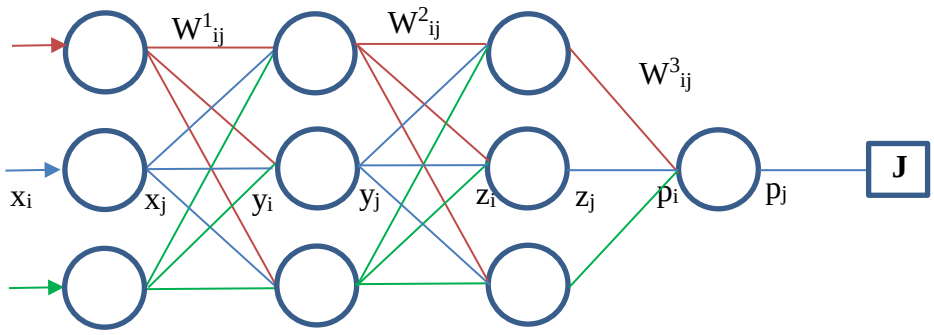
birçok pratik uygulama için çok geri kalmaktadır. Bu algoritmanın en büyük problemi, çok uzun eğitim süresine sahip olmasıdır.

Bir geri yayılımlı ağ modelinde giriş, gizli ve çıkış olmak üzere 3 katman bulunmakla birlikte, problemin özelliklerine göre gizli katman sayısını artırabilmek mümkündür. Geri yayılım çok katmanlı ağlarda kullanılan delta kuralı için genelleştirilmiş bir algoritmadır. Bu algoritma çok katlı ağlarda hesap işlerini öğrenmede kullanılabilmektedir. Geri yayılım ağında hatalar, ileri besleme aktarım işlevinin türevi tarafından, ileri besleme mekanizması içinde kullanılan aynı bağlantılar aracılığıyla, geriye doğru yayılmaktadır. Öğrenme işlemi, bu ağda basit çift yönlü hafıza birleştirmeye dayanmaktadır [24].

Bir ağ üzerindeki veriler eğitildikten sonra bir giriş vektörü yani bir nitelik üzerinde ağın gelecekte tahminde bulunmasına yardımcı olabilecek işlemler gerçekleştirilir ve bir çıkış vektörü oluşturulur.

Girdi Katmanı (Input Layer); bir ağa girdi olarak gelen ve öğrenilmesi istenen örneğin özniteliklerin giriş olarak verildiği katmandır. Bu katmanda nöron sayısı, giriş veri sayısı kadardır ve her bir giriş nöronu bir veri alır. Girdi katmanında öğretilecek örneklerin öznitelik sayısı kadar giriş nöronu bulunmalıdır [24].

Şekil 3.2'de gösterilen ağda üç giriş nöronu ve bir çıkış nöronu bulunmaktadır. Toplamda da bir giriş, bir çıkış ve iki gizli katman olmak üzere dört katmandan oluşmaktadır. Her gizli katmanda üç nöron yer almaktadır ve gizli katmanlar tamamen ağa bağlıdır. Her nöron/düğüm bir aktiviteyi işaret etmektedir.



Şekil 3.2. Katmanların gösterimi

x, girdi katmanına, y gizli katman 1'e, z gizli katman 2'ye ve p de çıktı katmanına karşılık gelmektedir.

xi herhangi bir giriş nöronunu temsil etmektedir. Bir değişken, i alt dizisine sahipse, değişken bu katmandaki ilgili nöronun girişi demektir. Eğer bir değişken j alt üyesine sahipse, değişken bu katmandaki ilgili nöronun çıktısı demektir. Örneğin, xi, ağa girdiğimiz herhangi bir giriş değerini ifade eder ve xj aslında xi'ye eşittir.

Gizli Katmanlar (Hidden Layers); giriş katmanı ile çıkış katmanları arasındaki katmanlardır. Ağın temel işlevini görür. Katman sayısı ve üzerindeki nöron sayısı problemden probleme değişebilmektedir. Bu katmalarda ileri yönlü hesaplamalar ve geri yönlü hata yayılımı yapılır. Giriş katmanından aldığı ağırlıklandırılmış veriyi probleme uygun bir fonksiyonla işleyerek bir sonraki katmana iletir. Katman sayısının çok olması hesaplama karmaşıklığına ve hesaplama süresinin artmasına sebep olur. Karmaşık problemlerde problem çözümü için genelde katman sayıları ve katmanlardaki nöron sayıları fazladır.

yi, ilk gizli katmandaki herhangi bir nöronun girişidir; önceki tüm nöronların ağırlıklı toplamıdır. yj ise gizli katmandaki herhangi bir nöronun çıktısıdır.

activation_function (yi) = sigmoid (yi) = sigmoid (weighted_sum_of_xj)

Çıktı Katmanı (Output Layer); yapay ağdaki öğrenilmesi istenen örneklerin sınıf bilgisinin veya etiket değerinin çıkış olarak hesaplandığı en uç katmandır [24]. Gizli katmandan aldığı veriyi ağın kullandığı fonksiyonla işleyerek çıktısını verir. Çıkış katmanındaki nöron sayısı, ağa sunulan her verinin çıkış sayısı kadardır. Bu katmandan elde edilen değerler söz konusu problem için çıkış değerleridir.

Ağırlıklar (Weights); girdilerin çıktılar üzerindeki etkisini ayarlayabilmek için kullanılan parametrelerdir. Ağırlıklar girdi değerleri ile çarpılarak ileri doğru iletilir. Ağdaki en kritik nokta, verilen ağırlıkların eğitim kümesine göre hatayı yayarak olması gereken optimum ağırlık değerlerini hesaplamaktır. Bu ağırlıklar pozitif, negatif veya sıfır olabilir. Ağırlığı sıfır olan nöronların çıktı üzerinde etkisi olmamaktadır [24].

Şekil 3.2'de ağırlıklar üç ayrı değişkende düzenlenmiştir: W^1, W^2 ve W^3. Her W, verilen

katmandaki tüm ağırlıkların bir matrisidir. Örneğin, W^1, giriş katmanını gizli katman 1'e bağlayan ağırlıklardır. $W^{layer}ij$, verilen bir katmandaki herhangi bir isteğe bağlı ağırlık anlamına gelmektedir. $W^{layer}i$, belirli bir katmandaki isteğe bağlı nöron i'yi bir sonraki katmana bağlayan tüm ağırlıklardır. $W^{layer}ij$, belirli bir katmandaki rastgele nöron i'yi, bir sonraki katmandaki rastgele nöron j'ye bağlayan ağırlıktır.

Geri Yayılma Algoritması bir ağı eğitirken aşağıdaki adımları izlemektedir:

1. Tasarlanan ağ oluşturulur ve eğitim verileri hazırlanır.
2. Tüm ağırlıklar rastgele değerlere sıfırlanır.
3. Test verileri kullanılarak bir ileri besleme gerçekleştirilir.
4. Ağda yer alan her ağırlıktaki hata türevlerini almak için geri yayılım gerçekleştirilir
5. Her ağırlığı, güncellemek için gradyan inişi yapılır ve yineleme sayısı artırılır.
6. Azami yineleme sayısına ulaşıldığında eğitim tamamlanır. Aksi takdirde, 3. adımdan başlayarak işlemler tekrarlanır.

Transfer Fonksiyonu; bir nörona gelen net girdiyi hesaplar. Bunun için değişik fonksiyonlar kullanılmaktadır. En yaygın olan ise ağırlıkların girdiler ile çarpımının toplamıdır.

Aktivasyon Fonksiyonu; hücreye gelen net değeri işleyerek nöronun bu girdiye karşılık aktivasyon çıktısını üretir. Aktivasyon fonksiyonunun probleme göre uygun seçilmesi ağın performansını ve başarı oranını önemli derecede etkiler. Hatayı geri yayarken her katmanda türev işlemi yapılır. Dolayısıyla seçilecek aktivasyon fonksiyonlarının türevlenebilir olması gerekmektedir [24].

