T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

AKILLI YÖNTEMLER İLE SİBER SALDIRI TESPİT SİSTEMİ GELİŞTİRİLMESİ

Abdulkadir BİLEN

Doktora Tezi

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

Yazılım Bilim Dalı

MAYIS 2021

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Bilgisayar Mühendisliği Anabilim Dalı

Doktora Tezi

AKILLI YÖNTEMLER İLE SİBER SALDIRI TESPİT SİSTEMİ GELİŞTİRİLMESİ

Tez Yazarı
Abdulkadir BİLEN

Danışman
Prof. Dr. Ahmet Bedri ÖZER

MAYIS 2021
ELAZIĞ

T. C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Bilgisayar Mühendisliği Anabilim Dalı

Doktora Tezi

Başlığı:	Akıllı Yöntemler ile Siber Saldırı Tespit Sistemi Geliştirilmesi
Yazarı:	Abdulkadir BİLEN
İlk Teslim Tarihi:	5.4.2021
Savunma Tarihi:	20.5.2021

TEZ ONAYI

Fırat Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına göre hazırlanan bu tez aşağıda imzaları bulunan jüri üyeleri tarafından değerlendirilmiş ve akademik dinleyicilere açık yapılan savunma sonucunda OYBİRLİĞİ ile kabul edilmiştir.

		İmza
Danışman:	Prof. Dr. Ahmet Bedri ÖZER Fırat Üniversitesi, Mühendislik Fakültesi	Onayladım
Başkan:	Prof. Dr. Davut HANBAY İnönü Üniversitesi, Mühendislik Fakültesi	Onayladım
Üye:	Doç. Dr. Taner TUNCER Fırat Üniversitesi, Mühendislik Fakültesi	Onayladım
Üye:	Doç. Dr. Fatih ÖZKAYNAK Fırat Üniversitesi, Teknoloji Fakültesi	Onayladım
Üye:	Dr. Öğr. Üyesi Ali ARI İnönü Üniversitesi, Mühendislik Fakültesi	Onayladım

Bu tez, Enstitü Yönetim Kurulunun ...../...../20..... tarihli toplantısında tescillenmiştir.

İmza

Doç. Dr. Kürşat Esat ALYAMAÇ
Enstitü Müdürü

BEYAN

Fırat Üniversitesi Fen Bilimleri Enstitüsü tez yazım kurallarına uygun olarak hazırladığım "Akıllı Yöntemler ile Siber Saldırı Tespit Sistemi Geliştirilmesi" Başlıklı Doktora Tezimin içindeki bütün bilgilerin doğru olduğunu, bilgilerin üretilmesi ve sunulmasında bilimsel etik kurallarına uygun davrandığımı, kullandığım bütün kaynakları atıf yaparak belirttiğimi, maddi ve manevi desteği olan tüm kurum/kuruluş ve kişileri belirttiğimi, burada sunduğum veri ve bilgileri unvan almak amacıyla daha önce hiçbir şekilde kullanmadığımı beyan ederim.

20/5/2021

Abdulkadir BİLEN

ÖNSÖZ

Siber saldırıların ve suçların ortaya çıkarılması ve tespit edilmesi kişiler, kurumlar ve kuruluşlar için hayati öneme sahiptir. Bu tezde gerçek verilerin elde edilmesi ve bu verilerin veri işleme süreçlerinden geçirilerek analiz edilebilir hale getirilmesi çeşitli zorlukları da beraberinde getirmiştir. Tez çalışması, Elazığ ili ve elde edilen verilerdeki bazı teknik detayların eksikliği ile sınırlıdır. Siber saldırıların zararlarını en aza indirmek ve bu suçlarla mücadele etmeye katkı sağlamak en büyük motivasyon kaynağı olmuştur. Yine bu verileri kullanmama izin veren daha önce görev yaptığım Elazığ Emniyet Müdürlüğüne teşekkürü bir borç bilirim.

Bu tez çalışmasında desteği ve tecrübesiyle her zaman bana katkı sağlayan danışmanım Prof. Dr. Ahmet Bedri Özer'e ve bana destek olan diğer öğretim üyelerimize teşekkür ederim.

Bugünlere gelmemde her zaman maddi ve manevi desteklerini eksik etmeyen annem, babam, ablalarıma, ayrıca eşim Büşra'ya ve kızım Alya'ya da sevgi ve saygılarımı sunarım.

Abdulkadir BİLEN
ELAZIĞ, 2021

İÇİNDEKİLER

Sayfa

ÖZET

Akıllı Yöntemler ile Siber Saldırı Tespit Sistemi Geliştirilmesi

Abdulkadir BİLEN

Doktora Tezi

FIRAT ÜNİVERSİTESİ
Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

Mayıs 2021, Sayfa: xiii + 100

Siber saldırılar dünyanın en büyük sorunlarından biri haline geldiği için ülkelere ve insanlara ciddi maddi zararlar vermektedir. Siber saldırıların artması, siber suçları da beraberinde getirdiğinden suç ve suçlularla mücadelede temel faktörler, siber suç faillerinin belirlenmesi ve saldırı yöntemlerinin anlaşılmasıdır. Siber saldırıları tespit etmek ve bunlardan kaçınmak zor bir görevdir. Ancak araştırmacılar son zamanlarda siber güvenlik modelleri geliştirerek ve yapay zekâ yöntemleriyle tahminler yaparak bu sorunları çözmektedir. Literatürde çok sayıda suç tahmin yöntemi mevcuttur, fakat siber suç ve siber saldırı yöntemlerini tahmin etme daha az çalışılmıştır. Bu sorun, gerçek verileri kullanarak bir saldırıyı ve bu tür saldırının failini belirlemek için ele alınmıştır. Kolluk birimlerinden elde edilen veriler suçun türü, mağdurun bilgileri, zarar ve saldırı yöntemleri gibi özniteliklerden bazılarını içermektedir. Bu tezde siber suçları makine öğrenmesi ve derin öğrenme yöntemleriyle üç farklı modelde analiz edilmiş ve tanımlanan özelliklerin siber saldırı amacı, yöntemi ve failin tespiti üzerindeki etkisini ortaya koyulmuştur. Tezde makine öğrenmesi, derin öğrenme ve yapay sinir ağları yöntemleri kullanılmıştır ve doğruluk, kesinlik, duyarlılık ve F1-skor açısından değerlendirilmiştir. Saldırı amacı tahmin modelinde Destek Vektör Makinası, Karar Ağaçları ve Rastgele Orman algoritmaları %77.35 doğruluk oranıyla en iyi performansı sergilemiştir. Saldırı faili tahmin modelinde %69.61 doğruluk oranıyla En Yakın Komşu algoritması en iyi performansı göstermiştir. Saldırı yöntemini tahmin eden modelde ise %96.69 doğruluk oranıyla Lojistik Regresyon algoritması en iyisi olmuştur. Sonuçlar, mağdurun eğitim ve gelir düzeyi arttıkça siber saldırı olasılığının azaldığını ortaya koymuştur. Ayrıca siber saldırıların tespit edilmesini kolaylaştıracak ve bu saldırılarla mücadeleyi daha kolay ve etkili hale getirecektir.

Anahtar Kelimeler: Siber Saldırı-Suç, Makine Öğrenmesi, Derin Öğrenme, Suç Tahmini, Yapay Zekâ

ABSTRACT

Cyber Attack Detection System Development with Intelligent Methods

Abdulkadir BİLEN

Ph.D. Thesis

FIRAT UNIVERSITY
Graduate School of Natural and Applied Sciences

Department of Computer Engineering

May 2021, Pages: xiii + 100

As cyber-attacks have become one of the biggest problems in the world, they cause serious financial damage to countries and people. Since the increase in cyber-attacks brings along cyber-crimes, the key factors in the fight against crime and criminals are the identification of cybercriminal perpetrators and understanding the attack methods. Detecting and avoiding cyber-attacks are hard tasks. However, researchers have recently been solving these problems by developing cyber security models and making predictions with artificial intelligence methods. There are many crime prediction methods in the literature but predicting cyber-crime and cyber-attack methods has been less studied. It addresses this issue to determine an attack and the perpetrator of this type of attack using real data. The data got from law enforcement units include the type of the crime, the information of the victim, the harm and attack methods. In thesis, it analyzed cyber-crime in three different models with machine learning methods and revealed the effects of the defined features on the purpose, method, and detection of the perpetrator. Machine learning, deep learning and artificial neural networks methods used in our approach and evaluated in terms of accuracy, precision, recall and F1-score. In the attack aim prediction model, the Support Vector Machine, Decision Trees and Random Forest algorithms performed the best with 77.35% accuracy. The Nearest Neighbor algorithm showed the best performance with an accuracy rate of 69.61% in the offender prediction model. In the model that predicted the attack method, the Logistic Regression algorithm was the best with an accuracy rate of 96.69%. Results show that as the education and income level of the victim increases, the probability of cyberattack decreases. It will also make it easier to detect cyberattacks and make the fight against these attacks easier and more effective.

Keywords: Cyber Attack-Crime, Machine Learning, Deep Learning, Crime Prediction, Artificial Intelligence

ŞEKİLLER LİSTESİ

Sayfa

TABLOLAR LİSTESİ

Sayfa

EKLER LİSTESİ

Sayfa

KISALTMALAR

Kısaltmalar

ABD : Amerika Birleşik Devletleri
AI : Artificial Intelligence
ANN : Artificial Neural Network
ARP : Address Resolution Protocol
ASP : Active Server Pages
BT : Bilgi Teknolojileri
DDoS : Distributed Denial of Service (Dağıtılmış hizmetin Reddi)
DL : Deep Learning
DNN : Deep Neural Network (Derin Sinir Ağı)
DoS : Denial-of-service (Hizmetin Reddi)
DÖ : Derin Öğrenme
DT : Decision Tree (Karar Ağaçları)
FISMA : Federal Information Security Management Act
HIPAA : Health Insurance Portability and Accountability Act
HTML : Hyper Text Markup Language
HTTPS : Hypertext Transfer Protocol Secure
ICMP : Internet Control Message Protocol
ICSA : Intelligent Cyber Security Assistant (Akıllı Siber Güvenlik Asistanı)
IoT : The Internet of Things (İnternet Şeyleri)
IP : Internet Protocol
KNN : K-Nearest Neighbors (K-En Yakın Komşu)
LR : Logistic Regression
ML : Machine Learning
MÖ : Makine Öğrenmesi
NB : Naive Bayes
PHP : Personal Home Page
RF : Random Forest (Rastgele Orman)
SCADA : Modern Supervisory Control and Data Acquisition
SQL : Structured Query Language
SSh : Secure Shell
SVM : Support Vector Machine (Destek Vektör Makinesi)
SYN : Synchronize
TCP : Transmission Control Protocol
XGB : eXtreme Gradient Boosting
VPN : Virtual Private Network
YSA : Yapay Sinir Ağları

1. GİRİŞ

Bilgisayarlar, cep telefonları, tabletler ve giyilebilir teknolojiler gibi ürünlerin hayatımızda yerini almasıyla birlikte, kişisel verilerin saklanması, siber güvenlik, yapay zekâ gibi kavramlar da hayatımıza girmiştir. Bu teknolojik gelişmeler, kişisel veri kavramını ve bunların mahremiyetini ön plana çıkarırken, bu verilerin korunması ve kötü niyetli üçüncü şahısların eline geçmemesi adına devletleri ve şirketleri siber güvenlik için büyük bütçeler ayırmaya zorlamıştır. Kişisel verilerin güvenliğini sağlamak için yeri geldiğinde bütün imkânlar kullanılarak siber güvenlik önlemleri doğru şekilde alınsa da yine ihlaller olabilmektedir. Dijital dünyada siber saldırıları önleme ve bu saldırılara müdahale etme çabaları özellikle finansal ve kritik altyapıya sahip kurumlarda giderek artmaktadır. Bu altyapılara yönelik saldırı ve tehditlerde siber olaylara zamanında ve etkin müdahale etmek gerekmektedir. Günlük hayatımızda cep telefonlarından banka uygulamalarına, sosyal medya hesaplarına ve internet sağlayıcılarına erişim arttıkça artık her an bir siber saldırı tehdidine maruz kalınmaktadır ve bununla birlikte maddi manevi zararlar meydana gelmektedir. Özel hayata dair birçok verinin internette olmasının sağladığı kolaylıkların yanında getirdiği sorunların başında da bu verilerin güvenliğini sağlamak gelmektedir.

Teknolojinin bize getirdiği kavramlardan bazıları da yapay zekâ, makine öğrenmesi, derin öğrenme ve yapay sinir ağlarıdır. Artık insanoğlu bazı şeyleri kendisi düşünmek yerine makinalara öğretip sonuç almaya başlamıştır. Bilgisayarlar, birçok alanda bizim kullandığımız araçlara göre veriler elde etmekte, örneğin internetten yapacağımız alışverişten girdiğimiz sitelere kadar verileri toplayarak bir sonraki aşamada bize öneriler sunmaktadır. Bu durum hayatımızı kolaylaştırıyor gibi görünse de önümüzdeki dönemde karşımıza daha ne gibi sorunların çıkacağı bilinmemektedir. Siber güvenliği sağlamak amacıyla, önceki meydana gelen siber saldırılardaki verileri kullanarak bir sonraki siber saldırının nereden ve nasıl gerçekleştirileceğini tahmin ederek insanların yerini bilgisayarların alması sağlanarak daha hızlı ve etkin çözümler sunulmaktadır. Yapay zekâ yöntemlerini kullanarak siber tehditleri algılamak, suç ve suçlu ile mücadele etmek insan gücüne katkı sağlayacak ve bizleri teknolojik gelişmede bir adım daha ileriye götürecektir. Elde edilen siber veriler, yapay zekâ algoritmalarına öğretilerek insan gücünden daha hızlı çalışan sistemler ile siber olaylara müdahale daha etkili hale getirilmektedir.

Siber güvenliği sağlamak, alan bilgisine dayanan ve büyük miktarda ağ verisinden olası tehditleri belirlemek için bilişsel yetenekler gerektiren karmaşık bir görevdir. Siber güvenlik konusunda bilgi sahibi olan kişilerin kötü niyetli olayları, bilgisi olmayanlara göre daha iyi algıladığı gözlemlenmiştir. Bilgisayarlar, akıllı telefonlar, tabletler, akıllı medya TV vb. ağ sistemlerine izinsiz giriş tespiti, bilgi ve ağ güvenliğinde kapsamlı alan bilgisi olmayan birçok son kullanıcının endişesi haline gelmiştir [1]. Siber saldırılara etkili bir şekilde yanıt vermek hem zor hem de kritiktir. Saldırıları otomatik olarak algılamak ve saldırganları yeniden saldırı

başlatamayacak hale getirmek için gelişmiş makine öğrenimi gerekmektedir. Bu zorlukları etkin bir şekilde ele almak için, siber uzay kaynaklarını ve hizmetlerini hem mevcut hem de yeni saldırılara karşı etkin ve hızlı bir şekilde savunarak güvenlik analistlerine yardımcı olacak ve günlük operasyonel yükleri hafifletecek akıllı siber asistan tasarlanmıştır. Akıllı Siber Güvenlik Asistanı mimarisiyle, siber güvenlik uzmanına yardım eden çeşitli savunma ve istihbarat uygulamalarında doğru tahminler yapılmıştır [2]. Otonom siber savunma sistemleri yapay zekâ yöntemleriyle birleştirilerek siber saldırılara nasıl daha etkin müdahale edileceği çalışılmıştır. Geleneksel siber savunma yöntemleri incelenmiş, siber savunmadaki yapay zekâ uygulamaları gözden geçirilmiş, siber savunma yöntemlerini daha iyi hale getirmek için yapay zekânın kullanılabileceği çeşitli yöntemler araştırılmıştır [3]. Tüm bilim ve araştırma alanları hızla gelişirken bilimin karşılaştığı zorlukları aşmak için yeni araçlar ve öğrenme teknikleri geliştirilmektedir. Bu araçlardan bir tanesi büyük veriyi incelemek için makine öğrenme tekniklerinin zorlukları ve çözümleridir [4]. Siber güvenlik uygulamaları için derin öğrenme algoritmalarından birkaç tanesi kullanılmıştır. Özellikle kötü amaçlı yazılım, gereksiz eposta, içeriden gelen tehditler, ağ müdahaleleri, yanlış veri enjeksiyonu ve botnet'ler tarafından kullanılan kötü amaçlı alan adları gibi çok çeşitli saldırı türleri ele alınmıştır. Siber güvenlik uzmanlarına yönelik olarak, ağ saldırılarındaki ihlalleri tespit etmek için makine öğrenmesi ve derin öğrenme yöntemleri kullanılarak kötü niyetli saldırıların önlenmesinde yapay zekâ yöntemlerinin kullanılmasının önemi vurgulanmıştır. Araştırmada siber güvenlik sistemleri için derin öğrenme teknolojisini ilerletmek üzere çalışmalar yapılmıştır [5].

Tanımlaması zor olan kötü niyetli yazılımları tespit etmek için derin öğrenme yaklaşımıyla gerçek veriler üzerinde akış tespiti gerçekleştirilmektedir. Kötü amaçlı yazılımları tespit etme yeteneğini değerlendirmek için derin öğrenme yöntemi kullanılmış üstün bir performansta doğru bir şekilde algılanabilmektedir [6]. Son yıllarda verdiği zararlarla gündeme gelen siber saldırıları ilkel yöntemlerle ele almak sıkıcı ve zor bir iş haline gelmiştir. Derin öğrenme alanı, veri analizindeki hemen hemen her soruna uygulanabilecek çok sayıda güçlü algoritmaya sahiptir. Siber risklere verilen yanıt ve çözümlere yönelik artan talep göz önüne alındığında, siber güvenlik sorunlarında diğer derin öğrenme tekniklerinin uygulanması gelecekteki çalışmalar için çok büyük bir olasılıktır [7].

Siber güvenlik alanında makine öğrenmesi ve derin öğrenme yöntemleri, geleneksel kural tabanlı algoritmalara karşı daha üstün olduğunu göstermiştir. Saldırı tespiti, kötü amaçlı yazılım analizi ve spam tespitinde yapay zekâ teknikleri otomatize ettiği için daha güçlü olduğu gözlemlenmiştir. Akıllı öğrenme yöntemlerinin otonom yetenekleri fazla tahmin edilememektedir. İnsan denetiminin olmaması, yetenekli saldırganların sızmasını, verilerin çalınmasını ve hatta bir işletmenin sabote edilmesini zorlaştırabilir [8]. Derin öğrenme yöntemleri, konuşma tanıma, görüntü sınıflandırmaları ve dil işlemede öğrenme yöntemleri gibi çeşitli bilim ve mühendislik

alanlarında yaygın olarak uygulanır. Büyük Veri analizi, verileri gerçek zamanlı olarak yüksek doğruluk ve verimlilikle işlemek için makine ve derin öğrenme tekniklerine dayanan yeni ve gelişmiş algoritmalar gerektirir. Son zamanlarda araştırmalar, yüksek hızlı veri işleme hibrit öğrenme ve eğitim mekanizmalarına sahip çeşitli derin öğrenme tekniklerini içermektedir. Çalışmada büyük veri analizi için kullanılan çeşitli derin öğrenme tekniklerinin karşılaştırmalı bir analizi sunulmuştur. Derin öğrenme teknikleri, derin inanç ağlarına ve evrişim sinir ağlarına dayalı olarak büyük veri öğrenme ve eğitimi için geniş bir şekilde sınıflandırılmıştır. Derin öğrenme yöntemlerinde, büyük verilerin mevcut tekniklerle işlenmesinde bazı sınırlamaları vardır. Bir dizi nöron ve gizli katmanlarla farklı veri kümelerinde test edilerek daha da detaylandırılmıştır. Gizli katmanların sayısı altıdan artarsa, derin bir öğrenme sürecini çözmek imkânsızdır [9]. IoT (The Internet of things) ve yeni ortaya çıkan IoT uygulamaları, ağın kenarında dağıtılan yeni siber güvenlik kontrolleri, modelleri ve kararları gerektirir. Geleneksel kriptografik çözümlerin başarısına rağmen, sistem geliştirme kusurları, artan saldırılar ve hackleme becerilerinden dolayı algılama mekanizmalarının kurulması kaçınılmazdır. DÖ yöntemleriyle oluşturulan bu algılama yöntemleriyle, siber saldırı tespitinin doğruluğunu ve verimliliğini artırılabileceği sonucuna varılmıştır [10]. HTML5 tabanlı uygulamalarına karşı uygulanan kod enjeksiyonu saldırılarını tespit etmek için hibrit DÖ ağı oluşturulmuş ve saldırılar tespit edilmiştir [11]. Terör ve şiddet eylemleri, kritik altyapıya sahip tesislere saldırılar, salgın hastalıklar gibi acil durumlarda sistem felaketlerini tespit ve analiz etmek için DÖ teknolojisi kullanılmıştır. Bu sayede gelecekte, suç faaliyetlerini tespit etmek ve DÖ tabanlı tespitin geleneksel gözetim sistemlerine göre doğruluk ve tespit gecikmesi açısından avantajları ortaya koyulmuştur [12]. Elektrik altyapılarının eskimesinden dolayı siber saldırılara daha savunmasız hale gelmesiyle birlikte siber saldırıları etkili bir şekilde tespit etmek için yenilikçi bir şekilde aralık durum tahmini tabanlı bir savunma mekanizması geliştirilmiştir. DÖ durum değişkenlerinin aralıklarının genişliğini daraltabildiği ve böylece veri manipülasyonunu tespit etmede başarı oranını büyük ölçüde artırdığı gözlemlenmiştir [13].

Kim ve arkadaşları tarafından yapılan çalışmada makine öğrenimi temelli suç tahmini araştırılmaktadır. Son 15 yıllık Vancouver suç verileri iki farklı veri işleme yaklaşımı kullanılarak analiz edilmiş, MÖ algoritmalarından K-En Yakın Komşu ile artırılmış karar ağacı uygulanmış ve %39 ile %44 arasında doğru tahminler yapılmıştır [14]. Duan ve arkadaşları tarafından suç tahmini için yeni bir derin evrişimli sinir ağı tabanlı modelle, suç ve veriye dayalı 4 temel yöntemle karşılaştırılarak suç tahminleri yapılmıştır. Çalışmada farklı suç riski kategorilerini modelleme, suçları önlemek için önemli sosyoekonomik özellikleri belirleme gibi daha fazla araştırma için birçok yol olduğu sonucuna varılmıştır [15]. Chandrasekar ve arkadaşları Amerika Birleşik Devletleri San Francisco şehrinde meydana gelen 39 farklı suç kategorisini makine öğrenmesi yöntemleriyle sınıflandırılmıştır. 2 sınıflı bir sınıflandırma problemine çok uygun olan mavi yakalı

/ beyaz yakalı suç sınıflandırma probleminde yüksek doğruluk ve hassasiyet elde etmişlerdir [16]. Ahishakiye ve arkadaşları tarafından kolluk kuvvetlerinin suçla mücadele etmesini kolaylaştırmak için karar ağacı algoritması kullanılarak suç tahmini prototip modelinin geliştirilmesi amaçlanmıştır. Suç verilerinin tahmininde en verimli makine öğrenmesi olarak %94,25 doğrulukla tahmin etmişlerdir [17]. Suç ve suçlu takibinde doğru veri tabanı tutularak, suçların öngörülmesine ve çözülmesine yardımcı olmak için analiz edilmelidir. Alkesh Bharati ve Sarvanaguru Rak tarafından yapılan bu çalışmada çok sayıda suçtan oluşan veri kümesini analiz etmek ve çeşitli koşullara bağlı olarak gelecekte meydana gelebilecek suç türünü tahmin etmektir. Chicago şehrinde işlenen suçun yeri, türü, tarihi, saati, enlemi, boylamı gibi verileri makine öğrenimi yöntemlerinden K-En Yakın Komşu (KNN) sınıflandırması ve diğer çeşitli algoritmalar ile gelecekte meydana gelecek suçlar tahmin edilmiştir. Yapılan çalışmayla 0.789 hassasiyetle suç türünü tahmin ederek veri görselleştirme ile veri kümesinin analizine yardımcı olmaktadır [18]. Suç analizi ve tahmini, suçtaki farklı örüntüleri, ilişkileri ve eğilimleri analiz etmek ve tanımlamak için kullanılan sistematik bir yaklaşımdır. San Francisco, Chicago ve Philadelphia'daki suç verileri öncelikle sınıflandırılmış daha sonra görselleştirme ve analiz yapılmıştır. Suç tahminleri yapılırken veri madenciliği kullanılarak suç eğilimleri tahmin edilmiştir. Polis departmanı ve kolluk kuvvetleri için faydası, suçları çözme sürecini hızlandırmak, suç faaliyetlerini takip etmek, kaynakları etkili bir şekilde dağıtmak ve daha hızlı karar vermelerini sağlayacak öngörüler sunmasıdır [19]. Twitter aracılığıyla paylaşılan mesajlar, karar destek sistemleriyle suç tahmini için kullanılmaktadır. Alınan mesajlar bir suç tahmin modeline eklenmiş ve 25 suç türünden 19'u için suç yoğunluğu tahminine dayalı standart bir yaklaşımla suç tahmini yapılmıştır. Bu sonuçlar sayesinde adliye teşkilatının daha kolay karar vermesine ve polis devriyeleri gibi harcanan çabanın azalmasına ve suça müdahale sürelerinde azalmaya yol açacağı düşünülmektedir [20]. MÖ ve DÖ algoritmaları özellikle suç tahmininde kullanılmıştır. Yapay zekâ teknikleri, veri madenciliği yöntemlerinin yerini almaktadır. Suçlar eğitim ve test verileri olarak ikiye ayrılmış ve oluşturulan modelle tahmin edilmiştir. Gerçekleştirilen modelle gelecekte olabilecek olaylar önceden bilinirse daha etkili tedbirler alınabileceği düşünülmektedir [21].