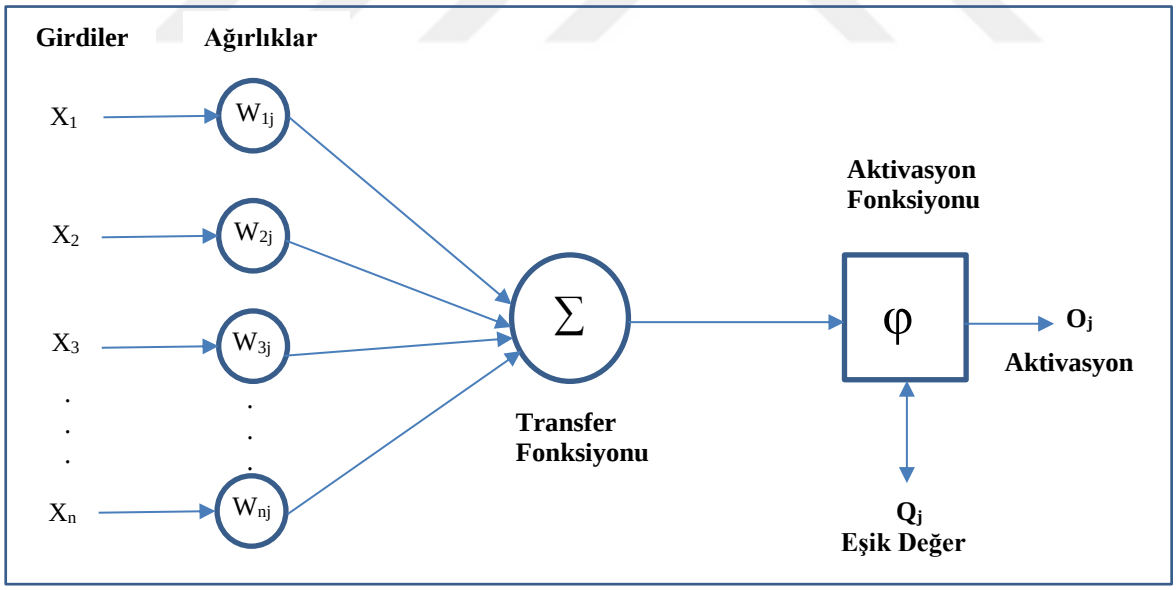
İleri besleme; çıktı aktivasyonlarını üretmek için girişler alınıp ağ üzerinde, katman bazında iletilir. Aslında ağımızı kullanırken, gerçekleştirmemiz gereken tek adım budur.

İleri besleme safhasında, giriş katmanındaki nöronlar veri değerlerini doğrudan gizli katmana iletir. Gizli katmandaki her bir nöron, kendi giriş değerlerini ağırlıklandırarak toplam değeri hesap eder ve bunları bir aktivasyon fonksiyonu ile işleyerek bir sonraki katmana veya doğrudan çıkış katmanına iletir. Katmanlar arasındaki ağırlıklar başlangıçta rastgele küçük rakamlardan seçilir [25].

Geri yayılım aşamasında hata fonksiyonu ele alınır ve her katman için ağırlık gradyanları hesaplanır.

Bir giriş verisi ağın ilk katmanında yer alan düğümlere uygulandığında en üst katman olan çıkış katmanına erişinceye kadar, bu veri üzerinde çeşitli işlemler gerçekleştirilir. Bu işlemlerin sonucunda elde edilen çıktı, olması gereken hedeflenen çıktı ile karşılaştırılır. Yapay Sinir Ağı çıkışı ve hedeflenen çıkış değerleri arasındaki fark, her çıktı düğümü için bir hata sinyali olarak hesaplanır. Hesaplanan hata sinyalleri, her çıktı düğümüne karşı gelen ara katmandaki düğümlere aktarılır. Böylece ara katmandaki düğümlerin her biri toplam hatanın sadece hesaplanan bir kısmını içerir.

Bu süreç her katmandaki düğümler toplam hatanın belirli bir kısmını içerecek şekilde giriş katmanına kadar tekrarlanır. Elde edilen hata sinyalleri temel alınarak, bağlantı ağırlıkları her düğümde yeniden düzenlenir [25]. Son olarak, ağırlıkları güncellemek için hesaplanan gradyanlar kullanılır. Algoritma Şekil 3.3'de ayrıntılı bir şekilde ifade edilmektedir.



Şekil 3.3. Geri yayılım algoritmasının gösterimi

Çıkış katmanındaki, her bir nöron ağırlıklandırılmış değeri hesaplandıktan sonra, bu değer yine aktivasyon fonksiyonu ile karşılaştırılarak mevcut hata minimize edilmeye çalışılır. Hata değeri belli bir mertebeye ininceye kadar iterasyon işlemine devam edilir. Böylece ağın eğitim aşaması tamamlanmış olur. Katmanlar arasındaki bağlantılardaki

ağırlık değerleri eğitimi tamamlamış ağdan alınarak deneme safhasında kullanılmak üzere saklanır.

Geri yayılım algoritması, yapay sinir ağı araştırmaları için önemli bir buluştur. Ama birçok pratik uygulama için çok geri kalmaktadır. Bu algoritmanın en büyük problemi, çok uzun eğitim süresine sahip olmasıdır.

3.6.2. Karar ağacı algoritması

Bir karar ağacı, belirli bir parametreye dayanarak verilerin sürekli olarak bölündüğü bir tür denetimli makine öğrenmesidir. Ağaç; verilerin nerede bölündüğünü belirten karar düğümleri ve nihai sonuç olan yapraklar olmak üzere iki varlıktan oluşmaktadır. Karar ağaçları, eğitim ve test sürelerinin hızlı olması, sonuçlarının daha kolay yorumlanabilmesi ve etkin olması sebebiyle sınıflandırmada sıklıkla kullanılan yöntemlerden biridir. Karar ağaçları eğitimli öğrenme için kullanılan çok yaygın bir yöntemdir.

Karar ağacı kullanmanın birçok avantajı vardır; örneğin, bir problemi anlamak ve yorumlamak kolaydır. Alternatif tekniklerin çoğunda, veriler küçük bir işlemden sonra kullanılabilir hale gelir. Karar ağacının ön işleme aşaması, diğer alternatif yöntemlerden daha kısa ve basittir. Bu yöntem hem sayısal hem de sınıf verilerini işlemek için kullanılabilir. Çoğu makine öğrenmesi algoritması, sayısal uygulamalar veya sınıflandırma problemleri için yararlıdır. Karar ağacı öğrenirken her ikisini de kullanabilir. Aynı zamanda düşük bir hesaplama karmaşıklığına sahiptir. Sadeliği ve hızı nedeniyle, kısa sürede çok miktarda veri işlenebilir ve bu yaklaşım veri miktarı arttıkça alternatif yöntemlere göre daha tercih edilebilir hale gelmektedir.

Karar ağacı verimli bir şekilde veri sınıflandırmak için kullanılan bir sınıflandırma algoritmasıdır. Karar ağacı terminal olmayan düğümler (bir kök ve iç düğümleri) ve terminal düğümlerden (yaprak) oluşur. Karar ağacının hazırlanırken başlangıç noktası (kök), vereceğimiz karara, ağaç hazırlanırken her karar düğümünden çıkan dallar ise karar alternatiflerine karşılık gelmektedir.

Karar ağaçlarına dayalı olarak geliştirilen birçok algoritma vardır. Bu algoritmalar birbirlerinden kök, düğüm ve dallanma kriterine göre farklı kategorilere ayrılırlar.

Yaygın olarak bilinen karar ağacı algoritmaları ID3, C4.5 ve C5'dir. C4.5, Karar Ağacı oluşturmak için kullanılan verimli ve popüler olarak kullanılan algoritmalardan biridir [26]. Karar ağacı algoritmasının adımları:

1. T öğrenme kümesi oluşturulur.
2. T kümesindeki örnekleri en iyi ayıran nitelik belirlenir.
3. Seçilen nitelik ile ağacın bir düğümü oluşturulur ve bu düğümden çocuk düğümleri veya ağacın yaprakları oluşturulur. Çocuk düğümlere ait alt veri kümesinin örnekleri belirlenir.
4. Bir önceki adımda yaratılan her alt veri kümesi için:

⇨ Örneklerin hepsi aynı sınıfa aitse

⇨ Örnekleri bölecek nitelik kalmamışsa

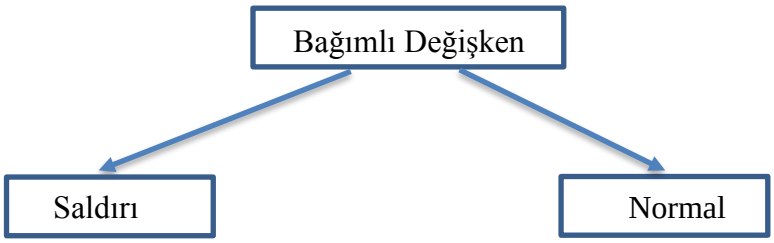
⇨ Kalan niteliklerin değerini taşıyan örnek yoksa işlem sonlandırılır, diğer durumda alt veri kümesini ayırmak için 2. adımdan devam edilir.

1. Adım: Veri seti ile Çizelge 3.3'de verilen öğrenme kümesi oluşturulur.

Çizelge 3.3. Öğrenme kümesi

Veri Seti (Example)	Bağımlı Değişken	Bağımsız Değişken	Karar (Decision)
W1	Saldırı	Bağlantı Uzunluğu (Duration)	Hizmet Engelleme
W2	Saldırı	Ftp oturumundaki komut sayısı	Port Tarama
W3	Saldırı	Son iki saniyede hedef sunucuya yapılan bağlantı sayısı (Count)	Hizmet Engelleme
W4	Saldırı	Tek bağlantıda kaynaktan hedefe aktarılan veri miktarı	Hizmet Engelleme
W5	Normal	Bağlantının normal veya hatalı olma durumu (Flag)	Normal
W6	Saldırı	Kaynak ve Hedef IP adresleri ve port numaraları aynı ise 1 değilse 0 ile ifade edilir. (Land)	Hizmet Engelleme
W7	Saldırı	Yetkili giriş yapabilmek için denenen giriş sayısı (Su_attempted)	Kullanıcı Hesabının Yönetici Hesabına Yükseltilmesi - U2R
W8	Saldırı	Misafir hesabı ile oturum açılma durumu (Is_guest_login)	Yönetici Hesabı ile Yerel Oturum Açma - R2L
W9	Saldırı	Bir bağlantıda "Root" olarak gerçekleştirilen işlem sayısı	Kullanıcı Hesabının Yönetici Hesabına Yükseltilmesi - U2R
W10	Normal	Alarm göstergelerinin sayısı	Normal

2. Adım: Veri setindeki en ayırt edici nitelik belirlenir ve ağacın kökü olarak alınır. Şekil 3.4'de gösterilen şekilde bir ağaç yapısı oluşturulur.



Şekil 3.4. Ağaç yapısı

3. Adım: Ağacın alt veri kümelerine ait yapraklar belirlenir.

3.6.3. Ripper kuralı algoritması

Ripper kuralı, Repeated Incremental Pruning to Produce Error Reduction (RIPPER) algoritmasının WEKA uygulamasıdır [27].

Çeşitli gürültülü veri setlerini işleyebilen verimli bir kural-tabanlı öğrenme algoritmasıdır. Ripper Kural algoritması iki aşamadan oluşmaktadır. İlk aşama kural koşullarını başlatmak içindir. Bir sonraki aşama, bir kural optimizasyon tekniğini kullanır. Kurallar, if-then-else şeklindedir. Bir kural, koşul sağlanıncaya kadar her sınıf tek tek kabul edilir veya veri bir sınıf içinde sınıflandırılır.

Eğer listedeki ilk kural örneği kapsamıyorsa, yani hem kural hem örnekteki nitelikler için eşleşen değerler yoksa o zaman bir sonraki kural denenir. Sırayla örnek bir kural tarafından sınıflandırılana kadar devam eder. Eğer kural hiçbir örneği kapsamıyorsa o zaman karar listesinin en altında varsayılan bir kural işletilir. Yani sınıflandırılamayan tüm örnekler bu sınıfta toplanır [28].

Bir kural oluşturulmadan önce mevcut eğitim örnekleri seti iki alt sete ayrılır. Bunlardan biri gelişen kurallar listesi diğeri de budama listesidir. Kural gelişen kurallar listesindeki örneklerden oluşturulur. Bütün olumsuz örnekler belirlenene kadar kurallar bu listeye ilave edilir [29].

Gelişen kurallar listesinde bir kural geliştirildikten sonra kural listesinin performansını arttırmak için budama (kural silme) yapılır. Bir kuralın budamasında RIPPER yalnızca bu kuraldan oluşan son koşulu göz önünde bulundurur [29].

Ripper algoritmasının avantajlı yönlerini aşağıdaki gibi sıralayabiliriz:

- ✓ Kural Kümesini yorumlamak diğerlerine göre daha kolaydır.
- ✓ Karar Ağacı öğrenmesine kıyasla daha iyi öğrenir.
- ✓ Birinci dereceden mantık gösterilemeyen uygulamalarda kolay uygulanabilir

Ripper kuralı, basit ve bilginin sunumu ve ifadesi kolay olduğu için karar ağaçları gibi popüler algoritmalardandır. RIPPER özellikle gürültülü büyük veri kümelerinde daha etkilidir. Ripper algoritmasında iç ve dış olmak üzere iki çeşit döngü bulunmaktadır. Dış döngü kural tabanına aynı anda bir kural ekler ve iç döngü geçerli kurala aynı anda bir koşul ekler. Bilgi kazanç değeri koşulların kurala eklenmesi ile maksimum olur. Bu işlem olumsuz örnek kalmayana kadar devam eder. Bu algoritmanın zaman karmaşıklığı ($N\log_2 N$) dir [30].

3.6.4. Rasgele orman sınıflandırması algoritması

Günümüzde büyük miktarda veriden oluşan çok sayıda sınıflandırma problemi gözlenmektedir. En sık kullanılan algoritmaların bile verileri tam olarak doğru sınıflandıramaması nedeniyle, büyük miktarda verinin sınıflandırılması ve regresyonu için rastgele orman algoritması oluşturulmuştur.

Rastgele orman algoritması bir makine öğrenme tekniğidir ve veri madenciliği uygulamalarından birisidir.

Veri madenciliği uygulamaları iki gruba ayrılmaktadır:

1. Tanımlayıcı Veri Madenciliği Teknikleri
2. Tahminleyici Veri Madenciliği Teknikleri

Tanımlayıcı veri madenciliği tekniği, verileri tanımlamaya, özetlemeye ve kategorilere ayırmaya odaklanır. Bu teknik denetimsiz makine öğrenme teknikleri kullanılarak uygulanır.

Tahminleyici veri madenciliği teknikleri ise geçmiş verilerin analizine odaklanır ve gelecek için sonuçları tahmin eder. Tahmini model, tahmini değişkenlerin özelliklerinin analizine dayanır. Tahminleyici veri madenciliği teknikleri, denetimli makine öğrenme teknikleri kullanılarak gerçekleştirilir. Veri seti, eğitim seti ve test seti olmak üzere ikiye ayrılmıştır. Modelin eğilmesi, eğitim seti kullanılarak uygulanır ve modelin doğruluğunu belirlemek için test seti kullanılır. Çıktı, hipotez olarak bilinen değişken şeklinde üretilir. Bu hipotez kabul ve reddedilme açısından test edilir.

Rastgele Orman, temel sınıflandırıcı olarak karar ağaçlarını kullanır. Breiman tarafından önerilen "Torbalama" ve Freund ve Schapire tarafından önerilen "Hızlandırma" öğrenme yöntemi için iki ana yöntemdir.

Bu teknikler, verilerin değiştirildiği her sınıflayıcı için farklı eğitim setleri oluşturmak amacıyla verilerin yeniden örneklenmesi için kullanılır. Torbalama tekniğinde, bir sınıflandırıcıyı eğitmek için yer değiştirmeli olarak orijinal eğitim veri setinden birden çok önyüklemeli eğitim veri setleri oluşturulur ve her bir önyüklemeli eğitim veri seti için bir ağaç üretilir. Ardışık gelen ağaçlar bir öncekinden bağımsızdır ve tahmin için en büyük oy baz alınır. Diğer taraftan, Hızlandırma iteratif olarak tekrarlı eğitimi kullanır ve yanlış sınıflandırılmış örnekler bundan sonraki iterasyonda daha önemli olsun diye bunların ağırlıkları artırılır. Hızlandırma genel olarak sınıflandırmanın hem varyansını hem de sapmasını azaltır ve çoğu durumda Torbalama'dan daha doğru sonuç verir. Ancak bazı dezavantajlara sahiptir. Yavaştır, tekrarlı eğitim olabilir ve gürültüye karşı hassastır [31].

Rasgele orman sınıflandırması birden fazla karar ağacını kullanarak daha uyumlu modeller üreterek daha isabetli sınıflandırma yapmaya çalışan bir sınıflandırma modelidir [32].

Rastgele orman, her biri birbirinden bağımsız olarak ve aynı dağılım kullanılarak eğitim verisinden rastgele elde edilmiş bir örnekleme dayanan karar ağaçlarından oluşan bir topluluktur. Bu yöntem eğitim sırasında birçok karar ağacı oluşturur ve daha sonra kestirim sırasında bu karar ağaçlarının sınıflandırma sonuçlarından yararlanılarak, girdinin sınıfına çoğunluk oyu aracılığıyla karar verilir. Rastgele ormanın en önemli avantajı, karar ağaçlarındaki aşırı uyum sorununa bir çözüm getirmiş olmasıdır [14].

Rasgele orman algoritmasının avantajlarını şu şekilde özetleyebiliriz:

- ✓ Doğruluk oranları yüksektir.
- ✓ Gürültüye karşı dayanıklıdır.
- ✓ Torbalama ve Hızlandırma işlemlerinden daha hızlıdır.
- ✓ Basittir.

4. ARAŞTIRMA BULGULARI

KDD99 veri kümesinde her veri 41 özelliğe sahiptir. Bu tez çalışmasında eğitim veri kümesi ve test veri kümesi olarak farklı veri setleri kullanılmıştır. Tez çalışmasında kullanılan eğitim veri seti Çizelge 4.1'de verilmektedir.

Çizelge 4.1. Eğitim veri seti

Normal Kayıt Sayısı	Hizmet Engelleme Saldırılarından Oluşan Kayıt Sayısı	Bilgi Tarama Saldırılarından Oluşan Kayıt Sayısı
5000	Back: 250	Satan: 1000
	Teardrop: 100	Ipsweep: 750
	Neptune: 600	Nmap: 0
	Land: 20	Portsweep:750
	Smurf: 1400	
	Pod: 130	

Yapılan çalışmada KDD99 veri setinden eğitim için 10.000 kayıt alınmıştır. Bunların 5.000'i normal kayıt, 5.000'i ise saldırı verilerinden oluşan kayıtlardır. Saldırı verilerinin 2500 tanesi Hizmet Engelleme (DoS) saldırılarından, 2500 tanesi ise Ağ Tarama (Probe) saldırılarından oluşmaktadır.