Alexander Stec ve Diego Klabjan tarafından yapılan çalışmada Chicago ve Portland suç verilerini kullanarak, zamansal ve mekânsal analizler yapılmıştır. Derin öğrenme algoritmalarıyla sırasıyla %75,6 ve %65,3 doğrulukla tahmin edilmiştir. Toplu taşıma ve hava durumu verisiyle desteklenen çalışmada suçların hafta içi ve hafta sonunda ve yoğunluğa göre farklılık gösterdiği tespit edilmiştir [22]. Chackravarthy ve arkadaşları tarafından yapılan çalışmada Hibrit Derin Öğrenme algoritması ile sinir ağları kullanılmıştır. Alınan video akışı verileri analiz edilmiş ve suçun tespiti için bir çözüm önerilmiştir. Denetleyici yetkililer üzerindeki iş yüklerini azaltacak, suç faaliyetlerini hızlı bir şekilde tanımlayabilecek, akıllı şehir altyapısı boyunca uygulandığında verimli ve uyarlanabilir olacak bir suç tespit sistemi sağlanması amaçlanmıştır. Çalışmadaki suç

tespit sisteminin, büyük bir veri seti üzerinde doğru bir performans sergilediği gözlemlenmiştir [23]. Kang ve arkadaşları tarafından yapılan çalışmada Chicago Illionis'deki suç istatistikleri ile demografik ve meteorolojik veriler toplanmış ve dört tür katmandan oluşan Derin bir sinir ağıyla (DNN) eğitilmiştir. Suç oluşumunu tahmin etmede diğer yöntemlere göre daha yüksek doğrulukta olduğu görülmüştür [24]. Shermila ve arkadaşları tarafından yapılan çalışmada suçu analiz etmek ve failini tahmin etmek amaçlanmıştır. Yıl, ay ve silah gibi faktörlerin suç üzerindeki etkisi ile failin cinsiyeti ve yaşı tahmin edilmiştir. Çoklu lineer regresyon, en yakın komşu ve sinir ağları algoritmalarıyla yapılan tahminlerde fail yaşı için %60, fail cinsiyeti için %96 ve ilişki için %97 doğruluk elde edilmiştir [25]. Chintal ve arkadaşları tarafından yapılan bu çalışmada Hindistan'daki siber suç eğilimlerini tahmin etmek ve analiz etmek için web sitesinden gerçek veriler toplanmıştır. Gradyan Azaltma yaklaşımıyla lineer regresyon algoritması kullanılarak siber suç bölge tahmini yapılmıştır. Siber suçları kontrol etmek için gelecek yıllardaki bölge suç haritaları oluşturulmuştur [26]. Arora ve arkadaşları tarafından analiz edilen siber suç verileri, Facebook ve Twitter gibi sosyal medya platformlarından halka açık verilerdir. Algoritmaları doğruluk ve kesinlik derecesi olan F-skor değerine göre karşılaştırılmıştır. %80 doğrulukla, senaryoya en iyi uyan Rastgele Orman algoritması bulunmuştur. Tehditler, siber suçları analiz eden model aracılığıyla otomatik olarak belirlenmiştir [27]. Ch ve arkadaşları tarafından Hindistan'da meydana gelen siber suç olay verileri, makine öğrenimi teknikleri kullanılarak sınıflandırılmıştır. Suçları %90 doğrulukla öngören model, analiz ve manuel raporlamada harcanan zamanı azaltmıştır [28].

Tezde, bilimsel problem olarak siber suçlardan yola çıkarak, bu suçların hangi amaçla, hangi yöntemle ve kimler tarafından gerçekleştirildiğinin ortaya çıkarılması hedeflenmektedir. Literatür incelemelerinde, siber saldırıların ve suçların kişilerle ve devletlere verdiği ciddi zararlar nedeniyle bunlarla mücadele etmenin önemine vurgu yapılmaktadır. Son yıllarda artan siber suçlarla ya da saldırılarla başa çıkabilmek için önemli düzenlemeler ve çeşitli araştırmalar yapılmaktadır. Bu sebeple siber saldırılarla mücadele etmek amaç olarak seçilmiştir.

Tezin amaçları, bir siber suç yönteminin tahmin edilmesi ve doğruluk sonuçlarını karşılaştırmak için girdi olarak gerçek siber suç verilerinin kullanılmasıdır. Siber suç faillerinin mevcut verilere dayanarak tahmin edilip edilemeyeceğinin ölçülmesi ve saldırgan / kurban profillerinin siber saldırılar üzerindeki etkisinin anlaşılmasıdır. Siber saldırıların hangi amaçla gerçekleştirildiğini bularak diğer parametreler arasındaki ilişkilerin ortaya çıkarılmasıdır. Mağdurlara ait yaş, cinsiyet, gelir, eğitim, medeni hal, meslek ve suçun verdiği zarar gibi özellikler ile failler ve suçlar arasındaki ilişkilerin ortaya çıkarılmasıdır. Tezde, olaylarla ilgili toplanan verilerin doğru analiz edilmesi, suçların engellenmesi ve faillerin yakalanması amaçlanmaktadır. Analiz edilen verilerden sonuçlar çıkarmak ve sonuca göre suçlarla mücadele etmektir. Bu sonuçlar, kolluk kuvvetleri tarafından yürütülen soruşturmaları ve gizlenmiş gerçekleri ortaya çıkararak geleceğe ışık tutacaktır. Mağdur hakkındaki bilgilere, siber saldırının yöntemine ve failin

tespit edilip edilmediğine bağlı olarak, siber saldırıyı aynı failin gerçekleştirip gerçekleştirmediğini belirlemek için makine öğrenimi yöntemleri kullanılmıştır.

Literatürde makine öğrenimi ve derin öğrenme yöntemlerinin hangi alanlarda kullanıldığı ve bu yöntemlerin tahminlerde başarılı olduğu gözlemlendiğinden çalışmamızda bu metotlar kullanılmıştır. Veri setindeki öznitelikler kullanılarak saldırıların amacı, faili ve yöntemi tahmin edilmiştir.

Tezde sunulan model, özellikle suç soruşturması yürüten birimlere ve literatüre kazandırılmıştır. Literatürde veri seti olarak genel suç verisi, siber suç verisi ve saldırı yöntem verisi kullanılsa da kişisel özniteliklere dayalı gerçek verilerin daha az çalışılmış olması sebebiyle çalışmadaki veri setiyle makine öğrenimi tabanlı model önerilmiştir.

Elde edilen sonuçlar, siber suçlarla ilgilenen kolluk birimleri tarafından siber suç modellemesi ve saldırı tespitinde fayda sağlayacaktır. Önerilen yaklaşımın başlıca katkıları; gerçek veri seti kullanıldığından siber suç departmanına avantaj sağlaması, mağdurların maruz kalabileceği saldırıların tahmin edilmesi ve makine öğrenimi algoritmalarını karşılaştırarak optimum performansın belirlenmesine olanak sağlamasıdır. Yine mağdurların özelliklerine göre gelecekte işlenebilecek suçlara karşı önleyici tedbirlerin alınmasıdır. Ayrıca siber suçları tahmin etme, bu suçların faillerini ve mağdurlarını profilleme konusunda polis departmanına avantajlar sağlayacaktır. Önerilen model sonucunda, kişiye özel tedbirlerin alınması sağlanacak ve insanların maruz kalabilecekleri suçlar hakkında bilgilendirilmesi kolaylaşacaktır.

2. SİBER GÜVENLİK

Siber güvenlik; veri, yazılım, donanım, sistem, süreç ve kapasite güvenliğinin siber dünyada korunmasıdır. Genel olarak bilişim sistemlerinin korunmasını ve bu ortamın gizliliğinin, bütünlüğünün ve erişilebilirliğinin sağlanması, meydana gelebilecek herhangi bir saldırının tespiti ve önlenebilmesidir [29]. Siber güvenlik, çeşitli siber saldırıları kapsayan konuların anlaşılması ve her türlü dijital ve bilgi teknolojisinin gizliliğini, bütünlüğünü ve erişilebilirliğini koruyan savunma stratejilerini tasarlama ile ilgilidir [30]. Bilgi güvenliği kapsamında değerlendirildiğinde, bilgi sistemleri içinde yer alan donanımlar, yazılımlar ve veriler sahibi için önem arz etmektedir. Yani bu kavramlar kişin sahip olduğu varlıklarıdır, bu varlıkların bozulmamasını ve değiştirilmemesini sağlamak güvenlik olarak değerlendirilmektedir [31]. Bilgi Güvenliği, fiziksel ve dijital verilerin yetkisiz kullanım, erişim, aksama, değiştirme, imha veya kayıttan korunmasını garanti eder. Herhangi bir kurum veya kuruluşun ilk başlaması gereken güvenlik programlarından biri bilgi güvenliğidir. Siber güvenlik, kuruluşun bilgisayarlarının, verilerinin ve ağının yetkisiz dijital saldırı, erişim veya hasardan çeşitli süreçler, uygulamalar ve teknolojilerin uygulanması yoluyla korunmasıdır [32]. Siber güvenlik yalnızca bilgi kaynaklarının değil, aynı zamanda kişinin kendisi de dahil olmak üzere diğer varlıkların korunmasını da içerecek şekilde geleneksel bilgi güvenliği sınırlarının ötesine geçmektedir. Bilgi güvenliği, bir varlık olan bilginin çeşitli tehdit ve güvenlik açıklarından kaynaklanan olası zararlardan korunmasıdır. Öte yandan siber güvenlik, yalnızca siber uzayın kendisinin korunması değil, aynı zamanda siber uzayda görev yapanların ve siber uzay yoluyla ulaşılabilen varlıklarından herhangi birinin korunmasıdır, yani siber saldırıların potansiyel hedefi olan insanlar ve hatta bilmeden siber saldırıya katılan insanlardan oluşmaktadır [33].

Şekil 2.1. Bilgi ve iletişim güvenliği, bilgi güvenliği ve siber güvenlik arasındaki ilişki [33]

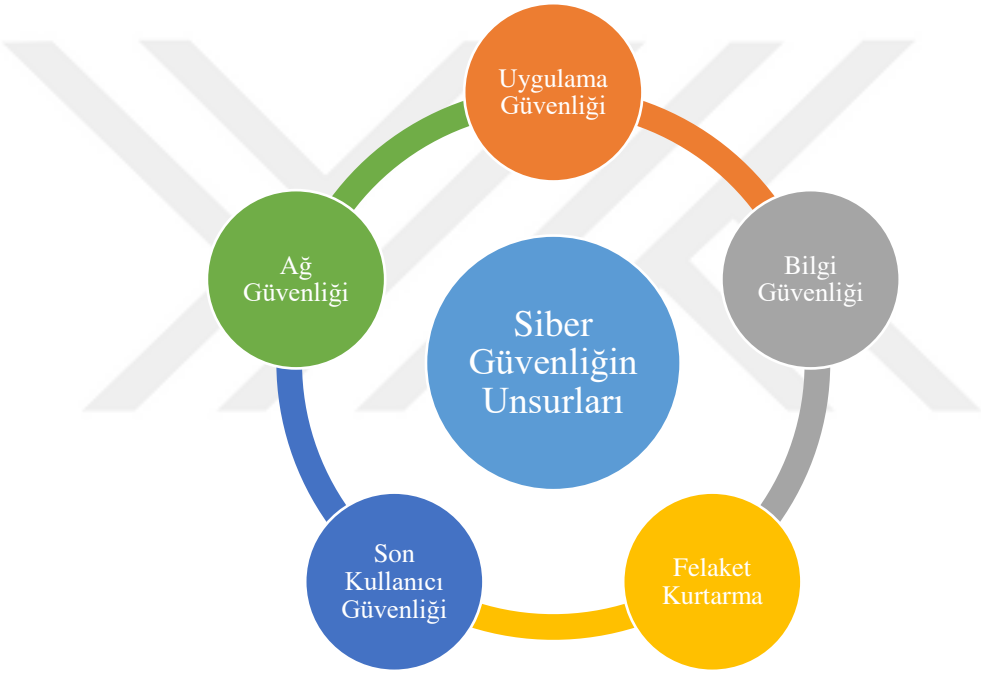
Şekil 2.1'de bilgi güvenliği ve siber güvenlik arasındaki ilişki ortaya koyulmuştur. Gelişmiş ülkelerde internet kullanımı arttıkça kişiler ve kurumlar güvenliklerini artırmalılardır. Kimlik doğrulama, güvenlik denetimi, yama yönetimi gibi tedbirleri alıp kritik altyapılarında kullandıkları yazılım ve donanımların güçlü ve zayıf yönlerini tespit etmeleri gerekmektedir. Sisteme gelebilecek saldırılardan korunma ve uygulamalarını iyileştirmeyi amaçlayan analiz çalışmalarını artırmalılardır [34]. Siber güvenliği sağlamada yeni tedbirler almanın yansıra, geniş bir ağ operasyonu ve bilgi güvenliği bilgisine sahip profesyonel siber güvenlik uzmanlarına görev vermek gereklidir. Siber güvenlik uzmanları sayesinde insan ile makine etkileşimi üst düzeye çıkarılarak, siber saldırılara müdahalede daha etkin bir hal alacaktır [35]. Bilgi ve ağ güvenliğindeki birçok teknolojik ilerleme saldırı tespit sistemlerini ortaya çıkarmış, analistler için gelişmiş izleme ve tehdit tespitini kolaylaştırmıştır, ancak gerçekleştirdikleri görevler tamamen otomatikleştirilememektedir. Analistin, meydana gelen saldırının kurumun ağına yönelik olup olmadığını belirlemesi önemli bir gerekliliktir [1]. Derin öğrenme yöntemlerinin son yıllarda çok hızlı bir şekilde gelişmesi, yeni bir yapay zekâ çağını başlatmış ve akıllı saldırı tespitleri geliştirmek için tamamen yeni bir yöntem ortaya çıkarmıştır. Derin öğrenme, çok daha iyi modeller oluşturmak için verilerden daha iyi tahminler çıkarma potansiyeline sahiptir [36].

Günümüzde tehdit ortamı sürekli değiştiğinden siber saldırılara karşı ulusal tedbirler almanın önemi her geçen gün daha önemli hale gelmektedir. Suç grupları, siber dolandırıcılıkları gerçekleştirmek ve sonuçlarını araştırmak için birçok siber savaşçıyı işe almaktadır. Bu gruplarla mücadele edebilmek için öncelikle kişisel düzeyde gelişmişlik ve sonrasında ulusal savunma şeklinde ilerleme kaydetmek bir gereklilik haline gelmiştir [37]. Son dönemlerde artan siber saldırılardan dolayı şirket ve kurumların maddi zararları da artmaktadır. Hatta yapılan ankete göre, katılanların %69'u yaşamlarında bir siber saldırıya maruz kaldıklarını ve her gün milyonlarca saldırının meydana geldiğini bildirmişlerdir [38].

2020 yılında yayınlanan Checkpoint güvenlik raporuna göre dünyanın en büyük ikinci ticari uçak üreticisi Airbus firması, bilgi sistemlerine girilerek çalışanlarının kişisel verilerini açığa çıkaran bir saldırı gerçekleştirildiğini açıklamıştır. Yine birçok hesap detayı çalınarak dark web pazarında satışa sunulmuştur. Mcafee, Symantec ve Trend Micro Anti virüs şirketlerine ait ağlar ihlal edilerek uzun vadeli uzaktan erişim sağlanmış ve satışa sunulan veriler çalınmıştır. Yeni Zellanda'da bir milyon kişinin tıbbi verilerini ele geçiren bilgisayar korsanı bu bilgileri satışa sunmakla tehdit etmiştir. New Orleans belediye başkanı, şehrin hizmetlerini kesintiye uğratan bir saldırının ardından olağanüstü hâl ilan etmiştir [39]. Siber saldırılar, bir ülkeyi veya bir kuruluşu derinden etkilediğinden saldırıları önlemek ve riski minimuma indirgemek oldukça önem arz etmektedir.

2.1. Güvenlik Unsurları

Güvenlik unsuru denildiğinde ilk akla gelen; Şekil 2.2'de gösterildiği gibi uygulama güvenliği, bilgi güvenliği, ağ güvenliği, felaket kurtarma planı, operasyonel ve son kullanıcı güvenliği gibi beş temel unsurdur. Siber güveliğin sağlanması için yapılması gereken en önemli elementler; bütünlük, gizlilik, kimlik doğrulama, inkâr edilemezlik ve erişilebilirliktir. Gizliliği, bütünlüğü ve kullanılabilirliği destekleyen teknolojiler genellikle birbirleriyle çelişmektedir. Örneğin; siber uzayda bilgi için yüksek düzeyde erişilebilirlik sağlamak gizliliği zor hale getirdiğinden, bu unsurların ne anlama geldiğini ayırma konusu siber güvenlik uzmanlık alanına girmektedir [30].



Şekil 2.2. Siber Güvenlik Unsurları

2.1.1. Gizlilik

Herhangi bir sistemdeki bilgilerin yetkisiz kişilere ifşa edilmesini engellemektir. Yetkisiz bir kişi depolanan bilgileri okuyabilir, değiştirebilir veya bunlardan yararlanabilir. Bu kaygılar sisteme sızanların yalnızca bilgi kullanım modellerini gözlemlediği trafik analizidir. Bu kişinin bilgileri çıkarıp özel alanımızda bulunan programları izinsiz kullanımını da içermektedir [40].

2.1.2. Bütünlük

Yetkisiz bir şekilde herhangi bir veri değişikliğini veya silinmeyi önlemek için kullanılan terimdir. Sistemde tutulan veya başka kişilere gönderilen bilgilerin değişmediğinin, tam ve eksiksiz

gönderildiğini doğrulamaktır. Veri gönderilirken özetleme alınarak verinin bozulmadan gidip gitmediği anlaşılabilmektedir [40].

2.1.3. Sürekli Erişebilirlik

Bilgilerin kişi veya kuruluşlara iletilmesinden, depolanmasından ve işlenmesinden sorumlu sistemlere, gerektiğinde ve bunlara ihtiyaç duyanlar tarafından sürekli olarak erişilebilir olmasıdır. Siber saldırı esnasında yetkili kullanıcının bilgiye erişmesi engellenebilmekte veya bilgi değiştirilebilmektedir [40].

2.1.4. Kimlik Doğrulama

Herhangi bir sisteme erişim yapmaya çalışan kullanıcının yetkili olduğu ve yetkilendirilen kişi olduğunun tespit edilmesi veya kesinlikle aynı kullanıcı olduğuna dair yazılım veya donanımla denetlenmesidir [29].

2.1.5. İnkâr Edilemezlik

Veriyi ileten veya iletilmesini sağlayan kişinin göndericiye ilettiği veriyi doğrulaması ve alıcı tarafından inkâr edilememesidir. Özellikle gerçek zamanda işlem gerektiren finansal alanda kullanılmaktadır. Bu unsuru sağlamak için öncelikle bütünlük ve kimlik doğrulama unsurlarının gerçekleşmesi gerekmektedir [29].

2.2. Siber Saldırılar

Siber saldırı, herhangi bir sistem içinde bulunan kritik veriyi ele geçirmek, değiştirmek veya yok etmek amacıyla bilgisayar ya da ağları hedef almaktadır. Kişi veya gruplar tarafından gerçekleştirilebilen saldırının amacı, bilgi sistemini ele geçirerek yıkıcı etkiler yaratmaktır. Saldırılar; bir kurum veya kuruluşun sistemi ele geçirilmeye çalışıldığında hedefli, geniş çapta birçok cihaz veya kullanıcıya saldırıldığında hedefsiz olabilmektedir [41]. Bir saldırgan saldırıyı düzenlemeden önce ilk olarak saldırının motivasyonunu ve faydalarını belirler. İkinci olarak hedefin güvenlik açığını belirledikten sonra farklı saldırı yöntemleri ve tehditleri hakkında kapsamlı bir çalışma yapılmalıdır. İnternet kullanıcı sayısının artmasıyla birlikte siber suç faaliyetleri, sistem korsanlığı, finansal veri hırsızlığı ve kötü niyetli saldırılar ciddi derecede maddi ve manevi kayıplara yol açmaktadır. Bilgi sistemlerinin güvenliğinde bazı terminolojiler vardır. Sıfırıncı gün saldırısı, bilgi sistemindeki açığın saldırının gerçekleştiği gün ortaya çıkmasıdır. Güvenlik açığı, saldırganın hedefe erişmesi için açığı tespit etmesidir. Exploit, sistemdeki güvenlik açığının istismarla sonuçlanmasıdır. Bot, hedefte gerçekleştireceği saldırıları uzaktan kontrol etmek için geliştirilmiş yazılımlardır [42].

Yine en önemli hususlardan bir tanesi siber saldırıya maruz kalındığında meydana gelen maddi zararı en aza indirmek için kuruluşun yaptığı yatırımın ne kadar olacağıdır. Gerçek dünyada bu maliyetlerin nasıl dengeleneceği ve saldırının maliyetinin maddi olmaması durumunda maliyet fayda analizi cevap olamamaktadır. Ama maddi kayıplara yol açacak saldırılarda fayda maliyet analizi yaparak siber güvenliğe yapılan yatırımı değerlendirmek faydalı olmaktadır [43]. Siber saldırıları önlemek için çeşitli yöntemler kullanılmaktadır. Son yıllarda derin öğrenme ve sinir ağlarını kullanarak saldırı tespit sistemleri ortaya çıkmıştır. Derin öğrenme yöntemlerinin geleneksel yöntemlerden daha başarılı olduğu görülmüştür [44]. Siber saldırılarda en önemli tehditlerden birisi de içeriden gelen tehditlerdir. İçeriden gelebilecek tehdidi bilebilmek gerçekten çok zordur çünkü içerideki kişinin imkân ve kabiliyetleri siber güvenlik uzmanlarından daha iyi olabilmektedir. Bu tehditlerle başa çıkmanın en önemli yollarından biri karakter analizi yaparak daha önce meydana gelen olaylarda bireysel ve ortak özellikleri bulabilmektir [45].

2.2.1. Hizmeti Engelleme (DoS / DDoS) Saldırıları

Hizmetin engellenmesi en yıkıcı saldırılardan biridir. Saldırganın internet ortamındaki bir hizmeti geçici veya kalıcı olarak kesintiye uğratarak bilgisayarı ya da ağı kullanılmaz hale getirmesidir. Bu saldırıları başlatmanın iki yolu vardır. Birincisi saldırgan kötü amaçlı paketler göndererek hedef makinada çalışan protokolü veya uygulamayı gizlemektedir, ikincisi ise hedef makinanın tüm kaynakları ele geçirilerek kullanıcılarla iletişime geçmesi engellenmekte ve gereksiz paket trafiği oluşturulmaktadır [46]. Bu saldırı genel olarak ticaret, hükümet, bankacılık ve medya kuruluşlarının web sunucularına yapılmaktadır. Hassas bilgilerin kaybolmasına ve çalınmasına yol açmasa da büyük miktarda mali kayıplara ve saldırıyı ortadan kaldırma maliyetine yol açmaktadır. Hizmeti engelleme saldırı yöntemlerinden birkaç tanesi aşağıdaki gibi tanımlanır.

Arabellek Taşması Saldırısı (Buffer Overflow Attack): Ağ adresine büyük miktarda trafik gönderilerek gerçekleştirilen bu saldırı belirli ağların veya uygulamaların hatalarını istismar etmek için kullanılmaktadır.

ICMP Akın Saldırısı (ICMP Flood Attack): Sahte paket gönderilerek yapılan bu saldırıda yanlış yapılandırılmış ağ cihazlarına yönelik ping gönderen bir paketle hizmet engellenmektir. Ölüm pingi ve smurf saldırısı olarak da bilinmektedir.

SYN Akın Saldırısı (SYN Flood Attack): Saldırı başladığında veri alışverişinde isteklere yanıt veremeyecek duruma gelen sistemin tüm kaynaklarını kullanarak SYN paketleriyle gerçekleştirilmektedir [42].

Dağıtılmış hizmeti engelleme saldırısı, internet trafiği oluşturarak hedeflenen ağın veya sunucunun normal işleyişini bozan kötü niyetli bir girişimdir. Hizmeti engelleme saldırısından en büyük farkı tek bir yerden değil birden fazla yerden yapılan bir saldırı olmasıdır. Sağlam oluşturulan bir ağ güvenlik altyapısıyla, saldırı başladığında kullanılan araçlar ve açıklar

belirlenerek engellenebilmektedir [42]. DDoS atağında bir veya birden fazla kurbana karşı dağıtılmış saldırıyı başlatmak için zombi bilgisayarlar da kullanılabilmektedir. İlk DDoS saldırısı, 1999 yılında Minnesota üniversitesinde bilgisayar ağının iki günlüğüne kapatılmasıyla ortaya çıkmıştır. Günümüzde bu saldırıları başlatmak için daha karmaşık yöntemler kullanıldığından azaltma ve tespit tekniklerine rağmen saldırılar giderek artmaktadır. Giderek artan saldırı boyutlarıyla mücadele eden araştırmacılar, saldırıları azaltmak için etkili, sağlam ve kapsamlı bir savunma mekanizması geliştirmektedirler [46].

Bu saldırılar makineyi veya ağı yavaşlatır, ağı kapatır, verilere erişimi engeller, kaynak paylaşımını yapılmaz hale getirir, varlık kaybı yaşatır, hatalardan yararlanarak trafiği artırır ve yasal kullanıcıların erişmesini engeller. DDoS, sadece yaygın olmasının yansıra ABD Federal Araştırma Bürosuna göre mali kayıplara yol açan ikinci siber saldırı türüdür. Sınıflandırmalar ve şifreleme teknikleri gibi bu saldırıyı azaltabilecek birçok prosedür vardır. TCP üzerinden yapılan DDoS saldırısı iki farklı yolla yapılmaktadır. Bunlardan birisi kurban makinanın sistemindeki zayıflığını hedef alan doğrudan saldırılardır. Diğeri ise doğrudan kurbanı hedef almayan, makinanın ilişkili olduğu diğer unsurları hedef alarak çalışmasını engelleyen dolaylı saldırılardır [47].