Bu çalışmada kullanılan test veri seti Çizelge 4.2'de verilmiştir.

Çizelge 4.2. Test veri seti

Normal Kayıt Sayısı	Hizmet Engelleme Saldırılarından Oluşan Kayıt Sayısı	Bilgi Tarama Saldırılarından Oluşan Kayıt Sayısı
2500	Back: 200	Satan: 375
	Teardrop: 80	Ipsweep: 475
	Neptune: 70	Nmap: 0
	Land: 10	Portsweep: 400
	Smurf: 865	
	Pod: 25	

KDD99 veri setinden seçilen test veri setlerinin her biri ise 5.000 kayıttan oluşmaktadır. Bunların 2.500 tanesi saldırı kaydı ve 2.500 tanesi ise normal kayıttan oluşmaktadır. Her veri kümesi içindeki saldırı kayıtları Hizmet Reddi ve Bilgi Tarama saldırı çeşitlerini içermektedir. Diğer saldırı türleri veri setinden silinmiştir.

KDD 99 veri setinde yer alan ve bağlantı kayıtları için tanımlanmış 41 tane özellik (parametre) mevcuttur. KDD 99 veri setinde yer alan her bir veri için tanımlanan 9 temel özellik Çizelge 4.3'de, içerik özellikleri ve sunucu bazlı bağlantı özellikleri için tanımlanmış 13 özellik Çizelge 4.4, her ağ bağlantısı için zaman bazlı trafik özellikleri için tanımlanmış 9 özellik Çizelge 4.5 ve her ağ bağlantısı için host bazlı trafik özellikleri olarak tanımlanmış 10 özellik Çizelge 4.6'da verilmiştir.

Çizelge 4.3. Ağ bağlantılarında yer alan temel özellikler

No	Özellik Adı	Türü	Açıklama
1	Duration	Sürekli (continuous)	Süre: Bağlantının uzunluğu, süresi (saniye sayısı)
2	Protocol_type	Ayrık (discrete)	Protokol türü: Bağlantı için kullanılan protokol türü TCP ise 1, ICMP ise 2, UDP ise 3 değerini alır.
3	Service	Ayrık (discrete)	Servis tipi: Kullanılan hedef ağ servisi örneğin; http, telnet gibi
4	Flag	Ayrık (discrete)	Bayrak bitleri: Bağlantının normal veya hatalı olma durumu
5	Src_bytes	Sürekli (continuous)	Kaynak bayt: Tek bağlantıda kaynaktan hedefe aktarılan veri miktarı veya veri baytı sayısı
6	Dst_bytes	Sürekli (continuous)	Hedef bayt: Tek bağlantıda hedeften kaynağa aktarılan veri miktarı veya veri baytı sayısı
7	Land	Ayrık (discrete)	Kaynak ve hedef IP adresleri ve port numaraları aynı ise 1 değilse 0 ile ifade edilir.
8	Wrong_fragment	Sürekli (continuous)	Bağlantıdaki yanlış bölümlenme/parçalanma sayısı
9	Urgent	Sürekli (continuous)	Bağlantıdaki acil paket sayısı

Çizelge 4.4. Her ağ bağlantısında yer alan etki alanı bilgisindeki içerik özellikleri ve sunucu bazlı bağlantı özellikleri

No	Özellik Adı	Türü	Açıklama
10	Hot	Sürekli (continuous)	"Alarm/İkaz" göstergelerin sayısı; Örneğin; bir sistem dizinine giriş yapmak, program oluşturmak veya program çalıştırmak gibi durumlar
11	Num_failed_logins	Sürekli (continuous)	Başarısız olan giriş sayısı veya başarısız oturum açma girişimi sayısı
12	Logged_in	Ayrık (discrete)	Başarılı giriş sağlanmış ise 1, başarısız giriş yapılmış ise 0 değerini alır.
13	Num_compromised	Sürekli (continuous)	Gizliliğin ihlal edildiği tehlikeli durum sayısı
14	Root_shell	Ayrık (discrete)	Kök kabuğu “Root Shell” elde edildiyse 1 değilse 0
15	Su_attempted	Ayrık (discrete)	Yetkili giriş yapabilmek için denenen giriş sayısıdır. “Su Root” komutu girildiyse veya denendiyse 1 değilse 0 değerini alır.
16	Num_root	Sürekli (continuous)	“Root” erişim sayısı veya bir bağlantıda “Root” olarak gerçekleştirilen işlem sayısı
17	Num_file_creations	Sürekli (continuous)	Bağlantıdaki dosya oluşturma/dosya yaratma işlemlerinin sayısı
18	Num_shells	Sürekli (continuous)	Shell (kabuk) istemlerinin sayısı
19	Num_access_files	Sürekli (continuous)	Erişim kontrolü dosyalarındaki işlem sayısı
20	Num_outbound_cmds	Sürekli (continuous)	Bir ftp oturumunda giden komut sayısı
21	Is_hot_login	Ayrık (discrete)	Eğer oturum açma işlemi “hot” listesindeyse yani “root” veya “admin” hesabı ile oturum açılmış ise 1 değilse 0 değerini alır.
22	Is_guest_login	Ayrık (discrete)	Eğer misafir “guest” hesabı ile oturum açılmış ise 1 değilse 0 değerini alır.

Çizelge 4.5. Her ağ bağlantısı için zaman bazlı trafik özellikleri - aynı sunucu veya aynı servise iki saniyelik bir zaman birimi kullanılarak yapılan bağlantının trafik özellikleri

No	Özellik Adı	Türü	Açıklama
23	Count	Sürekli (continuous)	Son iki saniyedeki aynı hedef sunucuya yapılan bağlantıların sayısı
24	Srv_count	Sürekli (continuous)	Son iki saniye içinde aynı bağlantı üzerindeki aynı servise veya port numarasına yapılan bağlantı sayısı
25	Serror_rate	Sürekli (continuous)	Aynı sunucu bağlantılarındaki "SYN" hatası olan bağlantıların yüzdesi; s0, s1, s2 veya s3 bayrağını etkinleştiren bağlantıların yüzdesi
26	Srv_serror_rate	Sürekli (continuous)	Aynı servis bağlantıları için "SYN" hatası olan bağlantıların yani s0, s1, s2 veya s3 bayrağını etkinleştiren bağlantıların yüzdesi
27	Rerror_rate	Sürekli (continuous)	"REJ" bayrağını etkinleştiren bağlantıların yüzdesi
28	Srv_rerror_rate	Sürekli (continuous)	Aynı servis bağlantıları için "REJ" bayrağını etkinleştiren bağlantıların yüzdesi
29	Same_srv_rate	Sürekli (continuous)	Aynı servise yapılan bağlantıların yüzdesi
30	Diff_srv_rate	Sürekli (continuous)	Farklı servislere yapılan bağlantıların yüzdesi
31	Srv_diff_host_rate	Sürekli (continuous)	Aynı servis bağlantıları için farklı hedef makinelere yapılan bağlantıların yüzdesi

Çizelge 4.6. Her ağ bağlantısı için host bazlı trafik özellikleri

No	Özellik Adı	Türü	Açıklama
32	Dst_host_count	Sürekli (continuous)	Aynı IP adresine sahip hedef makine için yapılan bağlantı sayısı
33	Dst_host_srv_count	Sürekli (continuous)	Hedef sunucu için son iki saniye içinde geçerli servisle aynı port numarasına yapılan bağlantı sayısı
34	Dst_host_same_srv_rate	Sürekli (continuous)	Hedef sunucular için aynı servise yapılan bağlantıların yüzdesi,
35	Dst_host_diff_srv_rate	Sürekli (continuous)	Hedef sunucular için farklı servislere yapılan bağlantıların yüzdesi
36	Dst_host_same_src_port_rate	Sürekli (continuous)	Hedef ana bilgisayarlar için aynı kaynak portu ile yapılan bağlantıların yüzdesi
37	Dst_host_srv_diff_host_rate	Sürekli (continuous)	Aynı servis bağlantıları için farklı hedef makinelere yapılan bağlantıların yüzdesi
38	Dst_host_serror_rate	Sürekli (continuous)	Hedef sunucular için SYN hatası olan S0, s1, s2 veya s3 bayrağını etkinleştiren bağlantıların yüzdesi

Çizelge 4.6. (devam) Her ağ bağlantısı için host bazlı trafik özellikleri

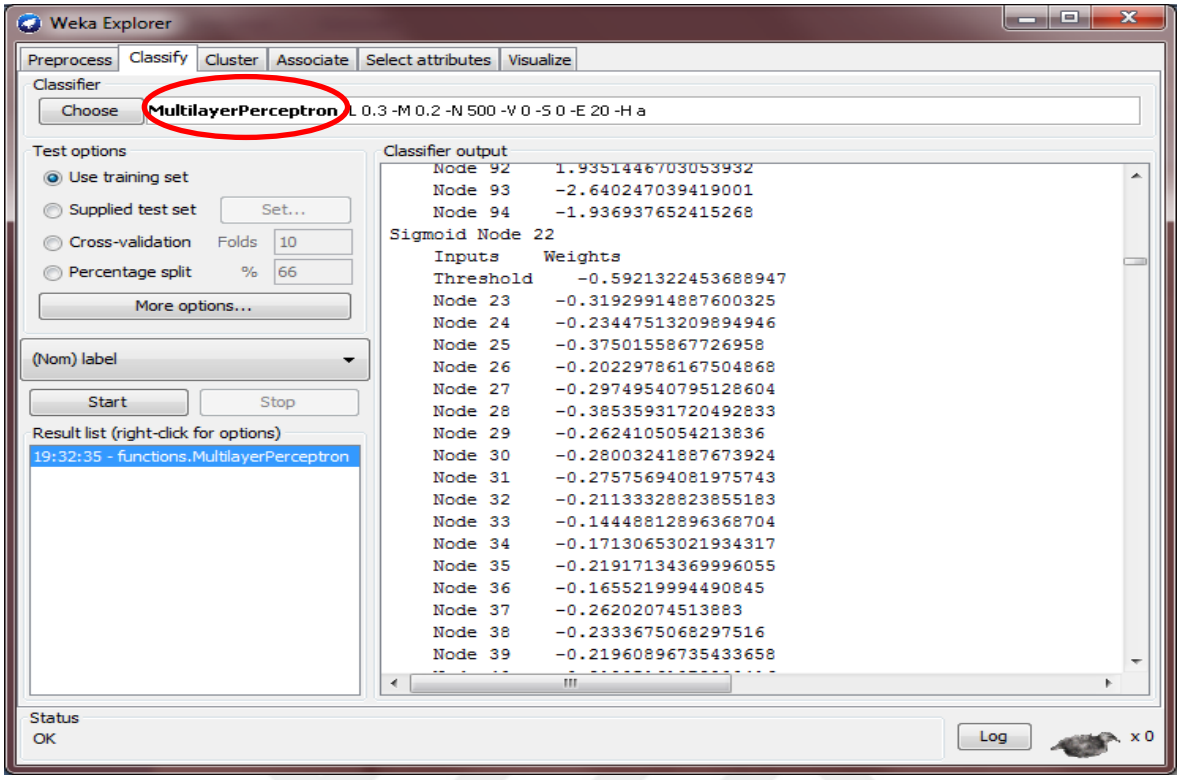
39	Dst_host_srv_serror_rate	Sürekli (continuous)	Hedef makinelerde aynı servis bağlantıları için "SYN" hatası olan s0, s1, s2 veya s3 bayrağını etkinleştiren bağlantıların yüzdesi
40	Dst_host_rerror_rate	Sürekli (continuous)	Hedef sunucular için REJ bayrağını etkinleştiren bağlantıların yüzdesi
41	Dst_host_srv_rerror_rate	Sürekli (continuous)	Hedef sunucularda aynı servis bağlantıları için "REJ" bayrağını etkinleştiren bağlantıların yüzdesi

Çizelgelerde kullanılan bazı ifadelerin açıklamaları aşağıdaki gibidir:

S0 **:** Bağlantı girişimi görüldü, cevap yok.
S1 : Bağlantı kuruldu, sonlandırılmadı.
SF : Normal kuruluş ve sonlandırma.
REJ : Bağlantı girişimi reddedildi.
RSTO **:** Bağlantı kuruldu, oluşturucu iptal edildi (bir RST gönderdi).
RSTR : Bağlantı kuruldu, cevaplayıcı iptal edildi.

4.1. Geri Yayılma (Back-Propagation) Algoritması İle Saldırı Tespiti

Çalışmamızda ilk olarak Geri Yayılma (Back-Propagation) sınıflandırma tekniği kullanılmıştır. WEKA aracında bu sınıflandırma Multilayer Perceptron sınıflandırma olarak geçmektedir. Şekil 4.1'de Geri Yayılma Algoritması'nın Weka üzerinde çalışması gösterilmektedir.



Şekil 4.1. Geri yayılma algoritması giriş katmanı

Eğitim veri kümesi eğitildikten sonra bir model olarak kaydedilir ve bu model Weka aracına yüklendikten sonra Test veri seti seçilerek test edilmeye başlanır. Test veri setinin oluşturulması önemlidir. Test veri seti, Eğitim veri setinden ayrı olmalıdır. Eğitim veri setinden seçilen veriler ile test yapılırsa %100'e çok yakın belki de %100 performans sağlanır. Çünkü bu şekilde gerçekleşen öğrenme çıkarımlardan oluşmaz. Bu nedenle KDD99 veri setinden rastgele alınan 2500 normal 1250 Hizmet Engelleme ve 1250 Bilgi Tarama saldırı verisi test için seçilmiştir.

Geri yayılma algoritması ile elde edilen sonuçlar Çizelge 4.7'de gösterilmektedir.

Çizelge 4.7. Geri yayılma algoritması ile elde edilen sonuçlar

Sınıflandırma Türü	TAO (%)	NAO (%)	STO (%)	Test Süresi (s)
Geri Yayılma	99,88	100	99,76	927.08

Toplam Algılama Oranı (TAO), Gerçek Zamanlı Saldırı Tespit Sisteminde doğru algılanabilen Hizmet Engelleme saldırıları, Bilgi Tarama saldırıları ve Normal ağ verilerinin yüzdesidir.

Normal Algılama Oranı (NAO), Gerçek Zamanlı Saldırı Tespit Sisteminde doğru algılanabilen Normal sınıf yüzdesidir.

Saldırı Tespit Oranı (STO), Gerçek Zamanlı Saldırı Tespit Sisteminde doğru algılanabilen tüm Saldırı sınıflarının yüzdesidir.

Geri Yayılma Algoritması çok katmanlı olduğu için eğitim süresi ve test süresi çok yüksektir. Algoritma normal verilerin tamamını doğru bulurken saldırı verilerinin bazılarını yanlış sınıflandırmıştır. Saldırı verisi olmasına rağmen 3 adet veri normal olarak gösterilmiştir. Test verilerinde kullanılan saldırı türleri ile eğitim verisinde kullanılan saldırı verileri aynı türdendir. Eğer test verisinde kullanılan verilerde saldırı türleri eğitim verisinden farklılık gösterirse alınan tespit oranlarında düşme yaşanır.

4.2. Karar Ağacı Algoritması İle Saldırı Tespiti

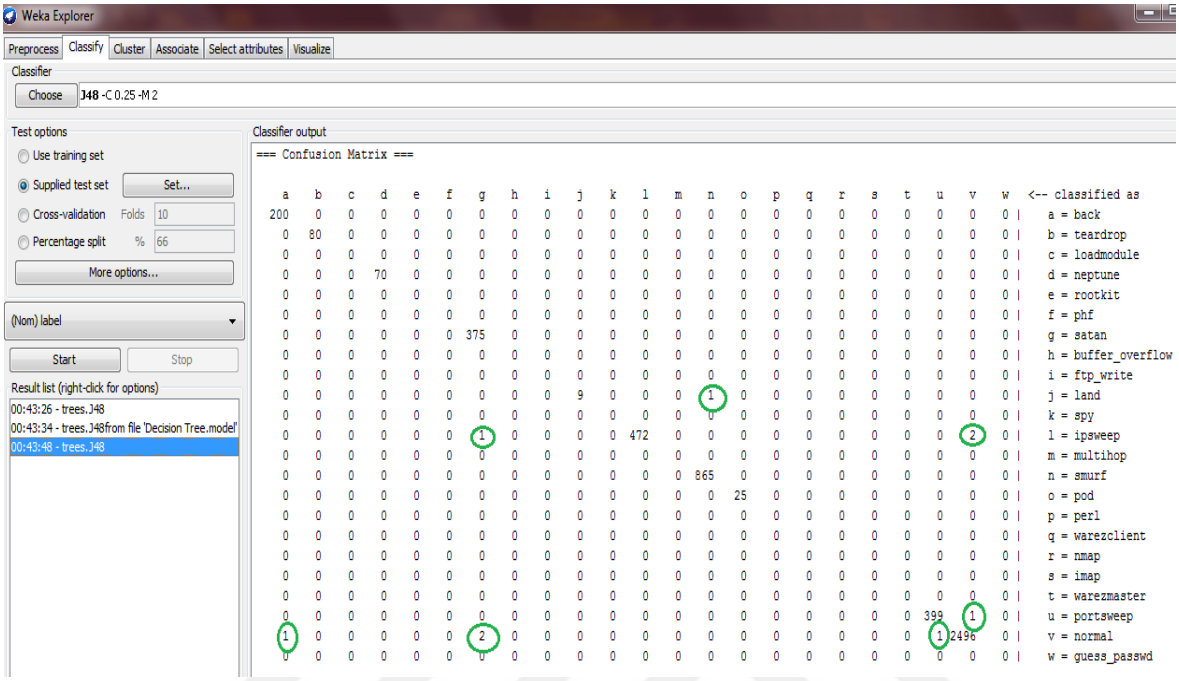
Weka aracında karar ağacı ile eğitim süresinin hesaplanması için J48 sınıflandırıcı seçilir. Karar ağaçlarının performansı yani eğitim süresi Geri Yayılma Algoritmasına göre çok daha iyidir. Karar Ağacı Algoritması tekniği kullanılan ve saldırı türleri aynı olan veri setlerinden elde edilen sonuçlar Çizelge 4.8'de gösterilmiştir.