2.2.2. IP Sahteciliği (IP Spoofing)

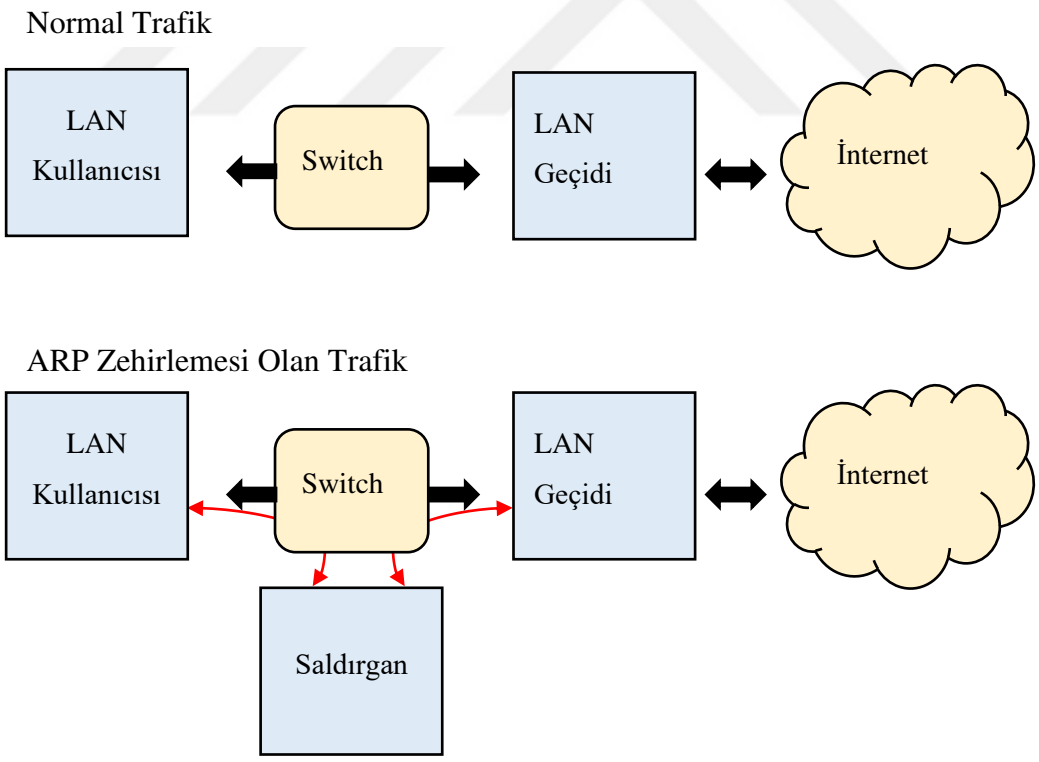
Sahtecilik kelime anlamı olarak başka biri gibi davranma anlamına gelmektedir. Bu güvenilir bir IP adresi yardımıyla uzaktaki bir sisteme yetkisiz erişmek için kullanılmaktadır. Başlamadan önce kötü niyetli kişi hedef istemcinin IP adresini alır ve TCP oturumuna sahte paketler ile kendi IP adresini enjekte eder. Kandırılan sunucu orijinal IP ile iletişim kurduğunu sandığından yetkisiz erişim elde edilir ve internet tabanlı saldırılar başlatılır [42]. IP sahtekarlığı, SQL enjeksiyonu gibi diğer sızma saldırılarıyla birleştirmede yeterlidir. Bu tür bir saldırı Endonezya genel seçimlerinde tablo sisteminde meydana gelmiştir ve bu hassas verileri koruyan kuruma karşı kamuoyu güveni ciddi derecede azalmıştır. Birçok saldırgan IP sahteciliği yaparken iz bırakmamak ve soruşturmacılara yakalanmamak için VPN veya Proxy sunucusu kullanmaktadır [48].

Sahteciliği önlemek için, sinyal gücü göstergesi ve gidiş dönüş süresi olmak üzere taklidi zor olan fiziksel ağ özellikleri kullanılarak saldırı tespit sistemleri geliştirilmiştir. Zaman serisi benzerlik ölçüsü kullanılarak sonraki ağ davranışlarının benzerliği elde edilmiştir. Sahtecilik saldırıları özellik çıkarma yöntemleriyle tespit edilerek etkisiz hale getirilebilmektedir [49].

2.2.3. ARP Zehirlemesi (ARP Poisoning)

Adres çözümleme protokolü (ARP), ağ katmanı adreslerini bağlantı katmanı adreslerine göre çözümleyen protokoldür. Bir istek ve yanıt protokolü olan ARP tek bir fiziksel ağın sınırları içindedir. ARP isteği, gönderenin MAC adresine çözmesi gereken hedefin IP adresini içermektedir.

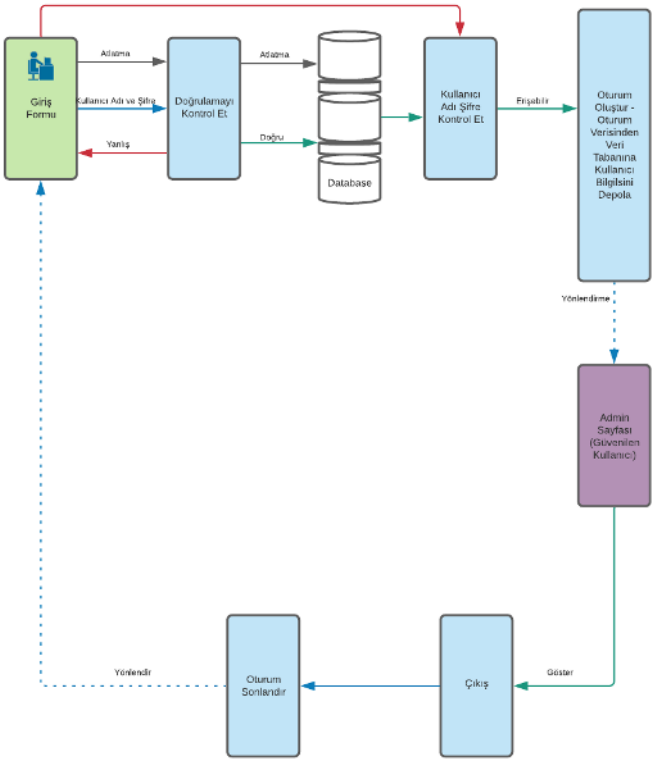
Tüm ağa yayınlanan istek, eşleştiği IP ile gönderene kendi MAC adresini içeren mesajı gönderir. ARP zehirlemesi, yerel ağa sahte bir ARP mesajı göndererek ana bilgisayarın ARP tablosunu bozar ve saldırdığı ağ trafiğini kendine doğru yönlendirir. Bu tür saldırılar saldırgan ve hedefin aynı yerel ağda olduğunda gerçekleştirilmektedir. Mağdurdan hedefe giden tüm trafik saldırgana yönlendirilecektir. Bu saldırı türünde saldırgan, aynı zamanda yönlendiriciye kurbanın IP adresiyle ilişkili kendi MAC adresi ile sahte ARP yanıt mesajı gönderdiğinde, saldırgan kurban ile internet trafiği arasında olacaktır. Veri ihlali araştırmalarına göre yerel alan ağında gerçekleşen bu saldırıların en önemli tehditlerden olduğu vurgulanmıştır. Yerel ağ ihlallerinde bunun şaşırtıcı olmamasının sebebi güvenlik açıklarının olması ve bunlardan yararlanmanın kolay olmasıdır. Özellikle Ethernet'te yaygın güvenlik açıkları adres çözümleme protokolünden kaynaklanmaktadır. Bu saldırı türüyle mücadelede güvenlik açıklarını öğrenme, statik ARP önbellek girişi, işletim sistemlerinde geliştirilmiş ARP modülü, şifreleme, erişim kontrolü, izinsiz giriş algılama ve veri yedeklemeyi içeren tedbirler gerekmektedir. Hatta çeşitli laboratuvar ortamlarında görselleştirme ve simülasyonlarla güvenlik tedbirleri geliştirilmektedir [50]. Şekil 2.3'te normal bir trafik ile ARP zehirlemesi olan bir trafik arasındaki fark anlatılmıştır.



Şekil 2.3. ARP zehirlemesi saldırısı örneği

2.2.4. SQL Enjeksiyonu (SQL Injection) Saldırısı

SQL, veri tabanındaki bilgileri depolamak, işlemek ve almak için kullanılan bir bilgisayar dilidir. SQL dili, gerekli görevi gerçekleştirmek için seçme, güncelleme ve silme gibi komutları kullanmaktadır. Ayrıca veri tabanına yönelik sorguları yürütebilir, kayıtlar ekleyebilir ve yeni tablolar oluşturabilir. SQL enjeksiyonu saldırısı arka uçta veri tabanını manipüle ederek kötü amaçlı yazılımla bilgilere erişmektedir. Bu saldırı verilerin yetkisiz görüntülenmesine, tabloların silinmesine ve veri tabanının etkisizleşmesine sebep olabilmektedir. Saldırgan bir uygulamanın formları veya domain adresi gibi bir programın veri girişleri aracılığıyla sorgu içine özel ve gereksiz bir SQL ifadesi ekleyebilmektedir. Bu komutlar görüntülenir ve uygulamalardaki açıklar yüzünden veri tabanında değişikliğe ve yetkisiz erişime neden olmaktadır [51]. PHP ve ASP uygulamaları bu saldırılardan ciddi derecede etkilenebilmektedir. Web tabanlı ASP.Net uygulamaları da bu tür saldırılara karşı savunmasızdır çünkü uç taraftaki uygulamaların yanı sıra veri tabanı prosedürlerindeki kodlama açıklıklarının bir sonucudur. SQL enjeksiyonu; Microsoft SQL sunucusu, Oracle, MySQL ve diğer veri tabanı uygulamaları için bir sorundur. Sistemde kötü amaçlı kod barındıran bu saldırı, veri tabanında görünen bilgiler ve kişiler için ciddi tehditler oluşturmaktadır [42].



Şekil 2.4. SQL enjeksiyonu saldırısını tespit ve engelleme örneği [52]

SQL saldırılarını önlemenin ve bunlara karşı koymanın birkaç yöntemi vardır. Girdi doğrulama, kod uygulaması yapabilen yasadışı kullanıcı girişlerini tanımlamak için uygulanmaktadır. Ancak yasal ve yasadışı girdilerin haritalanması mümkün olmadığından bu yöntem uygun değildir. Bu saldırıyı engellemede güvenlik duvarı, imza tanıma gibi yöntemler kullanılmaktadır. Günlük log kayıtlarını kontrol eden tespit işlemi bu saldırıyı belirlemenin en etkili yollarından biridir. SQL enjeksiyonu gerçekleşip gerçekleşmediğini bildirmektedir. Log dosyaları sunucu tarafından sağlanan önemli bilgilerdir. Zaman zaman yapılan değişiklikler, hangi işlemin kim tarafından yapıldığı ve kimin veri tabanına ne zaman eriştiği gibi bilgiler bu dosyalarda tutulmaktadır. Yöneticiler bu işlemleri kim, ne zaman, ne şekilde yaptı sorularıyla saldırganı tespit edip işlemleri geri alabilmektedirler. SQL saldırılarını önlemek için çeşitli algoritmalar önerilmiştir. Bu algoritmalar sayesinde saldırganların önceden kaydedilmiş davranışları incelenerek saldırı engellenmeye çalışılmıştır [52]. Şekil 2.4'te SQL enjeksiyonu saldırısı ve bunu önleme modeli anlatılmıştır.

2.2.5. Oltalama (Pishing) Saldırısı

Oltalama saldırısı, bilinen en etkili siber saldırılardan biri olmakla birlikte birçok kullanıcının kimlik bilgilerini elde etmeye ve veri güvenliği ihlaline sebep olmaktadır. Dolandırıcılar bu yöntemi kullanırken genellikle sahte web siteleri, hedefe giden e-postalar ve sahte telefon numaraları gibi hedefi kandırarak sanki gerçek kişiden geliyormuş gibi saldırmaktadır. Saldırının başarısı doğru adresten gelen bilgileri ne derece taklit ettiğine ve kullanıcının neyin sahte neyin gerçek olduğunu anlayamamasıyla ölçülmektedir. Saldırgan eğer kişiyi hedef almışsa, kişi hakkında sosyal medyadan veya daha önce ele geçirdiği yerlerden elde ettiği bilgileri kullanarak mağduriyet olasılığını artırabilmektedir. Eğer kişi hakkında bilgisi yoksa, kişisel bilgileri ele geçirebileceği banka veya başka bir gerçek site görünümünde olabilmektedir [53]. Bu saldırının benzetmesi kurbanlar için balık tutmaktan türetilmiştir. İnternette bulunan popüler ve yasal birçok web sitesinin sahtesi oluşturularak, kişiler milyonlarca zarara uğratılmıştır. Bu sitelerin tespit edilmesinde kullanıcının dikkati çok önemlidir. E-posta veya sitenin URL adresine dikkatli bakan kullanıcı aslında adresin gerçek siteye ait olmadığını fark edecek ve saldırıdan kurtulacaktır. Yapılan bir çalışmada, kullanıcıların adres hakkında ayrıntılı bilgiye sahip olmamaları, hangi web sitesine güveneceğini bilmemesi, herhangi bir siteye girdiğinde yönlendirmeye maruz kalıp kalmadığını bilmemesi, adresi bulmak için fazla zamanı olmadığından rastgele girmesi ve sahte web adresinin ayrımını yapamamasından kaynaklandığı tespit edilmiştir. 2017 yılında yapılan çalışmada, oltalama saldırılarına maruz kalmada %47 ile Çin, ardından %42,8 ile Türkiye ve %39,9 ile Tayvan'ın takip ettiği görülmüştür [54]. Oltalama saldırılarının hedef kitlesini tespit etmek ve önleyebilmek için kişilere yönelik farkındalığa odaklanan çeşitli çalışmalar yapılmıştır. Yapılan çalışmalarda kullanıcıların temel rol oynadığı çok çeşitli metodolojik yaklaşımlar ve potansiyel

saldırı öznitelikleri dahil olmak üzere kullanıcı merkezli olması saldırıyı anlamaya olanak sağlamaktadır. Yapılan çalışmalar, kullanıcı ne kadar sofistike olursa olsun yine de sahte web sitelerinde yanılabileceğini göstermiştir. Saldırılarda standart ve geleneksel önlemlerin ötesinde daha etkili tedbirlerin alınması gerektiği vurgulanmıştır [53]. Son dönemlerde oltalama saldırılarının artmasıyla, sistemlerin ve kullanıcıların korunması için bazı ilave destek sistemlere ihtiyaç duyulmuştur. Bunlar kullanıcı için karar desteği sağlayan yazılım tabanlı oltalama tespit sistemi ve makine öğrenme yöntemleridir [54].

2.2.6. Gizli Dinleme Saldırısı

Gizli dinleme saldırıları, bir ağ üzerinden tüm veri paketlerini gözlemleme ve ele geçirme yöntemidir. Hedef iletişimini sertifikasız olarak tarayan bu yöntemde saldırgan, hesap bilgileri, kredi kartı bilgileri ve parolalar gibi hassas veri paketlerini ele geçirmektedir. Kurban verilerinin çalındığını veya tahrif edildiğini anlamakta zorlanır çünkü bu süreç normal kullanım ve veri iletim yöntemini nadiren etkilemektedir. Örnek olarak kullanıcıya bir eposta geldiğinde, eğer eposta dijital imza kullanılarak şifrelenmemişse ortadaki adam saldırısı kullanılarak bu bilgiler ele geçirilmektedir. Saldırgan, mesajı uygun şekilde değiştirebilir, alıcı alınan mesajın meşru gönderenden geldiğine inanır ve kolayca bilgiler ele geçirilir. Gizli dinlemeyi önlemek için iki yöntem vardır. Birincisi, güvenlik konusunda daha iyi ve şifreli bağlantı yöntemi olan SSH ve HTTPS protokolleri kullanılmalıdır. İkincisi ise güncellemeleri tam ve eksiksiz olan kişisel güvenlik duvarı, anti virüs programı ve açık anahtar altyapısına sahip karşılıklı kimlik doğrulamasına sahip internet hizmetleri kullanılmalıdır [42].

2.2.7. Kötü Niyetli Yazılımlar

Kullanıcının kişisel bilgisayarına veya sistemine bilgisi olmadan virüs, casus veya fidye yazılımların yüklendiği siber saldırıdır. Kötü amaçlı kod, yasal koda eklenir, yayılır ya da kodun kendisi tarafından yürütülmektedir. Bu yazılımlarla ağa erişilerek veya sistemi kesintiye uğratılarak hassas bilgileri ve diğer kişisel bilgileri ele geçirildikten sonra bundan menfaat sağlanmaktadır. Günümüzde bu yazılımlar kişisel hedeflerden çok ticari ve finansal hedeflere yönelmektedir. En yaygın olan kötü amaçlı yazılımlar şunlardır.

Virüs; herhangi bir sisteme dosya indirirken veya bir program çalıştırıldığında eklenen, çalıştırıldığında kodları kopyalayan, değiştirebilen yazılımlardır.

Solucan; e-posta ekleri aracılığıyla bilgisayara veya ağa yayılan hizmetin reddi saldırısına sebep olarak sistemi durduran veya yavaşlatan yazılımlardır.

Truva atı; virüs gibi çoğalmadan, yararlı bir yazılımın içinde saklanarak saldırgana hassas bilgileri göndermeye yarayan çok tehlikeli bir saldırı türüdür.

Fidye yazılımı; basit bir kodu olan bu yazılım türünde kullanıcının verileri kilitlenir ve fidye ödenmediği takdirde verilere erişim sağlanamaz. Bu saldırıyı önlemek gerçekten çok zordur.

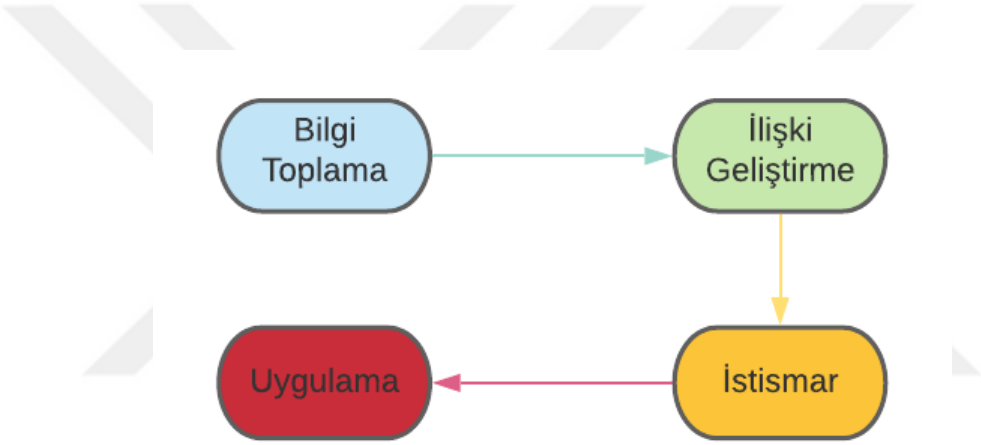
Casus yazılım; kullanıcının yaptığı tüm işleri bilgisi olmadan inceleyen ve saldırgana raporlayan kötü niyetli bir saldırı yazılımıdır [41].

Kötü amaçlı yazılımların bulaştığı bir bilgisayar çok sayıda bilgisayara bulaşmayı hedeflemekte ve başka kurbanlar bulmaktadır. Bu saldırıları önlemek için yazılımlar iyi anlaşılmalıdır. Ayrıca bu saldırıları tespit edebilmek için çeşitli analiz yöntemleri kullanılmalıdır. Özellikle gereksiz e-postalarla birlikte incelenerek nasıl bulaştığı ve yayıldığı anlaşıldıktan sonra saldırı azaltılabilmektedir [55].

2.2.8. Sosyal Mühendislik Saldırısı

Sosyal mühendislik, her kuruluşta çalışanların insani zayıflıklarından yararlanarak bir hedef hakkında bilgi toplamanın en basit yöntemi olarak tanımlanmaktadır. İnsan hedeflerin kasıtlı olarak manipüle edilerek aldatıcı tekniklerin tasarlanması ve uygulanmasıdır. Saldırgan sistemdeki en zayıf halkayı bulduktan sonra ondan aldığı bilgilerle, güvenlik politikalarını ihlal etmekte ve sisteme yetkisiz bir şekilde erişmektedir. Zayıflığı kullanılan kişi etkileşim boyunca bundan habersizdir ve yıkıcı eylem gerçekleştiğinde olayın ne olduğunu anlayabilmektedir. Yetenekli ve zeki saldırgan insan zihninin psikolojik yönlerini manipüle ettiğinden, uzun zaman ve emek harcayarak yapacağı başka bir saldırı yöntemine gerek kalmadan kolayca yapmak istediği saldırıyı gerçekleştirmektedir. Sosyal mühendislik saldırısını kullanan gruplar genellikle zeki insanlardır. Saldırgan mağdurun zaafından yararlanarak sistemine girer ve kişisel bilgilerini çalarak bunu menfaate dönüştürmektedir. Genç bir hacker olan dünyanın en ünlü bilgisayar korsanı olan Kevin Mitnick, teknik becerilerinden çok kurnazlığı ve ikna ediciliğiyle bilgisayar ağlarını ihlal ettiği için hapse atılmıştır. Herhangi bir bilgisayar sistemi teknik olarak ne kadar güvenli olursa olsun sistemdeki kişiler en büyük güvenlik açığı olarak görülmektedir [56]. Siber saldırıların büyümesi büyük ölçüde sosyal mühendislik saldırılarındaki artışla karakterize edilmiştir. Sosyal mühendislik saldırılarının doğası karmaşıktır çünkü saldırganlar otomatik olarak kolayca güvence altına alınamayan insan zaaflarından faydalanmaktadır. Saldırgan, saldırı sonucunda elde edilebilecek hassas veri miktarı göz önünde bulundurulduğunda hedefini çeşitli faktörlere göre seçmektedir. Büyük müşteri tabanına sahip kurumlar, telekomünikasyon şirketleri ve finansal kuruluşlar saldırı için hedef olma eğilimindedir. Özellikle işyerlerinde sosyal medya kullanımındaki artış; insan, süreç ve teknoloji açısından siber güvenlik uzmanları için ciddi zorluklar oluşturmaktadır. Saldırgan dijital bir saldırı yapması durumunda, dijital alan sayesinde suiistimal gibi kimlik avı e-postaları göndererek kötü niyetli girişimi otomatikleştirebilmekte ve saldırı sürecinin maliyetini düşürmektedir. Bu saldırılar bireysel zayıflık, mağdur profili ve ilişkili güvenlik açığı gibi çeşitli çalışmalarla incelenmiştir [57].

Sosyal mühendislik saldırılarının, teknik saldırılardan farklı olmasının sebebi hedef kitlenin düzenli çalışma saatlerinden sonra çalışan temizlik görevlilerinden yöneticilere kadar olmasıdır. Hedef personelin teknik yeterliği olmadığında saldırı endişesi yaşamayabilmektedir. İnsan faktörünün olduğu her yerde yapılan birçok çalışmada, bilgisayar kullanıcılarına yönelik bu saldırıları önlemenin en etkili yollarından birisi personel farkındalığının artırılmasıdır. Çeşitli çalışmalar ve akademik ortamlarda, bilgi varlıklarını korumak, iş sürekliliğini sağlamak ve kendi bilgi düzeyini geliştirmek amacıyla farkındalık programları uygulanması önerilmiştir. Araştırmacılar güvenin, bilgi güvenliği sisteminin her yönü için hayati önem taşıdığını ve güvenliği önemli derecede etkilediğini savunmaktadır. Hükümetlerde zamanla sosyal mühendislik saldırısı konusunda farklı yasa ve yönetmelikler çıkararak yeni yöntemler aramaktadırlar [57]. Şekil 2.5'te sosyal mühendisliğin 4 aşaması anlatılmıştır.



Şekil 2.5. Sosyal mühendisliğin dört aşaması [57]

Sosyal mühendislik saldırılarıyla başa çıkmada katmanlı bir güvenlik yapısı oluşturmaya ve saldırıların felakete dönüşmesini engelleyecek bir risk azaltmaya ihtiyaç vardır. Her kuruluşun çok iyi yazılmış bir güvenlik politikası olmalı ve bu en üst yönetimden en alta kadar uygulatılarak teknik ve teknik olmayan yaklaşımları içermelidir. Kurum veya kuruluşta çalışmaya başlayan herkes ilk olarak siber farkındalık eğitimine tabi tutularak kurallar anlatılmalıdır. Kullanıcıların şifrelerini belirli periyotlarda değiştirmesi sağlanmalı ve beyaz listede olmayan ağ adreslerine erişim engellenmelidir. Güvenlik politikalarına uyulup uyulmadığı denetlenmeli, ağ günlükleri incelenmeli, çalışanların yetkilerini yeniden düzenlenmeli ve kullanıcıların masaüstü yapılandırmalarının kontrol edilmesi sağlanmalıdır. Ağ, veri ve temel altyapıyı korumak için saldırı tespit sistemleri ve güvenlik duvarı kullanılmalıdır. Web filtresi ve sanal özel ağ kullanılmalıdır. Kimlik doğrulama, biyometrik giriş ve erişim kontrolü gibi fiziksel tedbirler alınarak çok faktörlü kimlik doğrulaması yapılmalıdır [58].