Çizelge 4.8. Test veri seti ile eğitim veri setindeki saldırı türleri aynı olduğunda karar ağacı algoritması ile elde edilen sonuçlar

Sınıflandırma Türü	TAO (%)	NAO (%)	STO (%)	Test Süresi (s)
Karar Ağacı	99.82	99.84	99.80	0.71

Çalışmada kullanılan eğitim veri seti ile test veri setindeki saldırı türleri aynı olduğunda Saldırı Tespit Oranı daha iyi bir performans sergilemekte, ancak hiç karşılaşılmayan ve eğitilmeyen bir saldırı türü test veri setinde yer aldığında daha düşük bir Saldırı Tespit Oranı

elde edilmektedir. Şekil 4.2'de test veri seti ile eğitim veri setindeki saldırı türleri aynı olduğunda oluşan hata matrisi yer almaktadır.



Şekil 4.2. Test veri seti ile eğitim veri setindeki saldırı türleri aynı olduğunda hata matrisi

Eğitim veri kümesinde bilgi tarama saldırılarından olan Nmap yoktur. Ancak test veri kümesine Nmap taraması dahil edildiğinde algılama oranlarında farklılık oluşmakta ve saldırı tespit oranı düşmektedir. Çizelge 4.9'da test veri seti ile eğitim veri seti farklı olduğunda eğitim veri setine ait bilgiler Çizelge 4.10'da ise test veri setine ait bilgiler yer almaktadır.

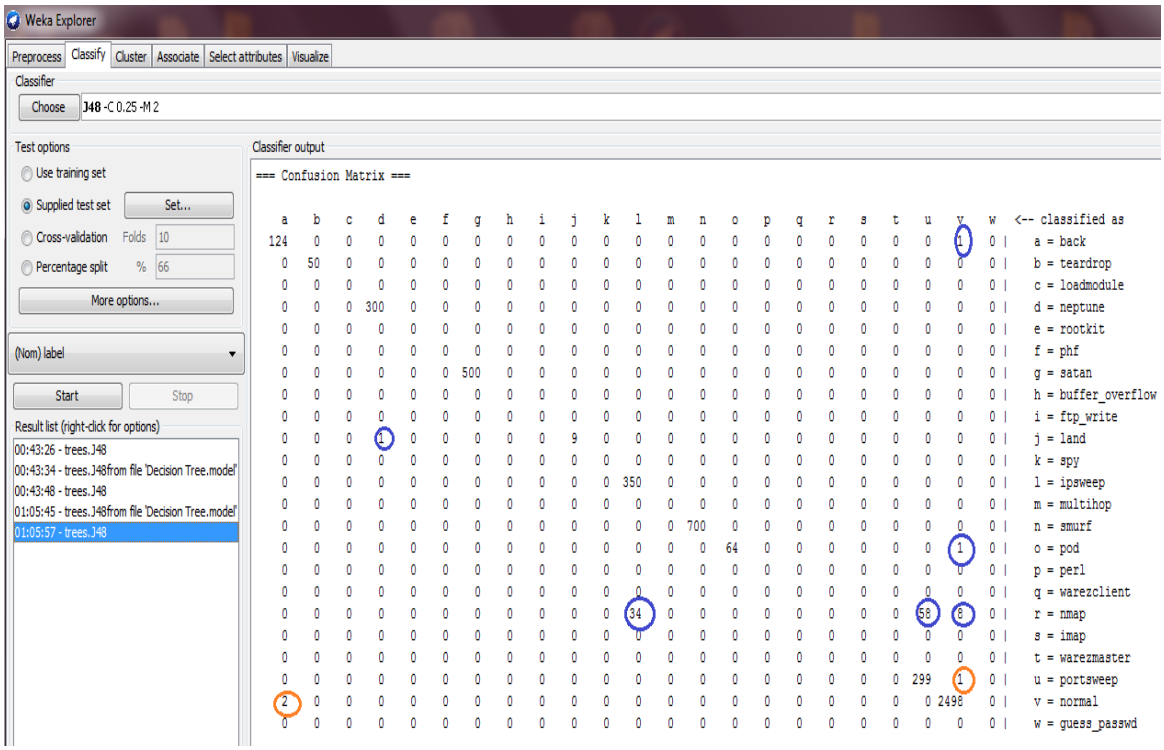
Çizelge 4.9. Test veri seti ile eğitim veri setindeki saldırı türleri farklı olduğunda eğitim veri seti

Normal Kayıt Sayısı	Hizmet Engelleme Saldırılarından Oluşan Kayıt Sayısı	Bilgi Tarama Saldırılarından Oluşan Kayıt Sayısı
5000	Back: 250	Satan: 1000
	Teardrop: 100	Ipsweep: 750
	Neptune: 600	Nmap: 0
	Land: 20	Portsweep:750
	Smurf: 1400	
	Pod: 130	

Çizelge 4.10. Test veri seti ile eğitim veri setindeki saldırı türleri farklı olduğunda test veri seti

Normal Kayıt Sayısı	Hizmet Engelleme Saldırılarından Oluşan Kayıt Sayısı	Bilgi Tarama Saldırılarından Oluşan Kayıt Sayısı
2500	Back: 125	Satan: 500
	Teardrop: 50	Ipsweep: 350
	Neptune: 300	Nmap: 100
	Land: 10	Portsweep: 300
	Smurf: 700	
	Pod: 65	

Veri kümesi Nmap ile eğitilmediği için Nmap verilerinin büyük bir bölümü diğer bilgi tarama saldırılarına benzetilmiştir. 8 tanesi ise normal veri olarak algılanmıştır. Saldırı verilerini algılamada bu nedenle azalma gerçekleşmiştir. Hatalı algılamalar Şekil 4.3'de yer alan ekran görüntüsünde yer almaktadır.



Şekil 4.3. Test veri seti ile eğitim veri setindeki saldırı türleri farklı olduğunda hata matrisi

Eğitim veri setinde olmayıp, test veri setinde yer alan bir saldırı türü ile çalışma tekrarlandığında elde edilen sonuçlar Çizelge 4.11'de gösterilmektedir.

Çizelge 4.11. Test veri seti ile eğitim veri setindeki saldırı türleri farklı olduğunda elde edilen sonuçlar

Sınıflandırma Türü	TAO (%)	NAO (%)	STO (%)	Eğitim süresi (s)
Karar Ağacı	97.88	99.92	95.84	0.41

4.3. Ripper Kuralı İle Saldırı Tespiti

Weka aracında Ripper Kuralı ile eğitim süresinin hesaplanması için JRip sınıflandırıcı seçilir. Ripper Kuralının performansı yani eğitim süresi Geri Yayılma Algoritmasına göre çok daha iyidir. Ancak Karar Ağaçlarına göre daha iyi bir performans sergileyememiştir. Ripper Kuralı ile elde edilen sonuçlar Çizelge 4.12'de yer almaktadır.

Çizelge 4.12. Ripper kuralı ile elde edilen sonuçlar

Sınıflandırma Türü	TAO (%)	NAO (%)	STO (%)	Test Süresi (s)
Ripper Kuralı	99.74	99.80	99.68	3.46

4.4. Rasgele Orman Algoritması İle Saldırı Tespiti

Weka aracında Rasgele orman algoritması ile eğitim süresinin hesaplanması için RandomForest sınıflandırıcı seçilir. Rasgele orman algoritmasının algılama yüzdeleri diğer sınıflandırıcılara göre daha iyi bir performans sergilemiştir. Rasgele orman algoritması kullanılarak elde edilen sonuçlar Çizelge 4.13'de yer almaktadır.