2.3. Kritik Sektörlerde Siber Güvenlik

Günlük hayatımızı destekleyen kritik altyapılardaki teknoloji olmadan yaşamanın bir hayal olduğu şu günlerde, kötü amaçlarını gerçekleştirmeye çalışan saldırganlar kişisel, kurumsal ve ticari verilere erişmek için çeşitli denemeler yapmaktadır. Birçok vatandaş bu saldırıların kurbanı olmakta ve kritik altyapılar bu saldırılara maruz kaldığında sonuçlar çok daha ağırlaşmaktadır [59]. Modern toplumun ilerleyişiyle devletlerin temel işlevi hem yönetişim hem de mal ve hizmet sağlayan (ulaşım, elektrik, iletişim vb.) altyapılarının kullanımıdır. Kritik altyapı denildiğinde akla bireysel unsurlardan oluşan alt sistemler ve sektörler gelmektedir. Sonuç olarak bu altyapılar hem iç hem de dış tehditlere karşı sağlam olmalı ve dayanıklılığın geliştirilmesi de sürekli devam etmelidir. Bu dayanıklılık saldırıya hazırlık, etki azaltma önlemleri, müdahale ve kurtarma yetenekleri gibi çeşitli yöntemlerle sağlanabilmektedir [60].

Son yıllarda kritik altyapılardaki siber saldırıların artması, açık standartlara ve basit web teknolojilerine geçişin sürdürülemez olduğunu göstermiştir. Özellikle siber uzaydaki saldırıların istihbarat amaçlı hükümet faaliyetleriyle sınırlı olmayıp aynı zamanda bankacılık ve hizmet sektöründeki malların taşınmasına da sebep olmaktadır. Bu alandaki saldırıları analiz etmedeki en büyük sorun saldırıların fark edilmemesidir. Hatta bazı kuruluşlar bu olayları bildirme konusunda isteksizdir çünkü bunları utanç verici olarak görmektedir. Kritik altyapılar genellikle büyük kamu yatırımlarıdır. Bu sistemdeki en ufak bir aksamanın ekonomik kayıplar gibi ciddi etkileri olabilmektedir. Hükümetler kritik altyapıları belirlemeli ve ulusal öneme sahip sektörlere öncelik vermelidir. Kritik altyapılara yapılan saldırılar genellikle hedefli ve hedefli olmayan olmak üzere iki şekildedir. Hedefli olmayan saldırı, internete bağlı olan herhangi bir bilgisayarın maruz kalabileceği aynı saldırıların neden olduğu olaylardır. Hedefli saldırı ise saldırıya uğrayan sistemin kontrolü altındaki fiziksel sisteme zarar verecek şekilde tasarlanmaktadır. Hedefli saldırılar çok tehlikelidir çünkü bunların büyük bir kesinti etkisinin olması beklenmektedir [61].

Kritik altyapıların korunması sadece ülkelerin tek başlarına bir sorunu değil aslında uluslararası bir problemdir. Kritik bilgi altyapılarının güvenliğini sağlamaya yönelik görüşlerin bir araya getirilmesi önemlidir. Çünkü bir ülkede olan bir kritik altyapının diğer ülkelere de etkileri olabilmektedir. Örneğin, nükleer bir santrale düzenlenen siber saldırı sonrası meydana gelen patlamanın etkisi birçok ülkeyi de etkileyebilmektedir. Yine birlik kurmuş ülkelerden birine yapılan saldırı diğer ülkeleri de yakından ilgilendirecektir. Kritik altyapıları tamamen izole ederek bilgisayar sistemlerinden ve insanlardan uzaklaştırmak mümkün değildir. İnternete doğrudan bağlantı olmasa bile, modern kontrol sistemlerine ve diğer kritik altyapı nesnelerine dışarıdan erişilebilmektedir. Uzaktan bakım ve teşhis bağlantıları, iş kullanıcılarıyla paylaşılan yürütme sistemleri, uzaktan erişim modemleri, seri bağlantılar, kablosuz bağlantılar, mobil bilgisayarlar, usb cihazları ve veri dosyaları gibi birçok etken vardır [62].

Şekil 2.6. Kritik altyapı sektörleri

Şekil 2.6'da gösterilen kimyasal, kritik imalat, taşıma sistemleri, sağlık, bilgi teknolojileri, finans, acil servisler, enerji, ticari tesisler, nükleer santraller, hükümet tesisleri, telekomünikasyon sistemleri, barajlar, yiyecek ve tarım alanları, su ve sulama sistemleri ile savunma sanayi üsleri hayati öneme sahip kritik altyapılardandır [63]. Kritik altyapıların güvenliği sağlanırken krize hazırlık düzeyi, saldırı tespit kabiliyeti, saldırıya yanıt yeteneği, fiziksel direnç, teknik araçlar, organizasyonel önlemler ve saldırı anında kurtarılması gerekenler belirlenmelidir. Yine en önemli konular arasında risk yönetimi, yenilikçilik, eğitim, geliştirme ve siber farkındalık faktörleri bulunmaktadır. Her bir kritik altyapının kendi özelinde saldırı tespit sistemleri ve saldırı anında yapılacaklar planı bulunmalıdır [60].

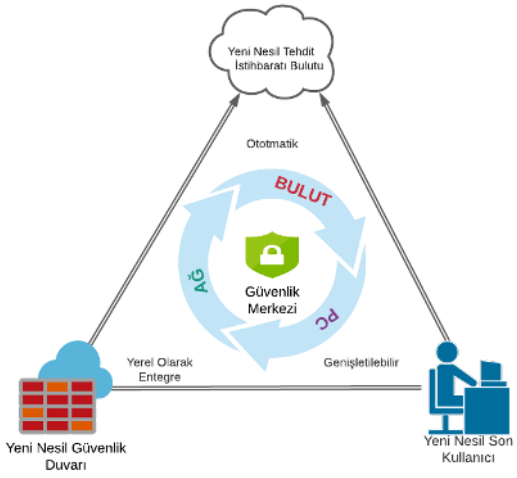
2.3.1. Kişisel Bilgisayar ve Ağ Sistemleri

Siber uzayda her zaman ve her yerden karmaşık, öngörülemeyen ve nerden çıktığı belirlenemeyen saldırılar gelebilmektedir. Çeşitli saldırı tespit sistemleri, güvenlik duvarları, virüs yazılımları, açık taramaları ve olay taramaları gibi birçok yöntem ile savunma geliştirilmiştir. Bu yöntemlerle siber güvenlik analistleri, bilgisayarları ve ağları korumak için tehditlere karşı zamanında ve etkili eylemlerle karşı koymalıdır [64]. Siber güvenlik analistleri bilgisayar kullanıcılarının günlük işlemlerini dikkatlice incelemeli ve bu işlemlere göre savunma politikası geliştirmelidir. Kullanıcıların eylemleri neticesinde, güvenlik mekanizmalarının ürettiği uyarılara göre kötü amaçlı yazılım tespiti, kullanıcının güvenlik politikalarına uyup uymadığı gibi hususlar tespit edilebilmektedir. Bu tespitlere göre çeşitli farkındalık eğitimleri, bilgisayarı nasıl kullanacağı ve ağ sistemine gelebilecek zararları nasıl engelleyebileceği anlatılmalıdır [2].

Siber güvenlik, verilerin kendisini, ağ ortamında iletilen verileri ve bu sistemlere yetkisiz erişimleri koruyarak sağlanmaktadır. Herhangi bir kurum ya da kuruluşun elektronik iletişim ağını ve bilgisayarlarının belirli güvenlik standartlarını karşılaması gerektiğini anlatan politikalar vardır. Bu politikalara uymak hayati önem taşır çünkü herhangi bir bilgisayara bulaşan virüs tüm sisteme zarar verebilmektedir. Yani çalışanlara eğitim verilirken ilk olarak kişisel bilgisayarların korunması

gerektiği anlatılmalıdır. Sonuç olarak bilgisayardaki bir tehlike tüm ağa erişimi engelleyebilir, verilerin ifşa olmasına, sistemin aksamasına, kurumun kamuoyuna karşı güveninin sarsılmasına kadar birçok zararı olabilmektedir [65]. Siber casuslukların son yıllardaki artışı ile milyarlarca dolar zararlar ortaya çıkmaktadır. Özellikle ağ ve bilgisayar sistemlerini korumada güvenlik fonksiyonlarını ve kontrollerini iyi belirlemek, kullanılan yazılımları denetlemek ve iş birliği modellerini iyi oluşturmak gerekmektedir [66].

Son yıllarda ağ ve bulut altyapıların artmasıyla birlikte güvenliğin sağlanması daha karmaşık hale gelmiştir. 2025 yılına kadar dağıtılan uygulamaların, teknolojilerin ve hizmetleri içeren tüm altyapıların yüzde seksenin bulut tabanlı olacağı tahmin edilmektedir. Mevcut durumda bulut altyapılarının depolama kapasitesi, ağ verimliliği ve donanım ile ilgili talepler gelişirken güvenliğini yöneten hususlar geleneksel yöntemlerle sağlanmaktadır. Bu geleneksel yöntemlerin birincisi olan güvenlik açığını yama ile kapatmak zaman ve karmaşıklık içerdiğinden zayıf kalmaktadır. İkincisi müşteriye sağlanan hizmetin engellenmesi bulut hizmet sağlayıcıların mevcut sistemlerde değişiklik yapmasını engellemektedir. Son olarak saldırganın keşif aşamasında elde ettiği bilgileri kullanarak geliştirdiği sıfır gün saldırısı geleneksel yöntemleri savunmasız hale getirebilmektedir. Bu savunma yöntemlerini geliştirmek için düşmanlara karşı olaylar ortaya çıkmadan engelleyen hareketli hedef savunmaları gibi çözümler ortaya çıkmaktadır [67].



Şekil 2.7. Ağ ve bilgisayar güvenliği şematik diyagramı

Şekil 2.7'de yeni nesil ağların güvenliği gösterilmiştir. Güvenlik ölçüleri, bir sistemin veya ağın güvenlik düzeyini hem nitel hem de nicel yollarla sunmaktadır. Güvenlik analizini yapmak için siber güvenlik ölçütlerini sınıflandırmak, güvenlik ölçüleri geliştirmek, bu ölçüleri resmi olarak tanımak ve personelin bu ölçülere uymasını sağlamaktır. Kullanıcı parolası, yazılım güvenlik açıkları ve kullanılan şifreleme anahtarları ile sistemlerin güvenlik açıklarını ölçmektedir. Kara

listeye alma, saldırı tespit yeteneği, yazılım çeşitlendirme ve karşı önlemlerin genel etkinliği kuruluştaki savunmayı ölçmektedir. Tehditleri ölçmeye yönelik metrikler, sıfır gün saldırısı tehdidi, bireysel saldırıların gücü ve şaşırtmanın karmaşıklığı aracılığıyla kuruluşa yönelik tehditleri ölçmeyi amaçlamaktadır. Kullanılan metrikler, güvenlik yatırımı, güvenlik durumu ve güvenlik olayları mevcut durum değerlendirmesidir. Bu sınıflandırma kurumsal sistemlerdeki saldırganlar ve savunanlar arasındaki perspektife odaklanmaktadır. Bu metriklerle birlikte bilgisayarlar ve ağlar daha güvenli hale getirilebilmektedir [68].

2.3.2. Akıllı Sistemler

Teknoloji insanların hayatlarını kolaylaştırdıkça başka sorunlar ortaya çıkmaktadır. Artık oturduğu yerden kontrol ettiği sistemlere de siber saldırılar düzenlenmekte ve istenmeyen zararlar verilmektedir. Bununla birlikte internete bağlı objeler adı verilen nesnelerin interneti gibi kavramların ortaya çıkmasıyla birlikte yeni güvenlik risklerini de beraberinde getirmiştir. Mesela kapı kilitleri, termostatlar, televizyonlar, buzdolapları, sağlık cihazları, trafik ışıkları ve araba yıkama makinaları gibi akıllı cihazlardan sahiplerinin hassas verilerini çalan veya müdahale eden saldırganlar vardır [69]. Akıllı şebekelerin artmasıyla birlikte büyük ekonomik kayıplara ve çevresel kaygılara yol açan saldırılar ortaya çıkmıştır. Akıllı şebekelerdeki güvenlik saldırılarını, güvenlik ve gizlilik için gerekli olan güvenlik parametrelerini, güvenli veri alışverişini sağlayan iletişim protokolü belirlemek, enerji yönlendiricisindeki güvenlik yeteneklerini geliştirmek ve güvenlik açığını ölçeğe göre analiz etmek ve çözmek gibi kritik güvenlik tedbirlerini almak hayati önemlidir. Akıllı şebekenin kullanılabilirlik, bütünlük, gizlilik, kimlik doğrulama, yetkilendirme ve inkâr edememe gereksinimlerini sağlamak için yeni teknikler geliştirilmelidir [70].

Şekil 2.8. Akıllı sistemler

Günümüzde geleneksel araçların yerini daha modern ve akıllı sistemlerle yönetilen araçlar almıştır. Bu araçların sürekli artan teknolojik işlevi sürücü için kolaylık olsa da daha fazla güvenlik açıklarına neden olmaktadır. Saldırganların araca ve yolculara zarar verebilecekleri gibi sistemi kötüye kullanarak iletilen hassas bilgileri çalabilmektedirler. Saldırgan herhangi bir zamanda aracı durdurabileceği gibi, hız göstergesinden ışık sistemine kadar müdahale edip araç sahibine zarar verebilecektir. Bilgi-eğlence sistemleri, akıllı telefona arabadan bağlanma, kablosuz veya kızılötesi bağlantı gibi araçtaki kablosuz bağlantıları araştırmak gereklidir. Şifreleme algoritmaları ve gelişmiş saldırı tespit sistemleri bu modern araçlara zararı büyük ölçüde engelleyebilmektedir [71]. Şekil 2.8'de gösterildiği gibi kontrol cihazlarıyla akıllı evler, araçlar, yangın sensörleri, akıllı cihazlar, elektrik ve sulama sistemleri, kapı kilit sistemleri, ısı ve sıcaklık sistemleri uzaktan kontrol edilebilmektedir.

Robot terimi genel olarak otomatik görevleri daha hızlı, hassas ve verimli bir şekilde tekrar tekrar gerçekleştirmek için tasarlanmış insan dışı bir bileşen olarak tanımlanabilir. Robotik teknolojiler sadece depolama ve imalat sektörleri için değil aynı zamanda savunma, çiftçilik, hastaneler, ofisler ve hatta okullar gibi endüstriyel olmayan alanlarda da kullanılmaktadır. Robotik yöntemlerin en gelişmişi bilişsel öğrenme, analitik ve gelişmiş algoritmaların kullanılmasıyla insana benzer makine öğrenimi ve yapay zekâ içermektedir. Yine robotik sistemler nesnelerin interneti ile birleştirildiğinde çeşitli güvenlik açıklarından dolayı siber saldırılara maruz kalarak kontrolleri başkalarının eline geçebilmektedir. Robotların içinde bulunan tutucular, motorlar, dişliler, tekerlekler veya pistonlar gibi mekanik parçalar robotların öğeleri hareket ettirmesini, yakalamasını ve kaldırmasını sağladığından bu kontrolleri ele geçiren saldırganlar ciddi tehditler oluşturmaktadır. Bu riskleri tespit etmek, azaltmak robotik bir sistemin her açıdan güvenli olması için siber güvenliğinin de sağlanmasıyla mümkündür [72]. Akıllı sistemlerdeki yazılımların hızlı gelişmesi ve yeni işlevlerinin piyasaya sürülmesinden dolayı hiçbiri hatasız değildir. Özellikle bu yazılımlarda sızma algılama, kimlik doğrulama gibi çözümler gerektiği gibi kimlik avı, hizmet reddi ve virüs saldırılarına karşı güvenlik seviyelerini artırmak da gerekmektedir. Sektörde yazılım geliştirme politikalarına ağılık vermek yerinde ve faydalıdır [68].

2.3.3. Finans Sistemleri

Dijital ekonomi kavramı farklı ve bağımsız olarak geliştirilmiş kullanıma hazır dijital teknolojilerin entegrasyonu ve eşzamanlı olarak uygulanmasına dayanmaktadır. Dijital teknolojiler bilgi, hesaplama, iletişim ve bağlantı teknolojilerinin kombinasyonları olarak tanımlanmaktadır. Bilgi işlem, depolama, bant genişliği ve yazılım uygulamalarının fiyat-performans kapasitesindeki gelişmelerinin yeni nesil teknolojileri yönlendirmektedir. Nesnelerin interneti, büyük veri, yapay zekâ ve sanal gerçeklik kavramlarıyla hareket eden dijital ekonomi bankacılık ve finans sistemlerinin bir vazgeçilmezi olmuştur [73].

Bankacılık ve finans endüstrisi, para yönetimi için ürün ve hizmetlere odaklanan yani kredi kartı veren, ödeme işlemlerini gerçekleştiren, sigorta işlemleri yapan, menkul kıymet satışına aracılık eden, yatırım fonları ve kredi sağlayan çok çeşitli kurumlardır. Sektör bilgi teknolojilerini kullanarak parayı yönettiği için sürekli siber tehditler altındadır. Yetenekli siber suçlular hırsızlık ve dolandırıcılık yapmak için giderek daha organize ve sofistike yaklaşımlarla finans sektörünün endişesi haline gelmiştir. Bankacılık sektörü siber güvenlik politikalarına uyum konusunda en hassas kurumlardan biridir. Yeni çıkan tüm standartlar uygulanır ve buna rağmen hala saldırılara karşı kendini geliştirmek zorundadır. Finans sektöründe uygulanan politikalar olası kritik faaliyetleri izlerken, müşteriyi korumak ve zararı engellemeye yönelik çözümler üretmektedir [30]. Çok sayıda finansal hizmet uluslararası ticaret, e-ticaret ve mobil ödeme gibi web hizmetleriyle müşterilerinin işini kolaylaştırmaktadır. Bu çözümlerin muazzam fayda ve kolaylıklarına rağmen kötü niyetli saldırılar, veri ihlalleri ve ağa sızma gibi saldırılara maruz kalmaktadırlar [74].

Finansal hassas verilerin depolanması bulut bilişim, büyük veri ve nesnelerin interneti gibi birçok yeni teknolojinin uygulanmasını mümkün kılmıştır. Özellikle finans sektöründe siber saldırıları önlemenin en etkili yollarından biri siber farkındalıktır. Yine siber riskleri sınıflandırmak, anlamak, kurum-varlık ilişkilerinin yüksek karmaşıklığını çözmek zordur. Fiziksel altyapı hasarları, veri aktarım katmanında ve veri tabanı katmanında saldırılara uğrama şansını artıracaktır [75]. Siber suçların maliyeti her geçen gün artmakta ve birçoğu tespit edilememektedir. Bazı finans ve banka kurumları riski ve siber dolandırıcılıkları azaltmak için blockchain tabanlı teknolojiyi kullanmaktadır. Örneğin Nasdaq, hisse senedi yönetim yeteneklerini artırmaya yardımcı olması için blockchain tabanlı dijital defter teknolojisini başlatacağını duyurmuştur. Bu sistem nesnelerin interneti ve akıllı sistemlerin geliştirilmesinde önemli rol oynamaktadır çünkü alışverişi yapılan bir veriyi takip ederek bireysel cihazların geçmişini takip etmektedir. Bu teknolojiyle cihazların maliyeti düşürülmekte, bilgi teknolojilerinin gücü artırılmakta ve güvenlik konusunda muazzam bir olasılık sunulmaktadır [76].

Özellikle Amerika'da kritik altyapıları korumak ve siber güvenliğinin sağlanmasında fayda sağlamak amacıyla çeşitli kurumlar kurulmuştur. Bu kurumların genel amacı politika oluşturmak ve hükümet finans sektörü arasında yakın çalışma ilişkileri sağlamak olmuştur. Bu kurum güvenli telekomünikasyon ekipmanı, şifreli eposta erişimi, güvenlik açıklarını ele alarak değerlendirme, ödeme ve sigorta sistemi gibi kritik konularda kararlar alarak kritik altyapının korunması için ulusal stratejiler geliştirmiştir [77]. Siber saldırıların genel amacının maddi menfaat elde etmek olduğunu düşünülürse korunması gereken en önemli sektör finanstır. Dünya ekonomi forumu son zamanlardaki büyük ölçekli siber güvenlik ihlallerini, dünyanın karşı karşıya olduğu en ciddi beş riskten biri olarak değerlendirmektedir. Tahminlere göre özellikle finans sektöründe siber güvenlik maliyetinin çok ciddi oranlarda artacağı düşünülmektedir. Bu sebeplerle siber saldırıların tespit edilmesi ve önlenmesi zor olduğundan daha bütünsel yaklaşımlara ihtiyaç vardır. Siber saldırıları

önlemek için öncelikle bankacılık personelinin farkındalığı artırılmalı ve yetkisiz kullanımı, yanlış verileri, etkisiz işlemleri, yanlış algoritmaları ve hatalı sistem girdilerini tespit edebilecek güvenlik kontrolleri sağlanmalıdır [73].

2.3.4. Diğer Kritik Sistemler

Siber dünyada en önemli tehdit kritik altyapıları korumaya odaklanmaktadır. Kritik altyapılar toplumun kesintisiz işleyişi için hayatı öneme sahip yapıları ve onları kesintisiz sürdürmeyi kapsamaktadır. Ülkelerin kritik altyapıları koruma düzeyleri farklıdır fakat kritik altyapıları birbirine benzerdir [78]. Siber tehditler, bu altyapıların artan karmaşıklığından ve uzaktan erişimlerinden yararlanarak, ulusun güvenliğini, ekonomisini ve sağlığını risk altına sokmaktadır. Hatta bu saldırılar bazı kuruluşların itibarını zedeleyerek yenilik yapma ve müşteri kazanma potansiyelini bile etkileyebilmektedir [79]. Kritik altyapılardan biri olan elektrik şebekelerinin bilgisayar sistemlerine dahil olmasıyla birlikte çok sayıda güvenlik açığını da ortaya çıkarmıştır. Örneğin, 2015 yılında Ukrayna'nın elektrik şebekesine bir siber saldırı düzenlenmiştir. Saldırıdan birkaç ay önce oltalama saldırısı ile kötü niyetli yazılımlar sisteme yüklenmiş ve saldırı adımlarını planlamak için şebekenin operasyonları bir süre izlenmiştir. Saldırı gününde elektrik doğrudan kesilmiş ve saldırı 225000 müşteriyi etkilemiştir [80]. Birçok elektrik şebekesinde, elektrik enerjisi üretiminin, iletimi ile dağıtımının izlenmesi ve yönetilmesi için Modern Denetleyici Kontrol ve Veri Toplama (SCADA) sistemleri kullanılmaktadır. SCADA sistemlerinin güvenliğini güçlendirmek için korunan varlıkların riskini azaltmak ve güvenlik kontrollerini katmanlayarak derinlemesine savunmadır [81].

Yeni inşa edilmiş nükleer tesisler de siber güvenliğin sağlanması gereken kritik alanlardandır. Son yıllarda nükleer santrallerin de bilgisayar sistemleriyle entegre edilmesinden dolayı siber tehditler artmış ve bütünlük tehdit edilmiştir. Geçmiş yıllarda nükleer santrallere gerçekleştirilen saldırılardan dolayı birçok ülke ciddi zararlar görmüştür. Ülkeler bu saldırılardan sonra risk değerlendirme, varlık analizi, sistem analizi, tehdit analizi, güvenlik açığı analizi, risk analizi ve izinsiz giriş tespitleri gibi bir dizi önlemler almışlardır [82].

Kimyasal madde depolama ve üretim alanları ile bunlarla ilişkili altyapılar giderek artan bir güvenlik endişesi haline gelmektedir. Meydana gelen olaylar ve endişeler göz önünde bulundurulduğunda güvenlik politikaları oluşturarak, güvenlik olaylarına dikkat çekmek ve raporlama sistemi başlatmak bir gerekliliktir [83]. Dağıtım şebekeleri, atık su arıtma tesisleri gibi kritik altyapılarda güvenliği sağlamak ve verimliliğini artırmak için siber fiziksel sistemler kullanılmaktadır. Bu sistemler çalıştırılırken her an siber saldırıya maruz kalabileceği düşünülerek prosedürler oluşturulmalıdır [84].

Ulaşım, petrol, gaz ve acil durum hizmetleri de siber saldırganların hedefinde olan endüstriyel ve kamusal altyapılardır. Teknik olup olmadığını geniş bir yelpazede analiz eden

kapsamlı araçlar ve yeni yöntemlerle, bu altyapıların siber yönlerini tespit etmeye ihtiyaç vardır. Yapılan araştırmalarda kritik sektörlere karşı siber suç faaliyetlerinin arttığı görülmüştür. Bu alanlarda tespit edilen güvenlik açıkları sadece sektörleri değil vatandaşları ve toplumu da doğrudan etkilemektedir [85].

Kamuda ve özel sektörde siber güvenlik için güvenlik yazılımları ve donanım cihazlarına milyonlarca dolar harcansa da hala güvenlik zafiyetleri tam olarak önlenememiştir. Ülkeler kritik altyapıların korunmasında diğer ülkelerle iş birliği yapmaktadır ve bu güvenlik açısından çok önemli bir konudur. Genellikle kritik altyapılarda dört ana unsur olan gerçek zamanlı izleme, anomali tespiti, etki analizi ve saldırı azaltma stratejileri kullanılmaktadır. Kritik altyapıların güvenliğini sağlamak için kullanılan yönetim modülü Şekil 2.9'da gösterilmiştir [86]. Bu konudaki güvenlik gereksinimlerini, ağ açıklarını, saldırılara karşı önlemleri, güvenli iletişim protokollerini ve mimarilerini gözden geçirmek önemlidir. Özellikle bu altyapıların güvenliğini sağlamak oldukça zordur. Veri toplama protokolleri ile güvenlik sağlanmaya çalışılsa da bu protokoller paketleri yol boyunca bir araya getirerek güvenli iletişimin yükünü azaltmaktadır. Ayrıca iki güvenlik sorununu da beraberinde getirmektedir. Birincisi, bir saldırgan ağa sahte bilgi enjekte etmek için toplama sürecine aktif olarak katılmasıdır. Saldırganların nasıl doğru bir şekilde tanımlanacağı ve toplama yoluna katılmalarının nasıl önleneceği, güvenli veri toplama protokolleri için ele alınması gereken kritik bir sorundur. İkincisi saldırganın, aynı zamanda toplama paketinin şifresini tamamen çözerek büyük miktarda hassas bilgi elde etmesidir. Bu, daha fazla toplama sonucuna sahip paketlerin gizlilik için daha güçlü koruma gerektirdiği anlamına gelmektedir. Toplama yolu boyunca güçlü veri şifreleme şemalarının nasıl tasarlanacağı da akıllı şebekede zorlu bir konudur [87].

Şekil 2.9. Siber güvenlik yönetim modeli

2.4. Siber Güvenlik Stratejileri, Düzenlemeleri ve Uygulamaları

Ülkelerin siber saldırılara karşı zafiyetlerinin arkasındaki faktörleri bilmek önemli bir adımdır. Bazı ülkelerin neden siber saldırılara yoğun şekilde maruz kaldığını belirlemek, siber suçluları nasıl engelleyeceğine ilişkin politikalar hakkında bilgi vermektedir. Web saldırıları ve aldatıcı uygulamaların genellikle gelişmiş ülkelerde olduğu görülmektedir. Yapılan analizlerde saldırganların büyük bilgi teknolojilerini ve parasal kaynakları hedef aldığı tespit edilmiştir [88]. Ülkeler ve kurumlar saldırılara engel olmak için yayınlanmış belgelerle ortaya konan bazı teknik standartlar belirlemektedir. Bu standartlar ağlara doğrudan veya dolaylı olarak bağlanabilen kullanıcıları, ağ cihazlarını, yazılımları, süreçleri, uygulamaları, depolama veya geçişteki bilgileri içermektedir. Araçlar, güvenlik kavramları, çeşitli politikalar, güvenlik koruması, risk yönetimi, eğitim gibi kavramlar siber güvenlik standardının parçalarıdır. Belirli sektörlere odaklanan birkaç tane düzenleme vardır. Bunlar sağlık, finans kuruluşları ve federal kurumların düzenlemelerinin parçası olan Sağlık Sigortası Taşınabilirlik ve Sorumluluk Yasası, Gramm Leach-Biley Yasası ve Federal Bilgi Güvenliği Yönetimi Yasasıdır [89]. Siber güvenliğin sağlanması, ulusal ve uluslararası düzeyde devlet, iş dünyası ve toplum için merkezi bir zorunluluk haline gelmiştir. Birçok ülke kendi siber güvenlik politikalarını ve kanunlarını geliştirmiştir. Ülkelerin politikalarını, düzenlemelerini, yasal çerçevelerini ve yeteneklerini geliştirme durumları ortak olsa da yaklaşımları birbirinden farklı olabilmektedir. Güvenlik politikası ve standardı geliştirmek ne kadar önemli ise etik, hukuk ve kanun oluşturmak da o kadar önemlidir. Ayrıca siber güvenlik çözümleri sadece teknolojiyi değil aynı zamanda bilgisayar bilimi ve hukuk alanları arasındaki iş birliğini de gerektirmektedir [90]. Ülkeler için Şekil 2.10'da gösterilen siber risk yönetim süreci önemli olmakla birlikte yapılacak işlemlere öncülük etmektedir.

Şekil 2.10. Siber güvenlik risk yönetim döngüsü

2.4.1. Dünyada Siber Güvenlik

Avrupa Birliği, siber saldırılarla mücadele etmek için Avrupa Birliği Ağ ve Bilgi Güvenliği Ajansını (European Union Agency for Network and Information Security- ENISA) kurmuştur. Avrupa Birliği Siber Güvenlik Yasasını yürürlüğe koymuş ve aşağıdaki hükümlerin üyeler tarafından uygulanmasını istemektedir. Siber yeteneklerin artırılması, iş birliği ve koordinasyonu geliştirmek, vatandaşların ve kurumların siber güvenlik farkındalığını artırması, dijital pazarda sunulan ürünlerin siber güvenlik güvencesinin şeffaflığını artırması ve güvenlik gerekliliklerine en üst düzeyde riayet edilmesidir [91].

Amerika Birleşik Devletleri teknoloji, hukuk ve politika düzenlemelerinde dünyaya önderlik etmektedir. ABD İç Güvenlik Bakanlığı siber güvenlik stratejisine göre güvenli iletişim ve güvenlik konularında çeşitli tedbirler almayı amaçlamaktadır. Gelişen siber güvenlik risklerini değerlendirerek ülkenin duruşunu belirlemek, hükümetin bilgi sistemlerini korumak, kilit paydaşlarla ortaklık yaparak kritik altyapıları korumak, uluslararası suç örgütleriyle siber uzayda mücadele ederek saldırıları önlemek, koordine edilmiş yapılarla siber olayların sonuçlarını en aza indirmek, politika ve standartları desteklemek ve siber güvenlik çabalarını entegre etmektir. Çeşitli kurum ve kuruluşlarla en üst derecede tedbir almak ve siber olaylarla mücadele etmek birincil öncelikleri arasına girmiştir [92].

Fransa ulusal dijital güvenlik stratejisi, Fransızların dijital geçişini desteklemek ve ülkeyi dijital stratejik özerklik için lider konuma getirmeyi amaçlayan itici bir güçtür. Fransa Ulusal Siber Güvenlik Ajansı önderliğinde yürütülen bu strateji, dijital çağın sorunlarına yanıt vermek için koordineli şekilde çalışmaktadır. Siber suç, casusluk, propaganda, sabotaj ve kişisel verilerin güvenliği için beş stratejik önceliğe önem vermektedir. Devletin bilgi sistemlerinin ve kritik altyapılarının savunulması ve güvenliği için önlem almak, dijital güveni, kişisel verileri ve siber saldırıları engellemek, farkındalık yaratarak eğitimleri artırmak, dijital teknoloji işletmelerine önem kazandırmak ve dijital stratejik özerklik ve siber uzay istikrarını sağlamak için çalışmaktadır [93].

İngiltere Ulusal Siber Güvenlik Stratejisinde, İngiltere aleyhine yürütülen siber tehditleri tespit etme, araştırma ve bunlara karşı koyma yeteneğini geliştirmek, siber suçun etkisini azaltmak ve siber saldırı hedefi olmaktan uzaklaştırmak, saldırıları azaltmak ve düşmanlara karşı koyma yeteneğine sahip olmak, aktif siber savunmayla kötü amaçlı yazılımları engellemek, ülkeyi dijital açıdan daha güvenli hale getirmek, yüksek kaliteli tavsiyelerle siber riskleri etkin şekilde yönetmek, ülkeyi siber güvenlik araştırma ve geliştirmede evrensel önder haline getirmek ve hem ekonomik hem de teknolojik alanda lider hale getirmektir [94].

Çin Ulusal Siber Güvenlik Stratejisine göre temelde 5 ana hedef ortaya koyulmuştur. Bilgi güvenliğini tehlikeye sokanlar engellenmeli ve siber alandaki çatışmaların önlenmesi için çalışmalar yapılmalıdır. Toplumda siber farkındalık artırılmalı ve siber güvenlik yetenekleri artırılarak koruma sistemleriyle desteklenmelidir. Standart ve politikalar gayet şeffaf olmalı ve

üretilen ürünlerin dolaşımına destek verilerek hiçbir ayrım yapmadan gelişen fırsatlar bunlarla paylaşılmalıdır. Tüm vatandaşların siber güvenliğe destek verici şekilde katılımının sağlanması ve ulusal kritik altyapıların korunmasında öncelik halkın çıkarları olmalıdır [95].

Rusya Federasyonu Bilgi Güvenliği Doktrinine göre, hükümetin tüm organlarıyla düzenli yapılan tatbikatlar da dahil olmak üzere siber saldırılara karşı hazırlıklarını artırmak amacıyla yeni yöntemler geliştirmek, bilgi, analitik, bilimsel ve teknik yönleri geliştirmek, tüm kurum, kuruluş ve vatandaşlarla etkileşim halinde verimliliği artırmak gerekliliktir. Ayrıca bilgi güvenliği alanında faaliyet gösteren, güvenlik araçları geliştiren, bilgi güvenliği veren kuruluşlara devlet desteği uygulanmaktadır [96].

2.4.2. Türkiye'de Siber Güvenlik

Türkiye'de siber güvenlik alanında Ulusal Siber Güvenlik Stratejisi ve 2013-2014 Eylem Planı ve 2016-2019 Ulusal Siber Güvenlik Strateji Belgesi yayınlanmıştır ve çeşitli siber tatbikatlar, eğitim, kapasite artırımı ve farkındalık çalışmaları yapılmaktadır. Yine ulusal eylem planında siber suçlarla mücadele alanında ceza ve güvenlik tedbirlerine yönelik düzenlemeler yapılmaktadır. 185 sayılı Avrupa Konseyi Siber Suç sözleşmesi imzalanarak bu alanda yaşanabilecek hukuki uyuşmazlıkların en aza indirilmesi planlanmaktadır. Siber suçlarla mücadelede ulusal iş birliği yapılarak diğer kurum ve kuruluşların tecrübeleri ve teknolojileri paylaşılarak toplumsal faydanın artırılması amaçlanmaktadır [97].

Siber suç ve suçlularla mücadelede Emniyet Genel Müdürlüğü Siber Suçlarla Mücadele Daire Başkanlığı koordinesinde il birimleriyle birlikte görev yürütülmektedir. Siber suçlarla mücadele etmek, farkındalık oluşturmak, uluslararası iş birliği yapmak, ülkemize yönelik tehditleri analizler yaparak değerlendirmek ve karşı koymak, uzman soruşturmacı ve adli bilişim personeli yetiştirmek ve teknolojiyi takip ederek kapasiteyi artırmak gibi görevleri bulunmaktadır [98].

Türk Ceza Kanunu'nda siber suçlar Bilişim Alanında Suçlar başlığı altında (Madde 234), bir bilişim sisteminin bütününe veya bir kısmına, hukuka aykırı olarak giren veya orada kalmaya devam eden kimse bilişim sistemine girme olarak tanımlanmaktadır ve bu maddeye göre cezalandırılmaktadır. Bir bilişim sisteminin işleyişini engelleyen veya bozan, yok eden, değiştiren veya erişilmez kılan, sisteme veri yerleştiren, var olan verileri başka yere gönderen kişi, sistemi engelleme, bozma, verileri yok etme veya değiştirme (Madde 244) başlığında cezalandırılmaktadır. Başkasına ait bir banka veya kredi kartını, her ne suretle olursa olsun ele geçiren veya elinde bulunduran kimse, kart sahibinin veya kartın kendisine verilmesi gereken kişinin rızası olmaksızın bunu kullanarak veya kullandırtarak kendisine veya başkasına yarar sağlarsa, başkalarına ait banka hesaplarına ilişkilendirilerek sahte banka kartı üreten, satan, devreden, satın alan veya kabul eden kişi ve sahte oluşturulan veya sahtecilik yapılan bir banka kredi kartı kullanan kişiler banka veya kredi kartlarının kötüye kullanılması (Madde 245) başlığı altında cezalandırılmaktadır. Bilişim

sistemlerini kullanarak hırsızlık yapmak ise nitelikli hırsızlık (madde 142) olarak cezalandırılmaktadır. Bilişim sistemlerinin, banka veya kredi kurumlarının araç olarak kullanılması suretiyle kullanılması halinde nitelikli dolandırıcılık (Madde 158) suçuyla cezalandırılmaktadır. Bilişim sistemi kullanılarak kumar ve bahis siteleri vasıtasıyla kumar oynatan kişilerde kumar oynamak için yer ve imkân sağlama (Madde 228) suçuyla cezalandırılmaktadır. Özel hayatın gizliliğini ihlal (Madde 134) ve kişisel verilerin kaydedilmesi (Madde 135) ve verileri hukuka aykırı olarak verme veya ele geçirme (Madde 136) suçları da bilişim alanında işlenen suç yöntemine göre siber suç olarak cezalandırılmaktadır [99].

3. MATERYAL VE METOT

3.1. Materyal

İnsanlar bir suçun kurbanı olduklarında, bu tür suçlarla ilgilenen kolluk birimlerine başvurmaktadır. Bu veriler, birimin veri tabanına ayrıntılı olarak kaydedilir. Bu suçlar türüne, yöntemine, yılına vb. göre ayrılarak kolluk birimlerinin veri tabanlarında saklamaktadır. Bu bilgilere göre istatistikler hazırlanarak analiz edilmektedir. Veri tabanında çok sayıda suç bulunmasına rağmen, çalışmamızın odak noktası son yıllarda artan siber suçlardır. Siber suçlar ciddi maddi ve manevi zararlara neden olmakta ve önlenmesi de bir o kadar zor olmaktadır. Bu alandaki çoğu çalışmada gerçek veriler kullanılmadığı için siber suç konu olarak seçilmiştir.

3.1.1. Veri Seti

Veri seti, 2015-2019 yılları arasında Elazığ ilinde meydana gelen gerçek siber suç verileridir. Gerçek verilere erişim ve bu verilerin makine öğrenimi yöntemleriyle işlenmek üzere hazırlanması zorlu bir süreçtir. Veri seti elde edildiğinde, tüm siber suç ayrıntıları incelenmiştir. Gereksiz alanlar, çeşitli veri bilimi yöntemleri kullanılarak kaldırılmıştır. Veri setinde bulunan alanlara ait bilgiler Tablo 3.1'de verilmiştir.

3.1.2. Veri Hazırlama

Verilerle Python 3.7 programındaki çeşitli kütüphaneler kullanılarak tahminler yapılmıştır. Bu programın NumPy, Pandas, Matplotlib, Keras ve TensorFlow gibi ana kütüphaneleri kullanılmış ve veriler bu program aracılığıyla görselleştirilmiştir. Standardizasyon, normal bir dağılımdaki özelliklerin yeniden ölçeklendirilmesidir. Makine öğrenmesi yöntemleri kullanılmadan önce standartlaştırılarak verilerin uygun hale getirilmesi gereklidir. Verileri uygun hale getirmek için sütunlardaki veri çeşitliliğine göre 1'den 10'a kadar sayılar verilmiştir. Python kitaplığındaki StandardScaler (), algoritmalarda kullanılacak verileri optimize etmek için kullanılmıştır. Python programının rastgele özelliği ile verilerin %80'i eğitim verisi ve %20'si de test verisi olarak ayrılmıştır. Eğitim verisi ile modelin eğitimi gerçekleştirildikten sonra öğrenen sistem tahminler üretmiş ve test verisiyle karşılaştırılarak modelin doğruluk oranları elde edilmiştir. Yine veriler python kütüphanesindeki araçlar kullanılarak grafiksel şekillerle görselleştirilmiştir.

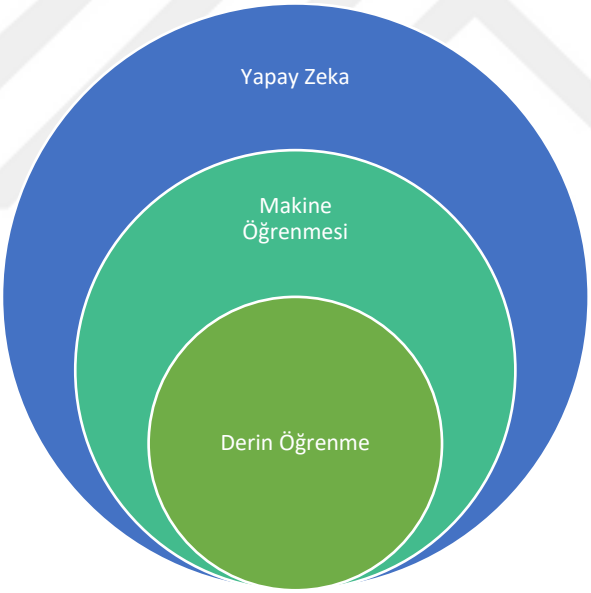
İlk modelde suç, cinsiyet, yaş, meslek, gelir, medeni hal, eğitim, saldırı yöntemi, saldırı şekli ve fail özellikleri verilerek saldırı amacı tahmin edilmiştir. İkinci modelde suç, cinsiyet, yaş, gelir, meslek, medeni durum, eğitim, saldırı amacı, saldırı şekli ve saldırı yöntemi özellikleri verilerek suç faili tahmin edilmiştir. Üçüncü modelde suç, cinsiyet, yaş, meslek, gelir, medeni durum, eğitim, saldırı amacı, saldırı şekli ve fail özellikleri verilerek saldırı yöntemi tahmin edilmiştir.

Tablo 3.1. Veri setindeki öznitelikler

Türü	Öznitelikleri	Sayısı
Suç	• Banka veya kredi kartlarının kötüye kullanılması	529
	• Bilişim yoluyla hırsızlık	90
	• Bilişim sistemine girme ve verileri ele geçirme	282
Cinsiyet	• Erkek	656
	• Kadın	245
Yaş	• 27 ve altı	224
	• 28-37 arası	254
	• 38-50 arası	274
	• 51 ve üstü	149
Gelir	• Düşük	436
	• Orta	389
	• Yüksek	76
Meslek	• Diğer	385
	• Öğrenci	104
	• Emekli	57
	• Adalet ve güvenlik	85
	• Sağlık sektörü	61
	• Yönetici	17
	• Ev hanımı	45
	• Eğitim öğretim	88
	• Teknik	37
	• Finans sektörü	22
Medeni Hal	• Evli	627
	• Bekar	274
Eğitim	• İlköğretim	191
	• Lise	351
	• Lisans	310
	• Lisansüstü	49
Saldırı Amacı	• Kendini Banka Görevlisi Olarak Tanıtarak İnternet alışverişi	68
	• Bilgisi Dışında İnternet Alışverişi	282
	• Bilgisi Dışında Para Çekilmesi	136
	• Kendini Banka Görevlisi Olarak Tanıtarak Para Çekilmesi	43
	• Dolandırıcılık	224
	• Manevi Zarar Vermek	127
	• Tehdit/Şantaj	21
Saldırı Şekli	• Elektronik Banka Hesaplarının Ele Geçirilmesi	475
	• Banka/ATM Kartı Kopyalanması	54
	• Sosyal medya hesabının ele geçirilmesi	214
	• Dijital Ortamdaki Verilerinin Elde Edilmesi ve kullanılması	112
	• Ürün Satışı Yapılması	45
Saldırı Yöntemi	• Sosyal Mühendislik Kullanarak	140
	• Hack Araçları veya Zararlı Yazılım Kullanarak	482
	• Kart Kopyalama, Üretme Cihazlarını Kullanarak	54
	• Phising (oltalama) Saldırısı Kullanarak	67
	• Sosyal Medyadaki Herkese Açık Verilerini Alarak	112
	• Sahte Alışveriş Sitesi Oluşturarak	46
Fail Durumu	• Faili bilinmeyen	510
	• Faili bilinen	391

3.2. Metot

Yapay zekâ, insanlar gibi davranmaya çalışan bilgisayar uygulamalarıdır. İnsanlar için deneyim ve yeni şeyler öğrenme ne kadar önemli ise makinalar içinde aynı durum söz konusudur. Aslında yapay zekâ teknolojileri de insanlar gibi pratik yapmakta ve sürekli öğrenerek performansını artırmaktadır. Bir yapay zekâ programı ne kadar iyi öğrenmiş ise deyimi o kadar artmakta ve insana en yakın sonuçları vermektedir. Makine öğrenmesi yapay zekanın bir alt sınıfı olup verileri insanlar gibi analiz edip sonuçlar üreten bir öğrenme türüdür. Makine öğrenmesinin bir altında Yapay sinir Ağları vardır. Bunun da bir alt dalı derin öğrenme yaklaşımlarıdır. Şekil 3.1'de gösterildiği şekilde yapay zekanın alt dalları birbiriyle ilişkilidir. Derin öğrenme bir programın genel çalışmasını tanımaya yarayan insan beynine benzer şekilde çalışan yeni bir dönüm noktasıdır. Her bilgisayar programı öğrenemez, bazıları kendisine ne yüklendiyse o talimatları yerine getirmektedir. Fakat önerilerde bulunan, önlemler almaya çalışan ve insana kararlarında yardımcı olan program ve uygulamalar öğrenme içermektedir [100].



Şekil 3.1. Yapay zekâ ve alt dalları

Son yıllarda makine öğrenmesi ve derin öğrenme algoritmaları birçok çalışmada kullanılmıştır. Siber saldırılar ile sistemlere izinsiz girişleri tespit etmede [101, 102, 103, 104, 105], hizmeti engelleme ve gizlilik saldırılarını tespit etmede [106], akıllı şebekeleri ve elektrik altyapılarına yönelik saldırıları tespit etmede [107], bilgisayarlara bulaşan kötü niyetli yazılım ile kod bloklarını ve bu yazılımların bulaşma ihtimalini tespit etmede [108, 109, 110, 111, 112], meydana gelen suçları veya acil durumları görüntüleyen kameralardan alınan video akışlarını incelemede [113], metin analizleri yaparak tahminler üretmede [114, 115], meydana gelen suç

verilerinden gelecekteki suçları tahmin etmede [116, 117] ve siber olaylarda kişilik profillemede [118, 119] ML ve DL uygulamaları kullanılarak başarılı sonuçlar elde edilmiştir.

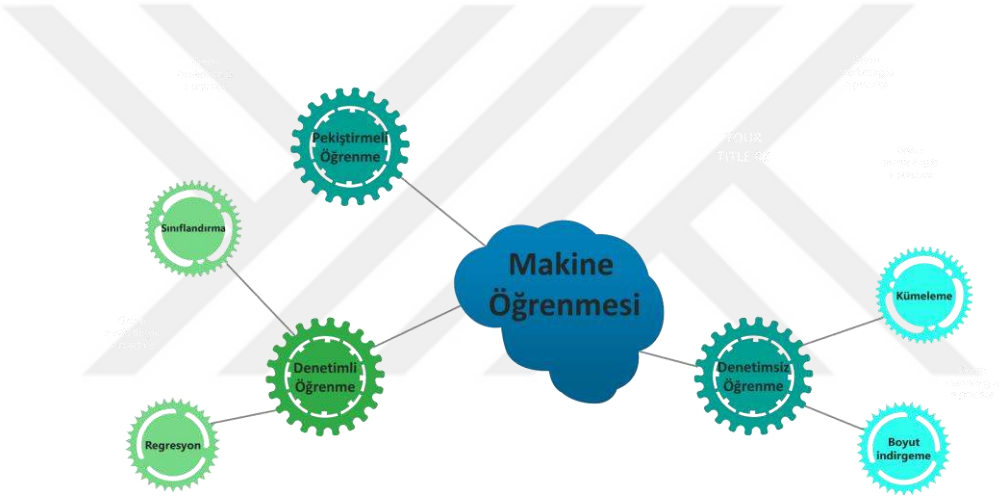
Özellikle makine öğrenimi ve derin öğrenme yöntemleri birçok çalışmada kullanılmış ve başarılı sonuçlar elde edilmiştir. Makine öğrenimi yöntemlerinin, birden çok yapılandırılmış ve yapılandırılmamış veri modelini tanıma, suç taktiklerini değiştikçe tespit etmede üst düzey başarı, karmaşık veriler arasındaki ilişkileri çıkarma ve insanlar tarafından öngörülemeyen sonuçlar üretme yeteneği gibi avantajlarından dolayı tezde kullanılmıştır. Yapay sinir ağlarının kullanılmasının ise yapısının lineer olmaması, girdi ve çıktı eşleştirmeleri ile tasarlanabilmesi, uyumlu olması ve hatalara toleranslı olmasıdır.

3.2.1. Makine Öğrenmesi

Bilgisayardaki problemleri çözmek için çeşitli algoritmalara ihtiyaç duyulmaktadır. Bu algoritmalar mağazalardan, süpermarketlere kadar birçok sektöre yardım etmektedir. Algoritmalar; finans sektörü, borsa, dolandırıcılık tespiti gibi birçok sektördeki büyük verileri analiz ederken kullanılmaktadır. Makine öğrenmesi de bu algoritmalarla çeşitli çözümler üretmektedir. Makine öğrenmesi yalnızca bir veri tabanı problemi değil aynı zamanda yapay zekanın bir parçasıdır. Bir sistemin akıllı olabilmesi için öğrenme yeteneği de olması gerekmektedir. Makine öğrenimi görme, konuşma tanıma ve robotikteki birçok soruna çözüm bulmamıza yardımcı olmaktadır. Makine öğrenimi, örnek verileri veya geçmiş deneyimleri kullanarak bir performans kriterini optimize etmek için bilgisayarları programlamaktır. Örneğin bazı parametrelere kadar tanımlanmış bir modelde; öğrenme, eğitim verilerini veya geçmiş deneyimleri kullanarak modelin parametrelerini optimize etmek için bir bilgisayar programının yürütülmesidir. Model, gelecekte tahminlerde bulunmak için tahmine dayalı veya verilerden bilgi elde etmek için tanımlayıcı ya da her ikisi birden olabilmektedir. Makine öğrenimi, matematiksel modellerin oluşturulmasında istatistik teorisini kullanmaktadır çünkü temel görev bir örnekten çıkarımda bulunmaktır. Bilgisayar biliminin rolü iki yönlüdür. Birincisi, eğitimde, optimizasyon problemini çözmek ve genel olarak sahip olduğumuz büyük miktardaki veriyi depolamak ve işlemek için verimli algoritmalara ihtiyacımızın olmasıdır. İkincisi ise, bir model öğrenildikten sonra, gösterim ve çıkarım için algoritmik çözümünün de verimli olması gerekmektedir. Belirli uygulamalarda, öğrenme veya çıkarım algoritmasının etkinliği, yani uzay ve zaman karmaşıklığı, tahmin doğruluğu kadar önemli olabilmektedir [120]. Öğrenme modelleri Şekil 3.2'de gösterilmiştir.