Çizelge 4.13. Rasgele orman algoritması ile elde edilen sonuçlar

Sınıflandırma Türü	TAO (%)	NAO (%)	STO (%)	Test Süresi (s)
Rasgele Orman	99.94	100	99.88	2.28

4.5. Değerlendirme

Yapılan çalışmada Hizmet engelleme saldırıları ve Bilgi tarama saldırılarının sınıflandırılmasında 6 çeşit Hizmet engelleme (DoS), 4 çeşit Bilgi tarama (Probe) saldırısı mevcuttur. Sınıflandırmanın doğru yapılması yalnızca Hizmet engelleme ve Bilgi tarama şeklinde değil saldırının türü şeklindedir. Örneğin Smurf saldırısı da Teardrop saldırıları da Hizmet engelleme saldırısıdır. Ancak bazı Smurf saldırıları Teardrop olarak algılanabilmektedir. Bu durumda yanlış sınıflandırma yapılmıştır. Deneysel çalışmamızda her bir saldırı çeşidini ayrı ayrı sınıflandırıp bu kapsamda hesaplamalar yapılmıştır. Her bir algoritma ile elde edilen sonuçlar Çizelge 4.14'de verilmektedir.

Çizelge 4.14. Farklı algoritmalar kullanılarak elde edilen sonuçlar

Sınıflandırma Türü	TAO (%)	NAO (%)	STO (%)	Test Süresi (s)
Geri Yayılma	99.88	100	99.76	927,08
Karar Ağacı	99.82	99.84	99.80	0.71
Ripper Kuralı	99.74	99.80	99.68	3.46
Rasgele Orman	99.94	100	99.88	2,28

4.5.1. Literatürde elde edilen sonuçlar ve bizim elde ettiklerimiz

Yapılan araştırmalar sonucunda birçok farklı teknikler kullanılarak yapılan yapay zekâ tabanlı saldırı tespiti ile ilgili birçok başarılı çalışma elde edilmiştir. Birbirinden farklı algoritmalar kullanılmış, en tanımlayıcı özellikler belirlenmiş ve farklı algoritmalar bir arada kullanılmıştır.

Hanifi ve arkadaşlarının farklı algoritmalar ile elde ettikleri performans yüzdeleri Karar Ağacı ile yüzde 81,05 oranında, Naive Bayes ile yüzde 76,56 oranında, NB Ağacı ile yüzde 82,02 oranında, Rasgele Orman ile yüzde 80,67 oranında, Çok Katmanlı Algılayıcı ile yüzde 77,41 oranında kendi çalışmaları ile de yüzde 80,119 oranında olmuştur.

Hasan ve arkadaşlarının Random Forest algoritması ile gerçekleştirdikleri ve 41 özellik yerine 25 özellik kullanarak gerçekleştirdikleri sistemde saldırı tespit oranı yüzde 91,90 olmuştur. 41 özellikli yaklaşım ile Random Forest algoritmasında 10,62 saniyelik zaman diliminde yüzde 91,41 oranında başarı sağlarken özellik sayısını azaltarak 7,98 saniyede yüzde 91,90 oranında başarı elde etmişlerdir.

Özgür ve Erdem 41 nitelik için %79 başarı elde ederken, Genetik Algoritma ile %91 oranında başarı elde etmişlerdir.

Phyu Thi Htun ve Kyaw Thet Khaing Rasgele Orman algoritması ile k-en yakın komşu algoritmasının birlikte kullanılması sonucunda %99.97 oranında doğruluk elde etmişlerdir. Ancak yalnızca Hizmet engelleme saldırılarını tespit edebilmişlerdir.

Sağıroğlu, Yolaçan ve Yavanoğlu'nun farklı eğitim ve test kümeleriyle yaptıkları testlerden elde ettikleri en iyi sonucun %97,92 başarılı olduğunu, en düşük sonucun ise %81,93 olduğunu ifade etmişlerdir.

Aggarwala ve Sharma çalışmalarında KDD veri setinin 41 niteliğini kategorize ederek Temel, İçerik, Trafik ve Host olarak dört gruba ayırmışlardır. Bu gruplara göre elde ettikleri başarı oranları Temel özelliklere göre yüzde 80,78 oranında, İçerik özeklerine göre yüzde 79,42 oranında, Trafik özelliklerine göre yüzde 78,59 oranında, Host özelliklerine göre ise yüzde 76,54 oranında olmuştur.

Sabhnani ve Serpen çalışmalarında Yönetici Hesabı ile Yerel Oturum Açma (Remote to Local Attack – R2L) saldırılarını tespit edebilmektedir. Warezmaster saldırısının tespit edilme başarısı yüzde 65,6597 oranında olmuştur. Warezclient saldırılarının tespit oranı ise yüzde 30,2352 olmuştur.

Bu tez çalışmasında elde edilen en iyi sonuç ise Rasgele Orman Algoritması ile yüzde 99,94 oranında olmuştur. Diğer çalışmalarda yalnızca verinin saldırı verisi mi yoksa normal veri mi olduğu tespit edilmiştir. Phyu Thi Htun - Kyaw Thet Khaing ve Sabhnani - Serpen ise çalışmalarında sadece tek bir saldırı verisinin sistem üzerinde eğitilip test edilmesi ile ezberletilmiş bir sistem tasarımı üzerinde çalışma tasarlamışlardır.

Biz de çalışmamızda hem saldırı verilerini ve normal verileri tespit ettik. Hem de bulunan saldırı verileri için Hizmet Engelleme ve Bilgi Tarama şeklinde sınıflandırmada bulunduk ve bu sınıf içerisinde yer alan saldırı çeşitlerine saldırının belirlenmesini sağladık.

Bizim tez çalışmamız ile yapılan araştırmalar sonucu elde edilen başarı oranlarının karşılaştırılması Çizelge 4.15'de gösterilmiştir.

Çizelge 4.15. Karşılaştırma

Çalışmalar	En İyi Başarı Yüzdesi	Tespit Edilen Veriler
Hanifi ve arkadaşları	80,119	Normal – Saldırı
Hasan ve arkadaşlarının	91,90	Normal – Saldırı
Özgür ve Erdem	91	Normal – Saldırı
Phyu Thi Htun ve Kyaw Thet Khaing	99,97	Hizmet Reddi Saldırısı
Sağıroğlu ve arkadaşları	97,92	Normal – Saldırı
Aggarwala ve Sharma	80,78	Özellik Bazlı
Sabhnani ve Serpen	65,6597	Yönetici Hesabı ile Yerel Oturum Açma Saldırıları
Bizim tez çalışmamızda	99,94	Normal – Saldırı Saldırı verisinin Hizmet Reddi ve Bilgi Tarama türlerine göre sınıflandırılması

5. SONUÇ VE ÖNERİLER

Weka uygulamasında yer alan çeşitli makine öğrenme algoritmaları denenerek en iyi sonuç elde edilmeye çalışılmıştır. Sonuç olarak büyük veri setleri için Geri Yayılma Algoritmasının kullanımı eğitim süresinin uzun sürmesinden dolayı uygun değildir. Bizim çalışmamızda en iyi tespiti Rasgele Orman Algoritması yapmıştır. Buna en yakın değerler Geri Yayılma Algoritması ile elde edilmiştir. Ancak buna rağmen eğitim süresinin çok uzun sürmesi zaman açısından maliyeti büyük ölçüde arttırmıştır.

Bir çıkarım yapmak gerekirse büyük veri setleri için Rasgele Orman Algoritması eğitim süresinin çok iyi olmasından ötürü ve iyi bir tespit, öğrenme yüzdesi vermesi nedeniyle kullanılabilir. Büyük veri setlerinde Geri Yayılma Algoritması kullanımı uygun değildir. Küçük veri setleri için kullanılabilir.

Yapılan çalışmada hem veriler saldırı verisi ve normal veri olarak ayrılmakta hem de Hizmet engelleme saldırıları ve Bilgi tarama saldırıları çeşitlerine göre kategorize edilmektedir. Sınıflandırmanın doğru yapılması yalnızca Hizmet engelleme ve Bilgi tarama şeklinde değil saldırının türü şeklindedir. Deneysel çalışma her bir saldırı çeşidini ayrı ayrı sınıflandırıp bu kapsamda hesaplamalar yapılmıştır. Literatürde yapılan çalışmalar verileri sadece saldırı verisi veya normal veri olarak ayırt etmiş veya sadece tek bir saldırı türü üzerinden tespit yapmıştır. Buna rağmen çok iyi başarı oranları elde edilememiştir.

En iyi başarı oranını Phyu Thi Htun ve Kyaw Thet Khaing 99.97 ile elde etmiş ancak onlarda çalışmalarında yalnızca Hizmet engelleme saldırılarını kullanmışlardır. Diğer saldırı türleri ve normal veriler üzerinde herhangi bir tespitte bulunamamışlardır.

Yapılan çalışma tüm saldırı türlerinin tespit edilmesi ile geliştirilebilir veya yeni bir algoritma tasarlanarak saldırı tespiti gerçekleştirilebilir.

KAYNAKLAR

1. İnternet: Siber Güvenliğin Sağlanması: Türkiye'deki Mevcut Durum ve Alınması Gereken Tedbirler. URL: https://www.btk.gov.tr/uploads/undefined/sg.pdf, Son Erişim Tarihi: 12.03.2019

2. İnternet: Kritik Altyapılara Siber Saldırı. URL: http://ylt44.com/bilimsel/siber.html, Son Erişim Tarihi: 29.04.2019.