Denetimli öğrenmede; makine, makineye verilerin neyi temsil ettiğini söylemek için etiketlenmiş örnek veriler kullanılarak eğitilmektedir. Bir dizi girdi değerinden bir çıktı değerini tahmin etme görevi, istatistiklerde doğrusal model için doğrusal regresyon olarak adlandırılmaktadır. Makine öğreniminde regresyon, denetimli öğrenmenin bir türüdür. Genel durumda öğrenme, bir performans kriterine göre daha iyi olmayı ifade eder ve regresyonda

performans model tahminlerinin eğitim verilerinde gözlemlenen çıktı değerlerine ne kadar yakın olduğuna bağlıdır. Buradaki varsayım, eğitim verilerinin temeldeki görevin özelliklerini yeterince iyi yansıttığıdır. Bu nedenle eğitim verileri üzerinde doğru çalışan bir modelin verilen görevi öğrenmiş olduğu söylenebilmektedir. Denetimli öğrenme algoritmasının amacı, her yeni girdi için yeni bir tahmin edilen çıktı yaratılacak şekilde bir tahmin haritalama işlevi elde etmektir. Başka bir deyişle öğrenme algoritması, karşılık gelen çıktılar ile bir dizi girdi almaktadır. Algoritma hataları bulmak ve öğrenme modelini buna göre modifiye etmek için somut çıktısını doğru çıktılarla eşitleyerek öğrenmektedir. Denetimli öğrenme algoritmaları, etiketlenmemiş verilerdeki etiketin değerlerini tahmin etmek için kalıplardan yararlanmaktadır. Bu, sınıflandırma, regresyon, tahmin vb. ile elde edilmektedir. Bu eğitime bağlı olarak, makine yeni verileri analiz edebilmeli ve doğru cevabı tahmin edebilmelidir [121].



Şekil 3.2. Makine öğrenmesi türleri

Denetimsiz öğrenmede; makine, etiketleri olmayan veriler kullanılarak eğitilmektedir. Denetimsiz öğrenme, eşdeğer çıktı değişkenleri olmadan yalnızca bir girdi verisinin mevcut olduğu yerdir. Denetimsiz öğrenmenin amacı, veriler hakkında daha fazla bilgi edinmek için verilerin yapısını modellemektir. Algoritmaların bir sonuca varmak için veriler içinde bir yapı, bir çıkarım ve anlam keşfetmesi gereklidir. Bu algoritmalar, denetimli algoritmalardan farklı olarak çıktıyı tahmin etmek için herhangi bir geçmiş veriye sahip değildir. Bu, verilerin neyi temsil ettiğini veya hangi yanıtların beklendiğini makinenin bilmediği anlamına gelmektedir. Makine, etiketlenmemiş girdinin modellerini ve yapısını kendi başına çözmeli ve beklenen çıktıyı keşfetmelidir. İstatistikte buna yoğunluk tahmini denir. Kümeleme yoğunluk tahmini için bir yöntem, amacın girdi kümelerini veya gruplarını bulmak olduğu kümeleme işlemidir. Bir kümeleme modeli, nitelikleri bakımından benzer müşterileri aynı gruba tahsis ederek şirkete müşterilerinin doğal gruplamalarını sağlar; buna müşteri segmentasyonu denir. Bu tür gruplar bulunduğunda şirket, farklı gruplara özgü stratejileri örneğin hizmet ve ürünleri müşteri ilişkileri yönetimi olarak tanımlamaktadır [121].

Pekiştirmeli öğrenmede, makine belirli bir hedefe ulaşmak için çevresi ile etkileşime girmektedir. Makine etiketlenmemiş veriler kullanılarak eğitildiği için denetimsiz öğrenmeye benzemektedir. Ancak, pekiştirmeli öğrenmede, makine sonuçla ilgili geri bildirim almaktadır. Bazı uygulamalarda sistemin çıktısı bir dizi eylemdir. Böyle bir durumda tek bir eylem önemli değildir önemli olan hedefe ulaşmak için doğru eylemlerin sırası olan politikadır. Herhangi bir ara durumda en iyi eylem diye bir şey yoktur ki iyi bir politikanın parçasıysa eylem iyidir. Böyle bir durumda, makine öğrenimi programı politikaların iyiliğini değerlendirebilmeli ve bir politika oluşturabilmek için geçmişteki iyi eylem dizilerinden ders alabilmelidir. Bu tür öğrenme yöntemlerine pekiştirmeli öğrenme algoritmaları denir. Öğrenme, tek bir hareketin kendi başına o kadar önemli olmadığı oyun oynamaktır ki iyi olan doğru hamlelerin sırasıdır. Hareket, iyi bir oyun oynama politikasının parçasıysa iyidir. Oyun oynama hem yapay zekâ hem de makine öğreniminde önemli bir araştırma alanıdır. Bunun nedeni, oyunların tanımlanması kolay ve aynı zamanda iyi oynamalarının da oldukça zor olmasıdır. Satranç gibi bir oyunun az sayıda kuralı vardır, ancak her durumda çok sayıda olası hamle ve bir oyunun içerdiği çok sayıda hamle nedeniyle çok karmaşıktır. Oyunları iyi oynamayı öğrenebilen iyi algoritmalara sahip olunduğunda, bunları daha belirgin ekonomik fayda sağlayan uygulamalara da uygulanabilmektedir. Bir hedef konumu aramak için bir ortamda gezinen bir robot, pekiştirmeli öğrenmenin başka bir uygulama alanıdır [120].

3.2.2. Lojistik Regresyon (Logistic Regression LR)

Lojistik regresyon, daha birçok karmaşık uzantısı mevcut olmasına rağmen, temel biçiminde bir ikili bağımlı değişkeni modellemek için lojistik bir işlev kullanan istatistiksel bir modeldir. Regresyon analizinde, lojistik regresyon (veya logit regresyon), bir lojistik modelin (bir ikili regresyon formu) parametrelerini tahmin etmektedir. Matematiksel olarak, bir ikili lojistik model, iki değerin "0" ve "1" olarak etiketlendiği, bir gösterge değişkeni ile temsil edilen başarılı / başarısız gibi iki olası değere sahip bir bağımlı değişkene sahiptir.

Lojistik regresyon fonksiyonu Denklem 3.1 ve 3.2'deki gibi iki şekilde ifade edilmektedir.

$$\pi(x) = \frac{exp\ (\beta_0+\beta_1 X)}{1+exp\ (\beta_0+\beta_1 X)} \tag{3.1}$$

$$\pi(x) = [1 + exp\ (-\beta_0 - \beta_1 X)]^{-1} \tag{3.2}$$

Burada β regresyon katsayısını ve X ise bağımsız değişkenlerdir.

İkili, ordinal ve nominal olmak üzere üç temel lojistik regresyon analizi vardır. İkili lojistik regresyon bir ya da daha çok açıklayıcı değişken ile arasındaki ilişkiyi ortaya koyduğu için denir. Ordinal lojistik regresyon, cevap değişkeninin üç veya daha çok özellik içerdiği ve değerlerin sıralı elde edildiği durumda neden ve sonuç ilişkisini çıkarmaya çalışan yöntemdir. Nominal lojistik

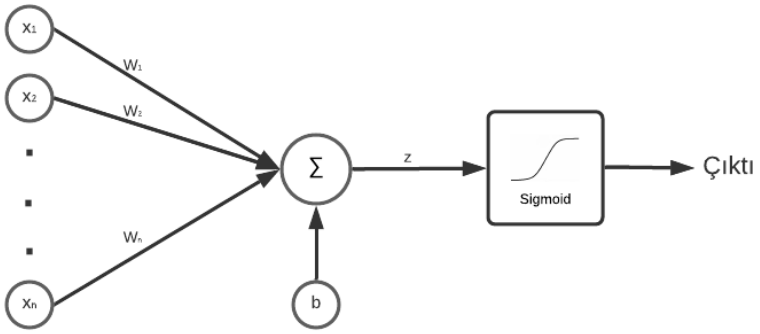
regresyon, cevap değişkeninin üç veya daha çok özellik içerdiği ve değerlerin isimsel ölçekle elde edildiği durumda neden ve sonuç ilişkisini çıkarmaya çalışan yöntemdir.

Lojistik regresyon, birçok farklı insan tarafından yaygın olarak kullanılmaktadır, ancak kısıtlayıcı dışavurumuyla mücadele etmektedir (örneğin, etkileşimler manuel olarak eklenmelidir) ve diğer modeller daha iyi tahmin performansına sahip olabilmektedir. Lojistik regresyon modelinin diğer bir dezavantajı, ağırlıkların yorumlanmasının katma değer değil çarpımsal olması nedeniyle yorumlamanın daha zor olmasıdır. Avantajı ise lojistik regresyon modeli yalnızca bir sınıflandırma modeli değil, aynı zamanda olasılıklar da vermesidir. Bu, yalnızca son sınıflandırmayı sağlayabilen modellere göre büyük bir avantajdır. Örneğin, %51'e kıyasla bir sınıf için %99 olasılığa sahip olduğunu bilmek büyük bir fark yaratmaktadır. Lojistik regresyon, ikili sınıflandırmadan çok sınıflı sınıflandırmaya kadar genişletilebilmektedir [122].

Lojistik Regresyon, temel bir sınıflandırma tekniğidir. Bu, onun devasa eğitimsel ve didaktik potansiyelini okumaktadır. LR kullanılırken, gelecekte uğraşılabilecek her sınıflandırma probleminde (özellikle daha güçlü algoritmalar dahil edildiğinde) yeniden kullanılacak tüm temel kavramları (maliyet fonksiyonları, model performans ölçütleri, karışıklık matrisi gibi ortak araçlar vb.) öğrenilmelidir. LR, kategorik bir değişkeni diğer değişkenler açısından modellemek için bir araçtır. Her bir değişkendeki değişikliklerin diğer değişkenlerdeki farklı sonuçları nasıl etkilediğini bulmanın yollarını vermektedir. LR hızlıdır ve nispeten karmaşık değildir. Bu nokta, modern makine öğrenimi ortamındaki LR tabanlı sınıflandırıcıların uygunluğunu vurgulamaktadır. Yeni örneklem dışı kayıtlarını gerçek zamanlıya yakın bir şekilde sınıflandırmaya ihtiyaç olduğunda (özellikle yanlış sınıflandırma maliyetiyle birleştirildiğinde, uzun sınıflandırma süresinin maliyetinden daha az olduğunda), Lojistik Regresyon gereken bir seçenek olmaktadır. LR, sonuçları yorumlamak için uygundur. LR kullanılırken model karar verme sürecini (özellik önemi, tahmin etmenleri vb. açısından) kolayca yorumlanmaktadır. Modern iş dünyası, yorumlanabilir ML için giderek daha fazla talepte bulunmaktadır ve LR eğitimi böyle bir yorumlanabilirlik çağrısına hizmet etmektedir. Lojistik regresyon, bir bağımlı değişken ile birkaç bağımsız değişken arasında çok değişkenli bir regresyon ilişkisi oluşturmaya izin vermektedir. Çok değişkenli analiz modellerinden biri olan LR, bir yordayıcı değişkenler kümesinin değerlerine dayalı olarak bir özelliğin veya sonucun varlığını veya yokluğunu tahmin etmek için kullanışlıdır. LR'nin avantajı, normal doğrusal regresyon modeline uygun bir bağlantı fonksiyonunun eklenmesi yoluyla, değişkenlerin sürekli veya kesikli olabileceği veya her iki türün herhangi bir kombinasyonu olabileceği ve mutlaka normal dağılımlara sahip olmamasıdır [123].

Şekil 3.3'te gösterilen grafikte ve Denklem 3.3'teki formülde w, ağırlıkları, b bias değerini, x, girdileri, ifade etmektedir. Z değerini bulmak için girdiler tek tek ağırlıklarla çarpılmaktadır ve b değeri eklenerek hesaplanmaktadır. Sigmoid işlevi, z'yi sıfır ile bir arasında yapmaktadır, (çıktı= Sigmoid(z)) bu da olasılıktır [124].

$$z = b + x_1 w_1 + x_2 w_2 + \ldots + x_n w_n \tag{3.3}$$



Şekil 3.3. Lojistik regresyon grafiği

3.2.3. En Yakın Komşu (K-Nearest Neighbors K-NN)

Veri setindeki her örneğin n boyutlu bir vektör oluşturmak için birleştirilen n özniteliğe sahip olduğu varsayıldığında Denklem 3.4'teki gibi ifade edilmektedir.

$$(\mathrm{x}) = (x_1,\ x_2,\ x_3, \ldots,\ x_n) \tag{3.4}$$

Bu n öznitelikler bağımsız değişkenler olarak kabul edilmektedir. Her örneğin, değeri diğer n özniteliğine bağlı olan, y (bağımlı değişken) ile gösterilen başka bir özniteliğe de sahiptir. Y'nin kategorik bir değişken olduğunu varsayıldığında ve bu tür her vektöre bir sınıf, y = f (x) atayan bir skaler fonksiyon f vardır.

F hakkında hiçbir şey bilinmiyorsa (aksi takdirde veri madenciliğine gerek yoktur), ancak bir bakıma sorunsuz olduğunu varsayılmaktadır. Bu tür vektörlerin karşılık gelen sınıflarıyla birlikte verildiği varsayılırsa Denklem 3.5'teki gibidir.

$$\mathrm{x}^{(i)}, \mathrm{y}^{(i)} \qquad i = 1, 2, \ldots, T. \tag{3.5}$$

Bu sete eğitim seti denir. Çözmek istenen sorun şudur. Varsay ki, x = u olan yeni bir örnek verilsin, bu örneğin ait olduğu sınıfı bulmaktır. F fonksiyonu bilinseydi, bu yeni örneği nasıl sınıflandırılacağını bilmek için sadece v = f (u) hesaplanırdı, ama elbette f hakkında yeterince pürüzsüz olması dışında hiçbir şey bilinmemektedir. K-En Yakın Komşu yöntemlerinde fikir, bağımsız değişkenleri x, u'ya benzeyen eğitim setindeki k örneklemini belirlemek ve bu k örneklemini bu yeni bir örnek olan v sınıfında kullanmaktır. Eğer f'nin düzgün bir fonksiyon olduğu varsayıldığında, mantıklı bir fikir eğitim verilerinde ona yakın olan (bağımsız değişkenler açısından) örnekleri aramak ve sonra y'nin değerlerinden v'yi hesaplamaktır. Komşular hakkında,

bağımsız değişkenlere dayalı olarak örnekler arasında hesaplanan bir mesafe veya farklılık ölçüsü olduğudur. Şu an için en popüler uzaklık Öklid ölçüsüdür. X ve u noktaları arasındaki Öklid mesafesi Denklem 3.6'daki gibi hesaplanmaktadır.

$$d(x,u) = \sqrt{\sum_{i=1}^{n}(x_i - u_i)^2} \tag{3.6}$$

Kümeleme yöntemlerini tartışırken bağımsız tahmin değişkenlerinin uzayındaki noktalar arasındaki mesafeyi ölçmektir. En basit durum, eğitim setinde u'ya en yakın olan (en yakın komşu) örneği ve v = y olarak ayarlanan, burada y en yakın komşu örneğin sınıfı olduğu k = 1'dir. Örnekleri sınıflandırmak için en yakın tek bir komşuyu kullanma şeklindeki bu basit, sezgisel fikrin eğitim setinde çok sayıda örnek olduğunda çok güçlü olabileceği dikkate değer bir gerçektir. Büyük miktarda veri varsa ve keyfi olarak karmaşık bir sınıflandırma kuralı kullanılırsa, yanlış sınıflandırma hatasını en iyi ihtimalle basit 1-NN kuralının yarısına indirebileceğini kanıtlamak mümkündür.

K-NN yaklaşımının pratik kullanımında iki güçlük vardır. İlk olarak, eğitim verilerinden parametreleri tahmin etmek için zaman gerekmemekle birlikte (yöntem parametrik bir yöntem olmadığından), büyük bir eğitim setinde en yakın komşuları bulma süresi engelleyici olabilmektedir. Bu zorluğun üstesinden gelmek için bir dizi fikir hayata geçirilmiştir. Ana fikirler şunlardır:

1. Temel bileşenler gibi boyut küçültme tekniklerini kullanarak azaltılmış bir boyutta çalışarak mesafeleri hesaplamak için harcanan zamanı azaltmaktır.
2. En yakın komşunun tanımlanmasını hızlandırmak için arama ağaçları gibi karmaşık veri yapıları kullanmaktır. Bu yaklaşım genellikle hızı iyileştirmek için "neredeyse en yakın" komşuya karar vermektir.
3. En yakın komşuyu aramayı hızlandırmak için eğitim setindeki fazlalık veya "neredeyse gereksiz" noktaları kaldıracak şekilde eğitim verilerini düzenlemektir. Bir örnek, hepsi aynı sınıfa ait olan örneklerle çevrelenmiş oldukları için sınıflandırma üzerinde hiçbir etkisi olmayan eğitim veri setindeki örnekleri kaldırmaktır.

İkinci olarak, eğitim veri setini büyük olarak nitelendirilmek için gereken örnek sayısı, n boyutlarının sayısıyla katlanarak artmaktadır. Bunun nedeni, eğitim veri setinin boyutu n ile üssel olarak artmadıkça, en yakın komşuya olan beklenen mesafenin n ile önemli ölçüde artmasıdır. Bu elbette boyutsallık laneti yüzündendir. Boyutsallık laneti, tüm sınıflandırma, tahmin ve kümeleme teknikleriyle ilgili temel bir konudur. Bu nedenle, model için yordayıcı değişkenlerin alt kümelerini seçmek veya bunları temel bileşenler, tekil değer ayrıştırması ve faktör analizi gibi yöntemler kullanarak birleştirmek gibi yöntemlerle sık sık yordayıcı değişkenlerin uzayının boyutsallığını azaltmaya çalışmaktır [125].

3.2.4. Destek Vektör Makinesi (Support Vector Machine SVM)

Destek vektör makinesi, bir tür denetimli makine öğrenimi sınıflandırma algoritmasıdır. SVM'ler ilk olarak 1960'larda tanıtılmış ve daha sonra 1990'larda rafine edilmiştir. Ancak, mükemmel sonuçlar elde etme yetenekleri sayesinde ancak şimdi son derece popüler hale gelmektedir. SVM'ler, diğer makine öğrenimi algoritmalarıyla karşılaştırıldığında benzersiz bir şekilde uygulanmaktadır. Doğrusal ve kernel olmak üzere iki destek vektör makinesi kullanılmaktadır. Bir SVM, iki kategori arasındaki boşluğun genişliğini en üst düzeye çıkarmak için eğitim örneklerini uzaydaki noktalarla eşlemektedir. Yeni örnekler daha sonra aynı alana eşlenmekte ve boşluğun hangi tarafına denk geldiklerine bağlı olarak bir kategoriye ait oldukları tahmin edilmektedir. Doğrusal sınıflandırma gerçekleştirmenin yanı sıra SVM'ler, çekirdek numarası olarak adlandırılan şeyi kullanarak girdilerini yüksek boyutlu özellik uzaylarına örtük olarak eşleyerek doğrusal olmayan bir sınıflandırmayı verimli bir şekilde gerçekleştirmektedir. Veriler etiketsiz olduğunda denetimli öğrenme mümkün değildir ve verilerin gruplara doğal olarak kümelenmesini bulmaya çalışan ardından yeni verileri bu oluşturulmuş gruplara eşleştiren denetimsiz bir öğrenme yaklaşımı gereklidir. Hava Siegelmann ve Vladimir Vapnik tarafından oluşturulan destek vektörü kümeleme algoritması, etiketlenmemiş verileri kategorize etmek için destek vektör makineleri algoritmasında geliştirilen destek vektörlerinin istatistiklerini uygulamaktadır ve endüstriyel alanda en yaygın kullanılan kümeleme algoritmalarından biridir. SVM diğer sınıflandırma algoritmalarından, tüm sınıfların en yakın veri noktalarına olan mesafeyi en üst düzeye çıkaran karar sınırını seçmesi bakımından farklılık göstermektedir. Bir SVM yalnızca bir karar sınırı bulmaz; en uygun karar sınırını bulmaktadır. En uygun karar sınırı, tüm sınıfların en yakın noktalarından maksimum marjı olan sınırdır. Karar sınırı ile noktalar arasındaki mesafeyi maksimize eden karar sınırından en yakın noktalara destek vektörleri denir. Destek vektörlerini bulmanın, karar sınırı ile destek vektörleri arasındaki marjın hesaplanmasının ve bu marjın maksimize edilmesinin arkasında karmaşık bir matematik vardır [126].

Örnek doğrusal olarak ayrılabilir olduğunda, SVM sınıflandırma ilkesi Şekil 3.4'te gösterilmiştir. İki tür verinin kenarında iki mavi çizgi bulunmaktadır, kenar çizgisindeki nokta destek vektörü olarak adlandırılır ve kırmızı çizgi en uygunudur.

Optimal sınıflandırma çizgisi Denklem 3.7'de verilmiştir.

$$\omega^T x + b = 0 \tag{3.7}$$

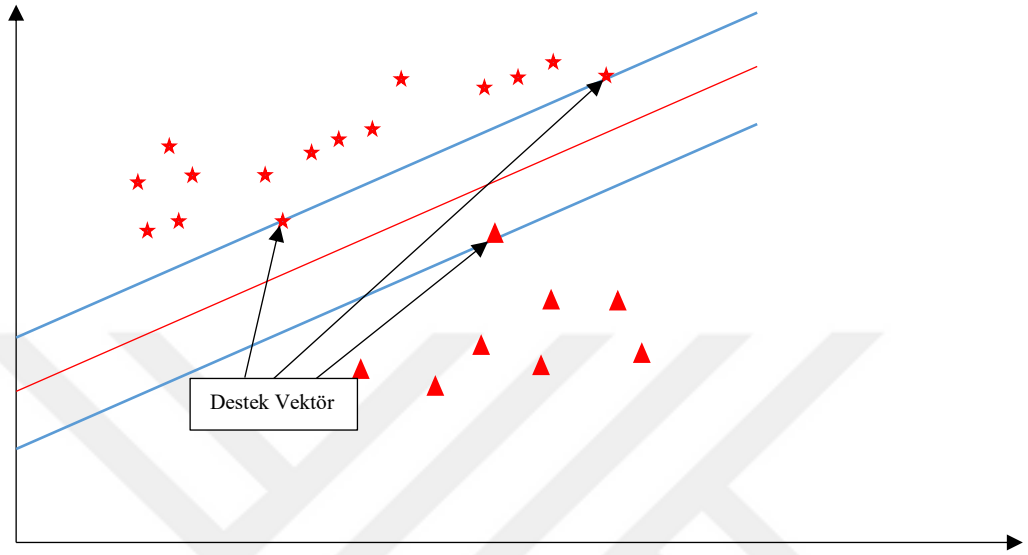
Destek vektörünün bulunduğu çizgiler Denklem 3.8 ve Denklem 3.9'da verilmiştir.

$$\omega^T x + b = 1 \tag{3.8}$$

$$\omega^T x + b = -1 \tag{3.9}$$

Sınıflandırma aralığı Denklem 3.10'daki gibi kaydedilir.

$$max \frac{2}{||\omega||} \tag{3.10}$$



Şekil 3.4. Destek vektör makinesi şematik gösterimi

Burada ω ağırlık vektörü, x girdi vektörü ve b sapmadır. Optimal sınıflandırma çizgisini bulma süreci, sınıflandırma aralığını maksimize etme sürecidir. Problem ikili probleme dönüştürülürse, Lagrange çarpanı yöntemi kullanılarak elde edilebilmektedir. Tüm akış verileri örnek noktalarının eğitim setinden en iyi sınıflandırıcıya sahip test örneklerini sınıflandırırken, sınıflandırıcı modelin eğitim örneğini aşırı uyarlaması ve tüm eğitim verilerini doğru şekilde sınıflandırması çok olasıdır. Test örneklerinin hata toleransı daha düşüktür. Destek vektör makinesi algoritmasına dayalı yumuşak kenar boşluğu örnek setini eğitmek için kullanılmaktadır, böylece sınıflandırıcı, parametre ceza faktörü ekleyerek daha iyi hata toleransına sahip olmaktadır.

Doğrusal ayrılabilirlik sorununu çözmek için, en uygun sınıflandırma çizgisi bulunmakta ve doğrusal ayrılmazlık sorununu çözmek için yüksek boyutlu uzayla eşleştirilmektedir. Çekirdek işlevi kullanılarak doğrusal olmayan örnek veriler için bir destek vektör makine sınıflandırıcısı oluşturulmaktadır. Çekirdek işlevi, yüksek boyutlu uzayda örnek verilerin haritalamasını temsil etmek zorunda kalmadan doğrudan orijinal düşük boyutlu uzayda hesaplanmaktadır, bu da yüksek boyutlu uzaydaki karmaşık hesaplamaları etkili bir şekilde önlemektedir. Destek vektör makinelerinde yaygın olarak kullanılan dört tür çekirdek işlevi vardır:

Doğrusal çekirdek işlevi Denklem 3.11'deki formülle ifade edilmektedir.

$$k(x_i, x_j) = x_i{}^T x_j \tag{3.11}$$

Polinom çekirdek işlevi Denklem 3.12'deki formülle ifade edilmektedir.

$$k(x_i, x_j) = (x_i{}^T\, x_j)^d \; d \geq 1 \tag{3.12}$$

Gauss kernel işlevi Denklem 3.13'teki formülle ifade edilmektedir.

$$k(x_i, x_j) = \; exp\; (-\frac{\left||x_i - x_j|\right|^2}{2\sigma^2}) \tag{3.13}$$

Sigmoid çekirdek işlevi Denklem 3.14'teki formülle ifade edilmektedir.

$$k(x_i, x_j) = \; tanh(\beta\, x_i{}^T x_j + \theta) \tag{3.14}$$

Ceza faktörü ve çekirdek parametrelerinin seçimi, örnek verileri sınıflandırmak için SVM algoritması kullanıldığında sınıflandırıcının performansını etkileyebilmektedir. SVM modelinin eğitim örneklerine daha iyi uyum sağlamasını ve eğitim örneklerini daha doğru bir şekilde sınıflandırmasını sağlamak ve parametrelerini optimize etmek için en uygun algoritma olarak kullanılabilmektedir [127].

3.2.5. Naive Bayes (NB)

Her biri bilinen bir sınıfa ait olan ve her biri bilinen bir değişken vektörüne sahip olan bir dizi nesne göz önüne alındığında, gelecekteki nesneleri bir sınıfa atamaya izin verecek bir kural oluşturmaktır. Denetimli sınıflandırma sorunları olarak adlandırılan bu tür sorunlar her yerde mevcuttur ve bu tür kuralları oluşturmak için birçok yöntem geliştirilmiştir. Bu yöntemlerden en önemlilerinden biri Naive Bayes'dir. Bu yöntem birkaç nedenden dolayı önemlidir. Herhangi bir karmaşık yinelemeli parametre tahmin şemasına ihtiyaç duymadan, inşa edilmesi çok kolaydır. Bu, büyük veri setlerine kolayca uygulanabileceği anlamına gelmektedir. Yorumlaması kolaydır, böylece sınıflandırıcı teknolojisinde kullanıcılar sınıflandırmayı neden yaptığını anlayabilmektedir. Ve son olarak, çoğu zaman şaşırtıcı derecede iyi sonuç vermektedir. Herhangi bir uygulamada mümkün olan en iyi sınıflandırıcı olmasa da genellikle sağlam ve oldukça iyi bir performans sergileyeceğine güvenilmektedir.

Burada açıklama kolaylığı sağlamak için, i = 0, 1 olarak adlandırılan yalnızca iki sınıf varsayılmaktadır. Daha büyük puanların sınıf 1 nesneleriyle ilişkilendirileceği ve sınıf 0 nesneleriyle daha küçük puanların ilişkilendirileceği bir puan oluşturmak için bilinen sınıf üyeliklerine (eğitim seti) sahip ilk nesne kümesini kullanmaktır. Sınıflandırma daha sonra bu puanın bir eşik t ile karşılaştırılmasıyla elde edilmektedir. P(i|x) 'i ($x = x_1, \ldots, x_p$) ölçüm vektörüne sahip bir nesnenin sınıf i'ye ait olma olasılığı olarak tanımlanırsa, P (i|x) 'in herhangi bir monoton fonksiyonu uygun bir skor oluşturmaktadır. Temel olasılık bize P(i|x) 'i f(x|i) P (i) ile

orantılı olarak ayrıştırabileceğidir. Burada f(x|i), x'in i sınıfı nesneler için koşullu dağılımıdır ve P(i) hakkında daha fazla bir şey bilinmiyorsa bir nesnenin i sınıfına ait olma olasılığıdır (i sınıfının 'önceki' olasılığı). Sınıflandırmaları üretmek için Denklem 3.15'teki formül kullanılmaktadır.

$$\frac{P(1|x)}{P(0|x)} = \frac{f(x|1)\,P(1)}{f(x|0)\,P(0)} \tag{3.15}$$

Burada f (x|i) ve P(i) 'yi tahmin etmek gerekir. Eğitim seti genel popülasyondan rastgele bir örnek ise, P(i) doğrudan eğitim setindeki sınıf i nesnelerin oranından tahmin edilmektedir. Naive Bayes yöntemi F(x|i)'yi tahmin etmek için $f(x|i) = \prod_{j=1}^{p} f(x_j|i)$ x'in bağımsız olduğunu varsayar sonra $f(x_j|i), j = 1, \dots, p; i = 0{,}1$ tek değişkenli dağılımların her birini ayrı ayrı tahmin etmektedir. Böylece p boyutlu çok değişkenli problem, p tek değişkenli tahmin problemlerine indirgenmiştir. Tek değişkenli tahmin benzerdir, basittir ve doğru tahminler elde etmek için daha küçük eğitim seti gerektirmektedir. Bu, Naive Bayes yöntemlerinin özel ve gerçekten de benzersiz cazibelerinden biridir. Ayrıca basit, çok hızlı ve karmaşık yinelemeli tahmin şemaları gerektirmemektedir.

Eğer marjinal dağılımlar f(x_j|i) ayrıksa ve her x_j sadece birkaç değer alıyorsa, o zaman f(x_j|i) tahmini çok terimli bir histogram tipi tahmin edicidir. Eğer f (x_j|i) sürekli ise, o zaman ortak bir strateji her birini az sayıda aralığa ayırmak ve tekrar çok terimli tahminciyi kullanmaktır ancak sürekli tahminlere dayanan daha ayrıntılı versiyonlarda (ör. çekirdek tahminleri) kullanılmaktadır.

Bağımsızlık varsayımı denildiğinde Denklem 3.16'daki formül kullanılmaktadır.

$$\frac{P(1|x)}{P(0|x)} = \frac{\prod_{j=1}^{p} f(x_j|1)\;P(1)}{\prod_{j=1}^{p} f(x_j|0)\;P(0)} = \frac{P(1)}{P(0)} \prod_{j=1}^{p} \frac{f(x_j|1)}{f(x_j|0)} \tag{3.16}$$

Şimdi, sadece P(i|x) ile monoton olarak ilişkili bir skor üretmek gerektiğini, aşağıdaki logaritma monoton artan bir fonksiyondur ve bu alternatif bir skor vermektedir. $\omega_j = \ln(f(x_j|1)/f(x_j|0)$ ve bir sabit $k = \ln(P(1)/P(0))$ tanımlandığında Denklem 3.17 ve Denklem 3.18'deki gibi basit bir toplamdır ve böylece sınıflandırıcı özellikle basit bir yapıya sahiptir.

$$ln\frac{P(1|x)}{P(0|x)} = ln\frac{P(1)}{P(0)} + \sum_{j=1}^{p} ln\frac{f(x_j|1)}{f(x_j|0)} \tag{3.17}$$

$$ln\frac{P(1|x)}{P(0|x)} = k + \sum_{j=1}^{p} \omega_j \tag{3.18}$$

Naive Bayes modelinde örtük olarak her bir sınıf içinde x_j|'nin bağımsızlığı varsayımı aşırı derecede kısıtlayıcı görülebilmektedir. Aslında çeşitli faktörler devreye girebilir, bu da varsayımın göründüğü kadar zararlı olmadığıdır. İlk olarak, sınıflar arasındaki ayrıma benzer bir şekilde katkıda bulunma ihtimalleri olduğu gerekçesiyle yüksek düzeyde ilişkili değişkenlerin elimine edildiği bir önceki değişken seçim adımı sıklıkla gerçekleştirilmiştir. Bu, kalan değişkenler

arasındaki ilişkilerin bağımsızlıkla tahmin edilebileceğidir. İkinci olarak, etkileşimlerin sıfır olduğunu varsaymak örtük bir düzenleme adımı sağlamaktadır ve modelin varyansını azaltarak daha doğru sınıflandırma sağlamaktadır. Üçüncüsü, değişkenlerin ilişkilendirildiği bazı durumlarda, optimal karar yüzeyi bağımsızlık varsayımı altında üretilenle çakışırken varsayımda bulunmak performansa zarar vermemektedir. Dördüncüsü ise, Naive Bayes modelinin ürettiği karar yüzeyi aslında karmaşık doğrusal olmayan bir şekle sahiptir. Yüzey ω_j'de doğrusaldır ancak orijinal değişkenler x_j'de oldukça doğrusal değildir, böylece oldukça ayrıntılı yüzeylere sığmaktadır.

Naive Bayes modeli sadeliği, zarafeti ve sağlamlığı nedeniyle son derece çekicidir. En eski resmi sınıflandırma algoritmalarından biridir ve yine de en basit haliyle bile çoğu zaman şaşırtıcı derecede etkilidir. Metin sınıflandırması ve spam filtreleme gibi alanlarda yaygın olarak kullanılmaktadır. İstatistiksel, veri madenciliği, makine öğrenimi ve örüntü tanıma toplulukları tarafından onu daha esnek hale getirmek amacıyla çok sayıda değişiklik yapılmıştır, ancak bu tür değişikliklerin zorunlu olarak karmaşıklıkları olduğunu ve bu değişikliklerin temelinden uzaklaştığını kabul etmek gerekmektedir [128].

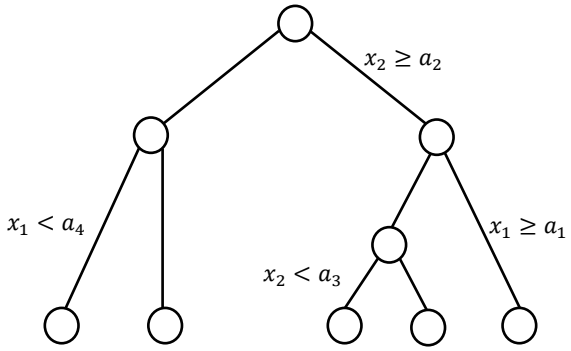
3.2.6. Rastgele Orman (Random Forest RF)

Rastgele Orman, ağaç işlemeyi kesmeden her biri tam olarak büyüyen birkaç karar ağacından oluşmaktadır. Ne kadar ağaç olursa o kadar doğru sonuç vermektedir ve aşırı uyum oluşmamaktadır. Rastgele orman algoritması genel tahmin yapmaktadır ve otomatik özellik seçimi vb. avantajına sahiptir. Dolayısıyla çözülmesi gereken aşağıdaki problemler vardır.

1) Küçük sınıf için değiştirme algoritması ile bir önyükleme örneklemesi tasarlamak ve ardından azınlık sınıf örneklerinin en yakın ana sınıf komşusunu bulmaktır.
2) Her örnekleme grubu için rastgele orman sınıflandırıcısını oluşturmak, her sınıflandırıcı yalnızca aynı düğümde gönderilen eğitim veri kümesini işlemek, ardından test veri kümesini tahmin etmek için modeli paralel olarak kullanarak tahmin sonuçlarını toplamaktır.

Yukarıdaki problemleri çözmek için global bir bölümleme algoritması tasarlamak, aynı anahtara sahip veri öğelerini aynı bölüme koymak ve güvenilmesi gereken bölümleri aynı düğüme yerleştirmek gerekmektedir. Topluluk rasgele orman algoritması, başlangıç veri kümesini birkaç kez örneklemek zorundadır ayrıca rasgele orman oluşturmak ve kullanıcı özellik bilgisi kazanımını teşvik etmek için bir yineleme işlemidir [129].

Ağaç temelli bir model, önceden belirlenmiş bir durdurma koşulu karşılanana kadar belirli bir kritere göre verilen veri setini yinelemeli olarak iki gruba ayırmayı içermektedir. Karar ağaçlarının altında yaprak düğümleri veya yapraklar vardır. Şekil 3.5'te bölümlere ayrılmış alt uzayların bir grafik temsilidir. Burada ilk bölünme $x_2 \geq a_2$'de meydana gelmektedir. Sonra, iki alt uzay tekrar bölümlenmektedir. Sol dal $x_1 \geq a_4$'e, sağ dal ilk önce $x_1 \geq a_1$'de bölünmekte ve alt dallarından biri $x_2 > a_3{}'$e bölünmektedir.



Şekil 3.5. Rastgele orman ağaç dağılımı

Bölme ve durdurma kriterlerinin nasıl belirlendiğine bağlı olarak, karar ağaçları hem sınıflandırma görevleri (örneğin, lojistik regresyon gibi kategorik sonuç) hem de regresyon görevleri (sürekli sonuç) için tasarlanmaktadır. Hem sınıflandırma hem de regresyon problemleri için, bir dahili düğümü bölmek üzere seçilen tahmin değişkenlerinin alt kümesi, bir optimizasyon problemi olarak formüle edilen önceden belirlenmiş bölme kriterlerine bağlıdır. Sınıflandırma problemlerinde yaygın bir bölme kriteri, rasgele bir değişkenin bit temsilinin uzunluğu üzerindeki alt sınırı belirten Shannon'ın kaynak kodlama teoreminin pratik uygulaması olan entropidir.

Rastgele orman modeli, ağaç temelli bir öğrenme algoritmasıdır; diğer bir deyişle algoritma birçok bağımsız ağaç üzerindeki tahminlerin ortalamasını almaktadır. Tek tek ağaçlar, orijinal örnek yerine önyükleme örnekleri üzerine inşa edilmiştir. Bu önyükleme toplama veya basitçe torbalama olarak adlandırılır ve fazla uydurmayı azaltmaktadır. Bireysel karar ağaçları kolayca yorumlanabilir, ancak bu yorumlanabilirlik rastgele ormanlarda kaybolmaktadır çünkü birçok karar ağacı bir araya toplanmaktadır. Bununla birlikte karşılığında, rastgele ormanlar genellikle tahmin görevlerinde çok daha iyi performans göstermektedir. Rastgele orman algoritması, karar ağaçlarına kıyasla hata oranını daha doğru tahmin etmektedir. Daha spesifik olarak hata oranının, ağaç sayısı arttıkça her zaman yakınsadığı matematiksel olarak kanıtlanmıştır. Rastgele ormanın hatası, eğitim süreci sırasında torba dışı hatasıyla tahmin edilmektedir. Her ağaç farklı bir önyükleme örneği üzerine inşa edilmiştir. Her önyükleme örneği rastgele gözlemlerin yaklaşık üçte birini dışarıda bırakmaktadır. Belirli bir ağaç için bu dışarıda bırakılan gözlemlere torba dışı hata örneği adı verilir. Düşük bir torba dışı hatası üretecek parametrelerin bulunması, genellikle model seçiminde ve parametre ayarlamasında önemli bir husustur. Rastgele orman algoritmasında, yordayıcı değişkenlerin alt kümesinin boyutu ağaçların nihai derinliğini kontrol etmek için önemlidir [130].

3.2.7. Karar Ağaçları (Decision Tree DT)

Karar ağaçlarında, bilgi kazanma yaklaşımı genellikle oluşturulan bir karar ağacının her düğümü için uygun özelliğini belirlemede kullanılmaktadır. Böylece mevcut düğümün test niteliği

olarak en yüksek bilgi kazancına sahip özniteliği (maksimum entropi azalması) seçilebilmektedir. Bu şekilde bölümlemeden elde edilen eğitim örneği, alt kümesini sınıflandırmak için gereken bilgiler en küçük olmaktadır. Diğer bir deyişle mevcut düğümde bulunan numune setini bölümlemek için bu özelliğin kullanılması üretilen tüm numune alt kümeleri için farklı türlerin karışım derecesini minimuma indirmektedir. Bu nedenle böyle bir bilgi teorisi yaklaşımının kullanılması nesne sınıflandırmasının gerekli bölme sayısını etkili bir şekilde azaltmaktadır. S tip özniteliği, C_i i $(1,2,3,\dots,m)$ tipine karşılık gelen m sayıda potansiyel farklı değerler alabilen veri örneklerinin sayısını içeren settir. C_i'nin örnek sayısı olduğu varsayıldığında daha sonra, belirli bir veriyi sınıflandırmak için gerekli bilgi miktarı Denklem 3.19'daki gibi ifade edilmektedir.

$$I(S_1, S_2, \dots, S_m) = -\sum_{i=1}^{m} p_i \log(p_i) \tag{3.19}$$

Burada $P_i = S_{ij} / |S_j|$ veri örneklerinin herhangi bir alt kümesinin C_i kategorilerine ait olma olasılığıdır. A'nın v farklı değerlere $\{a_1, a_2, \dots, a_v\}$ sahip bir özellik olduğu varsayılmaktadır. A'nın özelliği kullanılarak S, v sayıda alt kümeye $\{S_1, S_2, \dots, S_v\}$ bölünmektedir. Burada S_j özniteliği S kadar a_j'ye eşit olan veri örneklerini içermektedir. Test için özellik olarak A özelliği seçilirse, yani numune seti için bölümler yapmak için kullanılırsa, S_{ij}'nin S_i altkümesinde C_i tipi bir örnek set olduğu varsayıldığında, gerekli bilgi entropisi Denklem 3.20'deki gibi ifade edilmektedir.

$$E(A) = \sum_{j=1}^{v} \frac{(S_{1j} + S_{2j} + \dots + S_{mj})}{S} I(S_{1j}, \dots, S_{mj}) \tag{3.20}$$

Mevcut dal düğümünde A özelliğinin bu tür kullanımı, elde edilen bilgi kazancına karşılık gelen set bölümleme örnekleridir ve Denklem 3.21'deki formülle ifade edilmektedir.

$$Gain(A) = (s_1, s_2, \dots, s_m) - E(A) \tag{3.21}$$

Verilen veri setlerine göre bir karar ağacı oluşturmak için;

- Girdi: Eğitim örnekleri, her bir öznitelik ayrı bir değer almaktadır ve tümevarım için kullanılabilen bir aday öznitelik kümesi öznitelik listesidir.
- Çıktı: Bir karar ağacıdır.
- Bir düğüm N oluşturulur.
- Düğümün tüm örnekleri aynı C kategorisindeyse, o zaman bir yaprak düğüm olarak N döndürülmektedir ve C kategorisi ile işaretlenmektedir (başlangıç kök düğümü tüm eğitim örneklerine karşılık gelir).
- Öznitelik listesi boşsa N yaprak düğüm olarak döndürülmektedir ve düğümü örnekleri en fazla kategori içeren bir tür olarak işaretlenmektedir.
- Öznitelik listesinden en büyük bilgi kazanımına sahip bir test özniteliği seçilmektedir ve düğüm N test özniteliği ile işaretlenmektedir.

- Test özniteliğinin verilen her bir a_i değeri için, N düğümünde bulunan örnek kümesi kısımlara ayrılmıştır.
- Test özniteliği = a_i durumuna göre, test koşullarını belirtmek için N düğümünden karşılık gelen bir dal üretilmektedir.

Kurulan S_i, Test özniteliği = a_i koşulu altında elde edilen örnek settir. Boşsa ilgili yaprak düğümü en çok sayıda örnek türünü içeren kategori ile işaretlenmektedir. Aksi takdirde bir dönüş değeri ile işaretlenerek karar ağacı oluşturulmaktadır (S_i , öznitelik listesi – test özniteliği).

Bu, bir karar ağacı oluşturmak için yukarıdan aşağıya bölen ve ele geçiren yinelemeli bir şekilde kullanan açgözlü bir algoritmadır. Özyinelemenin sonlandırma koşulu şudur: bir düğüm içindeki tüm örnekler aynı kategoridedir. Mevcut numune setini bölmek için hiçbir öznitelik kullanılamıyorsa, oylama ilkesi onu zorunlu yaprak düğümü yapmak ve en çok sayıda numune türüne sahip kategorisi ile işaretlemek için kullanılmaktadır. Test özniteliği = a_i koşulunu karşılayan bir örnek yoksa bir yaprak düğüm oluşturulmaktadır ve onu en çok sayıda örnek türüne sahip kategorisi ile işaretlenmektedir [131].

Karar ağaçlarının avantajları; anlaması, yorumlaması ve görselleştirmesi basittir. Dolaylı olarak değişken tarama veya özellik seçimi gerçekleştirmesidir. Hem sayısal hem de kategorik verileri işleyebilmektedir. Çoklu çıktı sorunlarını da çözebilmektedir. Veri hazırlama için kullanıcılardan nispeten az çaba gerektirmektedir. Parametreler arasındaki doğrusal olmayan ilişkiler, ağaç performansını etkilememektedir. Dezavantajları ise; öğrenenler, verileri iyi bir şekilde genelleştiremeyen aşırı karmaşık ağaçlar oluşturabilmesidir. Buna aşırı uyum denir. Kararsız olabilir çünkü verilerdeki küçük değişiklikler tamamen farklı bir ağacın oluşturulmasına neden olabilmektedir. Buna, torbalama ve güçlendirme gibi yöntemlerle azaltılması gereken varyans denir. Açgözlü algoritmalar, küresel olarak en uygun karar ağacını döndürmeyi garanti edememektedir. Bu, özelliklerin ve örneklerin değiştirilerek rastgele örneklendiği birden fazla ağaç eğitilerek azaltılabilmektedir. Karar ağacı öğrenme, bazı sınıflar hakimse önyargılı ağaçlar oluşturmaktadır ve karar ağacına uydurmadan önce veri setinin dengelenmesi önerilmektedir [132].

3.2.8. eXtreme Gradient Boosting (XGB)

XGBoost, doğru modeller elde etmek için artırma adı verilen bir süreci uygulayan denetimli bir öğrenme algoritmasıdır. Denetimli öğrenme, bir dizi etiketli eğitim örneğinden tahmine dayalı bir modelin çıkarılması görevini ifade etmektedir. Bu tahmine dayalı model daha sonra yeni olmayan örneklere uygulanabilmektedir. Algoritmanın girdileri eğitim örneği çiftleridir $(\overrightarrow{x_0}, y_0), (\overrightarrow{x_1}, y_1) \ldots (\overrightarrow{x_n}, y_n)$ $(\overrightarrow{x_0}, y_0)$ burada $\vec{x}$, örneği açıklayan özelliklerin bir vektörü ve y ise etiketidir. Denetimli öğrenme, yeni giriş örneklerini doğru şekilde etiketleyecek bir $F(\vec{x}) = y$ işlevini öğrenmek olarak görülmektedir.

Denetimli öğrenme, sınıflandırma veya regresyon problemlerini çözmek için kullanılmaktadır. Sınıflandırma problemlerinde y etiketi ayrı bir değer almaktadır. Örneğin, $\vec{x}$ ile temsil edilen sıcaklık veya zaman gibi üretim sürecinden kaydedilen niteliklere dayalı bir üretim hatasının meydana gelip gelmediğini tahmin etmek istenmektedir. Regresyon problemlerinde hedef etiket y sürekli bir değer almaktadır. Bu, belirli bir gündeki sıcaklığı veya nemi tahmin etmek gibi bir sorunu çerçevelemek için kullanılmaktadır. XGBoost, özünde bir karar ağacı geliştirme algoritmasıdır. Artırma, her yeni modelin önceki modeldeki eksiklikleri gidermeye çalıştığı, birçok modeli sırayla oluşturmaya yönelik toplu öğrenme tekniğini ifade etmektedir. Ağaçta, topluluğa eklenen her yeni model bir karar ağacıdır. Bir karar ağacı modelinin nasıl oluşturulacağı ve bunun XGBoost algoritması ile genelleştirilmiş gradyan artırmaya nasıl genişletilebileceğidir [133].

XGBoost, aşırı uyumla mücadele etmek için kullanılan bir düzenlilik terimini ve ayrıca keyfi farklılaştırılabilir kayıp işlevlerini destekleyen genelleştirilmiş bir gradyan artırma uygulamasıdır. Düz kare hata kaybını optimize etmek yerine, iki bölümden oluşan bir hedef fonksiyon tanımlanmaktadır. Eğitim seti üzerinde bir kayıp fonksiyonu ve modelin karmaşıklığını cezalandıran bir düzenleme terimi Denklem 3.22'teki gibidir.

$$\mathcal{L}(\phi) = \sum_i l(\hat{y}_i, y_i) + \sum_t \Omega(f_t) \tag{3.22}$$

$\mathcal{L}(y_i, \hat{y}_i)$, belirli bir eğitim örneği için tahmin ve gerçek etiket arasındaki farkı ölçen herhangi bir dışbükey türevlenebilir kayıp işlevi olabilir. $\Omega(f_t)$ ağaç f_t'nin karmaşıklığını tanımlar ve XGBoost algoritmasında Denklem 3.23'teki formülle tanımlanır.

$$\Omega(f) = \gamma T + \frac{1}{2}\lambda||\omega||^2 \tag{3.23}$$

Burada t, f_t ağacının yaprak sayısıdır ve w yaprak ağırlıklarıdır (yani, yaprak düğümlerinde depolanan tahmin edilen değerler). $\Omega(f_t)$ objektif fonksiyona dahil edildiğinde, aynı anda $L(y_i, \hat{y}_i)$'yi en aza indiren daha az karmaşık bir ağaç için optimize etmeye zorlanmaktadır. Bu aşırı uyumu azaltmaya yardımcıdır. γt, her ilave ağaç yaprağı için sabit bir ceza sağlar ve $\lambda\omega^2$ aşırı ağırlıkları cezalandırmaktadır. γ ve λ kullanıcı tarafından yapılandırılabilir parametrelerdir.

Arttırmanın yinelemeli bir şekilde ilerlediğini göz önünde bulundurarak mevcut yineleme için amaç işlevini, en yeni ağaç f_t tarafından ayarlanan önceki yinelemenin $\hat{y}_i^{(t-1)}$ tahmini cinsinden Denklem 3.24'teki formülle ifade edilmektedir.

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \tag{3.24}$$

Daha sonra hedefi en aza indiren f_t'yi bulmak için optimizasyon yapılmaktadır. Yukarıdaki fonksiyonun Taylor açılımını ikinci sıraya almayı ve farklı kayıp fonksiyonlarını kolayca barındırmayı sağlar ve Denklem 3.25'teki gibi ifade edilmektedir.

$$\mathcal{L}^{(t)} \cong \sum_{i=1}^{n}[l(y_i, \hat{y}_i^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i {f_t}^2(x_i)] + \Omega(f_t) \quad (3.25)$$

Burada $g_i = \partial_{\hat{y}^{(t-1)}} l(y_i, \hat{y}^{(t-1)})$ ve $h_i = \partial^2_{\hat{y}^{(t-1)}} l(y_i, \hat{y}^{(t-1)})$ kayıp fonksiyonunda birinci ve ikinci derece gradyan istatistikleridir. T adımında aşağıdaki basitleştirilmiş hedefi elde etmek için sabit terimler kaldırılabilmektedir.