3. Şenkaya, Y., Adar, U. G. (2014). *Siber Savunmada Yapay Zeka Sistemleri Üzerine İnceleme.* Akademik Bilişim'14 - XVI. Akademik Bilişim Konferansı Bildirileri, 5 - 7 Şubat 2014 Mersin Üniversitesi.

4. Sağıroğlu, Ş., Yolaçan, E. N., Yavanoğlu, U. (2011). Zeki Saldırı Tespit Sistemi Tasarımı ve Gerçekleştirilmesi. *Gazi Üniversitesi Mühendislik –Mimarlık Fakültesi Dergisi*, 26(2), 325-340.

5. KarsligEl, M. E., Yavuz, A. G., Güvensan, M. A., Hanifi, K. ve Bank, H. (2017, May). *Network intrusion detection using machine learning anomaly detection algorithms*. In 2017 25th Signal Processing and Communications Applications Conference.

6. Hasan, M. A. M., Nasser, M., Ahmad, S. ve Molla, K. I. (2016). Feature selection for intrusion detection using random forest. *Journal of information security*, 7(03), 129.

7. Özgür, A. ve Erdem, H. (2018). Saldırı tespit sistemlerinde genetik algoritma kullanarak nitelik seçimi ve çoklu sınıflandırıcı füzyonu. *Gazi Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi,* 33(1).

8. Htun, P. T. ve Khaing, K. T. (2013). Detection model for daniel-of-service attacks using random forest and k-nearest neighbors. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, 2(5), 1855-1860.

9. Aggarwal, P. ve Sharma, S. K. (2015). Analysis of KDD dataset attributes-class wise for intrusion detection. *Procedia Computer Science*, 57, 842-851.

10. Sabhnani, M. ve Serpen, G. (2003, June). KDD Feature Set Complaint Heuristic Rules for R2L Attack Detection. *In Security and Management* (310-316).

11. İnternet: T.C. Ulaştırma, Denizcilik ve Haberleşme Bakanlığı, 2016 - 2019 Ulusal Siber Güvenlik Stratejisi. URL: https://www.uab.gov.tr/uploads/pages/siber-guvenlik/2016-2019guvenlik.pdf, Son Erişim Tarihi: 05.05.2019.

12. İnternet: Veri Madenciliği ile Saldırı Tespiti. URL: https://verimadencisi.blogspot.com/2014/06/veri-madenciligi-ile-saldr-tespiti.html, Son Erişim Tarihi: 16.02.2019.

13. İnternet: Makine Öğrenmesi SEO'yu Nasıl Değiştiriyor. URL: https://www.stradiji.com/makine-ogrenmesi-seoyu-nasil-degistiriyor/, Son Erişim Tarihi: 27.05.2019.

14. Kalaycı, T. E. (2018). Kimlik hırsızı web sitelerinin sınıflandırılması için makine öğrenmesi yöntemlerinin karşılaştırılması. *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi,* 24(5), 870-878.

15. İnternet: Saldırı Tespit ve Önleme Sistemi. URL: https://www.beyaz.net/tr/guvenlik/cozumler/saldiri_tespit_ve_onleme_sistemi.html, Son Erişim Tarihi: 21.01.2019.

16. Yıldırım, M. Z., Çavuşoğlu, A., Şen, B. ve Budak, İ. (2014). *Yapay Sinir Ağları ile Ağ Üzerinde Saldırı Tespiti ve Paralel Optimizasyonu.* Akademik Bilişim'14 - XVI. Akademik Bilişim Konferansı Bildirileri, 5 - 7 Şubat 2014 Mersin Üniversitesi, 671-677.

17. İnternet: KDD99 Veri Seti. URL: http://kdd.ics.uci.edu/databases/kddcup99/kddcup99.html, Son Erişim Tarihi: 13.04.2019.

18. İnternet: Bilgi Güvenliği ve Genel Güvenlik Kavramları. URL: https://www.cozumpark.com/bilgi-g-venli-i-ve-genel-g-venlik-kavramlar-b-l-m-2/, Son Erişim Tarihi: 16.05.2019.

19. İnternet: Hizmet Dışı Bırakma Saldırıları ve Korunma Yöntemleri. URL: http://bidb.itu.edu.tr/seyir-defteri/blog/2013/09/07/denial-of-service-(dos)saldırıları-ve-korunma-yöntemleri, Son Erişim Tarihi: 12.03.2019.

20. İnternet: Hizmet Dışı Bırakma Saldırıları ve Çeşitleri. URL: https://www.slideshare.net/bgasecurity/2-do-sddossaldirilarivecesitleri, Son Erişim Tarihi: 18.03.2019.

21. Kaya, Ç. ve Yıldız, O. (2014). Makine öğrenmesi teknikleriyle saldırı tespiti: Karşılaştırmalı analiz. *Marmara Fen Bilimleri Dergisi,* 3, 89-104.

22. Coşkun, C. ve Baykal, A. (2011). *Veri madenciliğinde sınıflandırma algoritmalarının bir örnek üzerinde karşılaştırılması.* Akademik Bilişim, 2011, XIII. Akademik Bilişim Konferansı Bildirileri, 2-4 Şubat 2011 İnönü Üniversitesi, Malatya, 51-58.

23. Keleşoğlu, Ö. ve Fırat, A. (2006). Tuğla Duvardaki ve Tesisattaki Isı Kaybının Yapay Sinir Ağları İle Belirlenmesi. *Fırat Üniversitesi Fen ve Mühendislik Bilimleri Dergisi,* 18(1), 31-139.

24. İnternet: Geri Yayılımlı Çok Katmanlı Yapay Sinir Ağları. URL: https://medium.com/@billmuhh/geri-yayılımlı-çok-katmanlı-yapay-sinir-ağları-147daa3856247, Son Erişim Tarihi: 06.05.2019.

25. Çetin, M., Uğur, A. ve Bayzan, Ş. (2006). *İleri beslemeli yapay sinir ağlarında backpropagation (geriye yayılım) algoritmasının sezgisel yaklaşımı.* Akademik Bilişim Kongresi, Pamukkale Üniversitesi, Denizli.

26. Tuncer, T. ve Tatar, Y. (2009). *Karar Ağacı Kullanarak Saldırı Tespit Sistemlerinin Performans Değerlendirmesi*, 4. İletişim Teknolojileri Ulusal Sempozyumu, Adana, Türkiye, 2009, 41-48.

27. Sasaki, M. ve Kita, K. (1998). *Rule-based text categorization using hierarchical categories*. In SMC'98 Conference Proceedings. 1998 IEEE International Conference on Systems, Man, and Cybernetics, 3, 2827-2830.

28. Özarslan, S. ve Barışçı, N. (2014). *Öğrenci performansının veri madenciliği ile belirlenmesi.* In ISITES2014: 2nd International Symposium on İnnovative Technologies in Engineering and Science, Karabük, Türkiye (1-8).

29. Uzun,Y. ve Tezel,G. (2012). Rule Learning With Machine Learning Algorıihms And Artificial Neural Networks, *Journal Of Selçuk University Natural And Applied Science 2012, 1(2).*

30. Vijayarani, S. ve Divya, M. (2011). An Efficient Algorithm for Classification Rule Hiding. *International Journal of Computer Applications*, 33(3), 39-45.

31. Akar, Ö. ve Güngör, O. (2012). Rastgele orman algoritması kullanılarak çok bantlı görüntülerin sınıflandırılması. *Jeodezi ve Jeoinformasyon Dergisi*, 1(2), 139-146.

32. İnternet: Karar Ağaçlarında Random Forest Tekniği ile Sınıflandırma. URL: http://www.datascience.istanbul/tag/random-forest-siniflandirma/, Son Erişim Tarihi: 21.05.2019.

ÖZGEÇMİŞ

Kişisel Bilgiler

Soyadı, adı	: AYTAN, Burcu
Uyruğu	: T.C.
Doğum tarihi ve yeri	: 05.04.1986, Ankara
Medeni hali	: Evli
Telefon	: 0 (312) 402 37 29
e-mail	: burcu.aytan@msb.gov.tr

Eğitim

Derece	Eğitim Birimi	Mezuniyet Tarihi
Yüksek Lisans	Gazi Üniversitesi / Bilgisayar Mühendisliği	Devam ediyor
Lisans	Başkent Üniversitesi /Bilgisayar Mühendisliği	2010
Lise	Çubuk Anadolu Lisesi	2004

İş Deneyimi

Yıl	Yer	Görev
2012-Halen	Milli Savunma Bakanlığı	Bilgisayar Mühendisi

Yabancı Dil

İngilizce

Yayınlar

1. Aytan, B., Barışçı, N.**,** *Siber Savunma Alanında Yapay Zeka Tabanlı Saldırı Tespiti ve Analizi*, 2nd International Symposium on Innovative Approaches in Scientific Studies**,** 1384-1390, November 30 - December 2, 2018, Samsun.

Hobiler

Yüzme, Gitar, Yürüyüş, Müzik, Sinema

GAZİ GELECEKTİR..