Sabitlerin kaldırıldığı basitleştirilmiş amaç işlevi Denklem 3.26'daki gibidir.

$$\tilde{L}^{(t)} = \sum_{i=1}^{n} g_i f_t(x_i) + \frac{1}{2} h_i {f_t}^2(x_i)] + \Omega(f_t) \quad (3.26)$$

$I_i = \{i|q(x_i) = j\}$, j yaprağının örnek kümesi olarak tanımlanmaktadır. Ω Denklem 3.27'deki gibi genişletilerek yeniden yazılmaktadır.

$$\tilde{\mathcal{L}}^{(t)} = \sum_{i=1}^{n}[g_i f_t(x_i) + \frac{1}{2} h_i {f_t}^2(x_i)] + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^{T} \omega_j^2 =$$
$$\sum_{j=1}^{T}[(\sum_{i \in I_j} g_i)\, \omega_j + \frac{1}{2}(\sum_{i \in I_j} h_i + \lambda)\, \omega_j^2] + \gamma T \quad (3.27)$$

Sabit bir q(x) yapısı için, j yaprağının optimal ağırlığı ω_j^* Denklem 3.28'deki gibidir.

$$\omega_j^* = \frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda'} \quad (3.28)$$

Karşılık gelen optimum değeri ise Denklem 3.29'daki formülle hesaplanmaktadır.

$$\hat{\mathcal{L}}^{(t)}(\mathrm{q}) = -\frac{1}{2} \sum_{j=1}^{T} \frac{(\sum_{i \in I_j} g_i)^2}{\sum_{i \in I_j} h_i + \lambda} + \gamma T \quad (3.29)$$

Bir ağaç yapısının q kalitesini ölçmek için bir puanlama fonksiyonu olarak kullanılmaktadır. Bu puan, daha geniş bir hedef işlevler yelpazesi için türetilmesi dışında, karar ağaçlarını değerlendirmek için safsızlık puanı gibidir. Normalde tüm olası q ağaç yapılarını saymak imkansızdır. Bunun yerine tek bir yapraktan başlayan ve yinelemeli olarak ağaca dallar ekleyen açgözlü bir algoritma kullanılmaktadır. I_L ve $I_R{}'$nin, bölünmeden sonra sol ve sağ düğümlerin örnek kümeleri olduğunu varsayılmaktadır. $I = I_L \cup I_R$ olduğunda, daha sonra bölünmeden sonra kayıp azalması Denklem 3.30'daki şekilde ifade edilmektedir. Bu, genellikle uygulamada bölünmüş adayları değerlendirmek için kullanılmaktadır.

$$\mathcal{L}_{split} = \frac{1}{2}[\frac{(\sum_{i \in I_L} g_i)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{(\sum_{i \in I_R} g_i)^2}{\sum_{i \in I_R} h_i + \lambda} - \frac{(\sum_{i \in I} g_i)^2}{\sum_{i \in I} h_i + \lambda}] - \gamma \quad (3.30)$$

Ağaç öğrenmedeki temel sorunlardan biri en iyi bölünmeyi bulmaktır. Bunun için, bir bölünmüş bulma algoritması, özellikler üzerindeki olası tüm bölünmeleri numaralandırmaktadır. Buna tam açgözlü algoritma denir. Sürekli özellikler için tüm olası bölmeleri saymak sayısal olarak

zordur. Bunu verimli bir şekilde yapmak için, algoritmanın önce verileri özellik değerlerine göre sıralaması ve gradyan istatistiklerini biriktirmek için verileri sıralı şekilde ziyaret etmesi gerekir.

Algoritmalar ya yalnızca yoğun veriler için optimize edilmiştir ya da kategorik kodlama gibi sınırlı durumları ele almak için özel prosedürlere ihtiyaç duymaktadır. XGBoost, tüm seyreklik modellerini birleşik bir şekilde ele almaktadır. Daha önemlisi, hesaplama karmaşıklığını girdideki eksik olmayan girişlerin sayısına göre doğrusal hale getirmek için seyreklikten yararlanmasıdır.

Önbellek erişim düzenleri, veri sıkıştırma ve parçalama ağacın güçlendirilmesi için ölçeklenmektedir. Bir uçtan uca sistem oluşturmak için temel unsurlar olduğunu göstermektedir. Bu diğer makine öğrenimi sistemlerine de uygulanabilmektedir. Bu iç görüleri birleştiren XGBoost, minimum miktarda kaynak kullanarak gerçek dünya ölçeğindeki sorunları çözebilmektedir [134].

3.2.9. Derin Öğrenme

Derin öğrenme, bilgisayarların dünyayı kavramlar hiyerarşisi açısından anlamasını sağlayan ve deneyimleyerek öğrenen bir makine öğrenimi biçimidir. Bilgisayar deneyimden bilgi topladığı için insanın, bilgisayarın ihtiyaç duyduğu tüm bilgileri resmi olarak belirtmesine gerek yoktur. Kavramlar hiyerarşisi, bilgisayarın karmaşık kavramları daha basit olanlardan oluşturarak öğrenmesine olanak tanımaktadır; bu hiyerarşilerin grafiği birçok katman derinliğinde olmasıdır. Derin öğrenme, endüstride ve çeşitli platformlarda, araştırma yapan ve kariyer planlayan lisans veya lisansüstü öğrenciler ve yazılım mühendisleri tarafından kullanılmaktadır. Bir web sitesi hem okuyucular hem de eğitmenler için ek materyaller sunmaktadır. Derin öğrenmenin, bilgisayarla görme, robotik, biyoinformatik ve kimya, video oyunları, arama motorları, konuşma ve ses işleme, doğal dil işleme, çevrimiçi reklamcılık ve finans dahil birçok yazılım alanlarında faydalı olduğu kanıtlanmıştır [135].

Derin öğrenme, birden çok işleme katmanından oluşan hesaplama modellerinin, birden çok soyutlama düzeyiyle verilerin temsillerini öğrenmesine olanak tanımaktadır. İlaç keşfi ve genomik, görsel nesne tanıma, nesne algılama ve konuşma tanıma gibi diğer birçok alanda, bu yöntemler sayesinde teknoloji önemli ölçüde gelişmiştir. Derin öğrenme, geri yayılma algoritmasını kullanarak büyük veri setindeki karmaşık yapıyı ortaya çıkarmaktadır. Bunu, bir makinenin önceki katmanındaki tasarımından her katmandaki tasarımı hesaplamak için kullanılan iç parametrelerini nasıl değiştirmesi gerektiğini göstererek yapmaktadır. Tekrarlayan ağlar metin ve konuşma gibi sıralı verilerde kullanılırken ön plana çıkarken, Derin Evrişimli Ağlar ise ses, görüntü, video ve konuşmaların işlenmesinde atılım sağlamıştır. Temsili öğrenme, ham veriyi beslemek ve tespit etme veya sınıflandırma için gereken temsilleri otomatik keşfetmek için makineye imkân veren bir dizi yöntemdir. Derin öğrenme yöntemleri, basit fakat yüksek oranda temsil edilen ve birazda soyut düzeyde lineer olmayan modüller oluşturularak elde edilen fazla sayıda temsile sahip temsili öğrenme metotlarıdır. Çok karmaşık işlevler de böyle dönüşümlerin birleşmesiyle öğrenilmektedir.

Daha yüksek temsil katmanlarını ayırmada önemli olan girdinin yönlerini güçlendirmek, alakasız çeşitlemeleri önlemek ve bunu sınıflandırma görevleri için yapmaktır. Mesela bir dizi piksel olarak gelen bir görüntü olduğunda, ilk temsil katmanındaki öğrenilen öznitelikler genellikle görüntüdeki belli konum ve yönlerde kenarların varlığını ya da yokluğunu temsil etmektedir. Sonraki katman genellikle kenarları belirli düzenlerde işaretleyerek, kenar pozisyonlarındaki küçük farkları dikkate almadan motifleri belirlemektedir. Motifleri, benzer nesnelerin parçalarına karşılık gelen daha büyük birleşimlerde bir araya getiren ise üçüncü katmandır ve son katmanda bu parçaların birleşimleri olarak nesneler tespit edilmektedir. Derin öğrenmedeki anahtar kavram, bu öznitelik katmanlarının mühendisler tarafından dizayn edilememesidir. Derin öğrenme yöntemlerinin en önemli avantajı literatürdeki zorlu problemleri çözmesidir. Bu yapay zekâ topluluklarının yıllardır çözmeye uğraştıkları problemlerdir. Çok büyük sayıda ve boyutu büyük karmaşık verileri analiz etmede gayet başarılı olduğu ortaya çıktığından beri bilimde, kurumlarda ve iş hayatında yoğun şekilde kullanılmaktadır. Ayrıca derin öğrenme yöntemleri konuşma ve görüntü tanımada çok iyi sonuçlar verirken aynı zamanda parçacık ivmelendirici verilerini analiz etmede, beyin devrelerini yeniden oluşturmada, ilaç moleküllerinin hareketliliğini tahmin etmede ve kodlamayan DNA'daki mutasyonların gen ekspresyonu ve hastalık üzerindeki etkilerini tahmin etmede makine öğrenmesi yöntemlerinden çok daha başarılı sonuçlar vermiştir. Derin öğrenme, doğal dil işleme, çeviri, duygu tespiti ve sorulara cevap verme gibi başlıca konulardaki problemlerde umut vadeden sonuçlara ulaşmıştır [136].

Derin öğrenmenin teorik temelleri, klasik sinir ağı literatüründe kök salmıştır. Ancak sinir ağlarının daha geleneksel kullanımından farklı olarak derin öğrenme, birçok gizli nöron ve katmanın (tipik olarak ikiden fazla) yeni eğitim paradigmalarıyla birleştirilmiş mimari bir avantaj olarak kullanılmasını açıklamaktadır. Birçok nörona başvurmak eldeki ham verilerin kapsamlı bir şekilde kapsanmasına izin verirken, çıktılarının doğrusal olmayan kombinasyonunun katman katman ardışık düzeni girdi uzayının daha düşük boyutlu bir projeksiyonunu oluşturmaktadır. Her düşük boyutlu projeksiyon daha yüksek bir algısal seviyeye karşılık gelmektedir. Ağın optimum şekilde ağırlıklandırılması koşuluyla, ham verilerin veya görüntülerin etkili bir üst düzey soyutlamasıyla sonuçlanmaktadır. Bu yüksek seviyeli soyutlama aksi takdirde el yapımı veya ısmarlama özellikler gerektiren otomatik bir özellik seti oluşturmaktadır. Derin öğrenmenin çeşitli metodolojik varyantları arasında popülerlikte birkaç mimari göze çarpmaktadır. Evrişimli Sinir Ağları, evrişimli filtreler uygulayan ve ardından indirgeme, düzeltme veya havuzlama katmanları uygulayan aralıklı ileri beslemeli katman kümesi olarak tanımlanmaktadır. Ağdaki her katman, yüksek seviyeli soyut özelliği ortaya çıkarmaktadır. Bu biyolojik olarak esinlenmiş mimari, görsel korteksin görsel bilgileri alıcılar biçiminde özümsediği prosedüre benzerdir. Diğer makul mimariler arasında, Derin İnanç Ağları gibi kısıtlı Boltzmann Makinelerinin, Derin Otomatik Kodlayıcılar olarak işlev gören Yığınlanmış Otomatik Kodlayıcıların, Derin Sinir Ağları gibi birçok katmanla

Yapay Sinir Ağları genişleten veya Tekrarlayan Sinir Ağları olarak yönlendirilmiş çevrimlerdir. Veri miktarı arttığında, makine öğrenimi teknikleri performans açısından yetersiz kalırken derin öğrenme doğruluk açısından daha iyi performans vermektedir. Derin öğrenmenin makine öğreniminden farkı ise makine öğrenimi derin öğrenmeyi kapsamasıdır. Özellikler, Makine öğreniminde manuel olarak verilirken derin öğrenme ise doğrudan verilerden öğrenmektedir [137].

3.2.10. Yapay Sinir Ağları (Artificial Neural Network ANN)

Sinir ağları bir bilgisayarın, insanların öğrendikleri şeyleri nasıl öğrendiklerine benzer şekillerde örneklerden öğrenmeye çalışmasına izin veren bir dizi yöntemdir. Sinir ağlarının tasarlanma şekli insan beynine benzer ve her katman sınıflandırma için gerekli bazı yararlı özellikleri çıkarmaktadır. İyi tahmin edici olan modellerle sonuçlanabilir ancak bunlar genellikle lojistik regresyon modellerinden çok daha karmaşıktır. Genel uygulamada ikili sınıflandırma için bir sinir ağı modeli, lojistik regresyon modelinden daha kötü olabilir çünkü sinir ağlarının eğitilmesi daha zordur ve LR gerilemeye göre aşırı uyuma karşı daha savunmasızdır. Ancak tamamen sınıflandırma uygulamamız gereken probleme ve veri kümesine bağlıdır. Sinir ağı temelde logistik regresyonu alıp en az iki kere tekrarlamaktır. LR'de girdi ve çıktı katmanları vardır fakat sinir ağlarında giriş ve çıkış katmanı arasında en az 1 gizli katman bulunmaktadır. Burada gizli katman denmesinin sebebi gizli katman eğitim setindeki gidileri görmemesidir [138]. YSA'daki temel unsur bir nörondur. Matematiksel terimlerle, bir nöron k, Denklem 3.31 ve Denklem 3.32'deki gibi tanımlanmaktadır.

$$\upsilon_k = \sum_{j=1}^{p} W_{kj} X_j \tag{3.31}$$

$$y_k = \phi\left(\upsilon_k - \theta_k\right) \tag{3.32}$$

Burada $x_1, x_2, \ldots, x_p$ giriş sinyalleridir; $w_{k1}, w_{k2}, \ldots, w_{kp}$ k nöronunun sinaptik ağırlıklarıdır ve υ_k doğrusal birleştirici çıktısıdır, θ_k ise eşiği gösterir. Ayrıca, $\phi(.)$ aktivasyon işlevidir ve y_k nöronun çıkış sinyalidir. Her nöron, bu bağlantıların güçlü yönlerini temsil eden çeşitli bağlantı katsayıları ile bazı komşularına bağlıdır [139].

Öğrenme, nöronların daha sonra katmanlara ayrılması için güçlerinin ayarlanmasıyla gerçekleştirilmektedir. Giriş katmanı, dış ortamdan girdi alan nöronlardan oluşmaktadır. Çıktı katmanı, sistemin çıktısını kullanıcıya veya dışarıdan çevreye ileten nöronlardan oluşmaktadır. Bu iki katman arasında genellikle birkaç gizli katman vardır. YSA modelinin gizli katmanı, girdi ve çıktı arasındaki ilişkiyi birbirine bağlayan bir kara kutu görevi görmektedir. Girdi ve çıktı değişkenleri arasındaki ilişki doğrusal olmadığında gizli katman, daha yüksek seviyeli özelliklerin çıkarılmasına yardımcı olurken çıktıların genelleştirilmesini kolaylaştırmaktadır. İki katmanlı sinir ağı Şekil 3.6'da gösterilmiştir [140]. Birkaç aktivasyon işlevi dikkate alınmaktadır [141].

Kimlik işlevi Denklem 3.33'teki formülle ifade edilmektedir.

$$\phi(x) = \quad x \tag{3.33}$$

Sigmoid işlevi (veya lojistik) Denklem 3.34'teki formülle ifade edilmektedir.

$$\phi(x) = \quad \frac{1}{1+exp(-x)} \tag{3.34}$$

Hiperbolik tanjant işlevi (tanh) Denklem 3.35'teki formülle ifade edilmektedir.

$$\phi(x) = \quad \frac{exp(x) - exp(-x)}{exp(x) + exp(-x)} = \frac{exp(2x) - 1}{exp(2x) + 1} \tag{3.35}$$

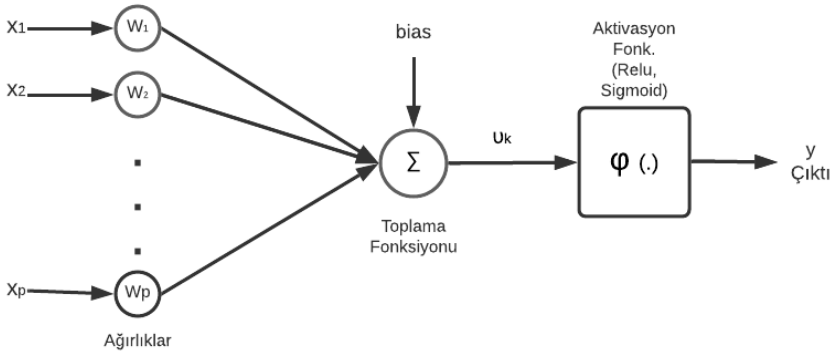
Zor eşik işlevi Denklem 3.36'daki formülle ifade edilmektedir.

$$\phi(x) = \quad 1_{x \geq \beta} \tag{3.36}$$

Doğrultulmuş Doğrusal Birim (ReLU) aktivasyon işlevi Denklem 3.37'deki gibidir.

$$\phi(x) = \quad max(0, x) \tag{3.37}$$

Tarihsel olarak sigmoid, türevlenebilir olduğu ve değerleri [0, 1] aralığında tutmaya izin verdiği için en çok kullanılan aktivasyon fonksiyonudur. Yine de |x| 0'a yakın olmadığında gradyanı, 0'a çok yakın olduğu için sorunludur.



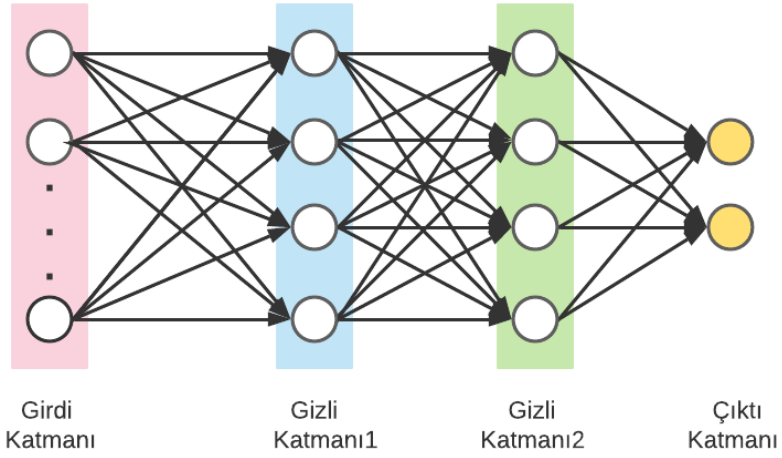
Şekil 3.6. Tek katmanlı YSA modeli

Çok sayıda katmana sahip sinir ağlarında (derin öğrenme için durum budur) bu, geri yayılım algoritmasının parametreyi tahmin etmesinde sorunlara neden olmaktadır. Bu nedenle sigmoid fonksiyonunun yerini düzeltilmiş lineer fonksiyon almıştır. Bu fonksiyon 0'da türevlenebilir değildir. Ancak pratikte bu gerçekten bir problem değildir, çünkü 0'a eşit bir girişe sahip olma

olasılığı genellikle sıfırdır. ReLU işlevi ayrıca bir seyreltme etkisine sahiptir. ReLU işlevi ve türevi negatif değerler için 0'a eşittir ve bu durumda böyle bir birim için hiçbir bilgi elde edilememektedir. Bu nedenle her birimin etkin olduğundan emin olmak için küçük bir pozitif önyargı eklenmektedir. ReLU fonksiyonunun çeşitli varyasyonları, tüm birimlerin kaybolmayan bir gradyanına sahip olduğundan ve x <0 için türevin 0'a eşit olmadığından emin olmak için kabul edilmektedir. Denklem 3.38'de ifade edilen formülde α, küçük bir pozitif değere ayarlanmış sabit bir parametre veya tahmin edilecek bir parametredir [141].

$$\phi(x) = \quad max(x, 0) \; + \; \alpha \, min(x, 0) \tag{3.38}$$

Çok katmanlı YSA'da ise bir katmanın bir nöronunun çıktısının bir sonraki katmanın bir nöronunun girdisi haline geldiği birkaç gizli nöron katmanından oluşan bir yapıdır. Dahası, bir nöronun çıktısı aynı katmandaki bir nöronun veya önceki katmanların nöronunun girdisi olabilmektedir. Çıktı katmanı olarak adlandırılan son katmanda, elimizdeki sorunların türüne (regresyon veya sınıflandırma) bağlı olarak gizli katmanlarda olduğu gibi farklı bir etkinleştirme işlevi uygulanmaktadır. Şekil 3.7'de üç giriş değişkenine, iki çıkış değişkenine ve iki gizli katmana sahip birçok katmanlı sinir ağını temsil etmektedir.



Şekil 3.7. Çok katmanlı YSA modeli

Çok katmanlı YSA'lar, bir katmanın her birimi bir sonraki katmanın tüm birimlerine bağlı olduğu için temel bir mimariye sahiptir ancak aynı katmanın nöronlarıyla bağlantı kurmaktadır. Mimarinin parametreleri, her katmandaki gizli katmanların ve nöronların sayısıdır. Aktivasyon fonksiyonları da kullanıcı tarafından seçilmelidir. Çıktı katmanı için, daha önce bahsedildiği gibi aktivasyon işlevi genellikle gizli katmanlarda kullanılandan farklıdır. Regresyon durumunda, çıktı

katmanına hiçbir aktivasyon fonksiyonu uygulanmayabilmektedir. İkili sınıflandırma için, çıktı bir ℙ(Y = 1/X) tahmini vermektedir, çünkü bu değer [0, 1] 'de olduğu için sigmoid aktivasyon fonksiyonu dikkate alınmaktadır. Çok sınıflı sınıflandırma için çıktı katmanı, ℙ(Y = i/X)'nin bir tahminini veren sınıf i başına bir nöron içermektedir. Tüm bu değerlerin toplamı 1'e eşit olmalıdır. Denklem 3.39'da gösterilen çok boyutlu fonksiyon softmax genellikle kullanılmaktadır [141].

$$softmax\,(\mathrm{z})_i = \frac{exp\,(z_i)}{\sum_j exp\,(z_j)} \tag{3.39}$$

L gizli katmanları olan çok katmanlı bir algılayıcının matematiksel formülasyonu $h^{(0)}(x) = x$ dir.

Burada $k = 1, \ldots, L$ (saklı katmanlar) için Denklem 3.40 ve Denklem 3.41'deki gibidir.

$$a^{(k)}(x) = b^{(k)} + W^{(k)} h^{(k-1)}(x) \tag{3.40}$$

$$h^{(k)}(x) = \phi(a^{(k)}(x)) \tag{3.41}$$

Burada $k = L + 1$ (çıktı katmanı) için Denklem 3.42 ve Denklem 3.43'teki gibidir.

$$a^{(L+1)}(x) = b^{(L+1)} + W^{(L+1)} h^{(L)}(x) \tag{3.42}$$

$$h^{(L+1)}(x) = \psi(a^{(L+1)}(x)) := f(x, 0) \tag{3.43}$$

Burada ϕ aktivasyon fonksiyonu ve ψ çıktı katmanı aktivasyon fonksiyonudur (örneğin çok sınıflı sınıflandırma için softmax). Her adımda, $W^{(k)}$ satır sayısı k katmanındaki nöron sayısı ve sütun sayısı k-1 katmandaki nöron sayısı olan bir matristir [141].

Gradyan Azalma işleminde ise ağırlıkları bireysel örnekler açısından aşamalı olarak güncellenmektedir. Bir eğitim örneği d verildiğinde, hata fonksiyonunu Denklem 3.44'teki gibidir.

$$\mathrm{e}_d(\vec{w}) = \frac{1}{2} \sum_{k \in \text{çıktılar}} (t_k - o_k)^2 \tag{3.44}$$

Buradaki fikir, bu örnekte hızlı bir hata azalması elde etmek için hata fonksiyonunun negatif gradyanına göre ağırlıkları değiştirmektir bu nedenle ağırlıkları gradyanlara göre yenilenmektedir. Denklem 3.45'teki formülle ifade edilmektedir.

$$\Delta w_{ji} = -\eta \frac{\partial E_d}{\partial w_{ji}}, w_{ji} = w_{ji} + \Delta w_{ji} \tag{3.45}$$

Burada η, gradyan aramasında adım boyutunu belirten öğrenme oranıdır [142]. Ağırlıkların güncellenmesi ise Denklem 3.46'daki şekilde ifade edilmektedir.

$$\frac{\partial E_d}{\partial w_{ji}} = \frac{\partial E_d}{\partial net_j} \frac{\partial net_j}{\partial w_{ji}} = \frac{\partial E_d}{\partial net_j} x_{ji} \tag{3.46}$$

$\delta_j = -\frac{\partial E_d}{\partial net_j}$, j biriminin hata terimi olarak ifade edildiğinde, ağırlık güncelleme kuralı $\Delta w_{ji} = \eta\, \delta_j x_{ji}$ şeklinde formüle edilmiştir [142].

3.2.11. Doğruluk, Kesinlik, Duyarlılık ve F1 Puanı

İkili bir sınıflandırma görevi gerçekleştiren bir model için birkaç farklı ölçüm vardır. Bu ölçümler arasında doğruluk, kesinlik, duyarlılık, yanlış pozitif oranı, F1 puanı ve eğri altındaki alan bulunur ve ölçümlerin birçoğunun birden fazla adı vardır. Tüm bu değerlendirme ölçütleri, hesaplanan tahmin edilen sınıfa karşı temel gerçeğe dayanan hata matrisinde bulunan dört değerden türetilmiştir. Bu değerler:

TP = Tahmin pozitif (normal) ve gerçek pozitif (normal),

FP = Tahmin pozitif (normal) ve gerçek negatif (anormal),

FN = Tahmin negatif (anormal) ve gerçek pozitif (normal),

TN = Tahmin negatif (anormal) ve gerçek negatif (anormal).

Doğruluk, doğru sınıflandırılmış örneklerin tüm öğelere oranıdır. Doğruluğun faydası, sınıflar dengesiz olduğunda daha düşüktür (yani, bir sınıftan diğerine göre önemli ölçüde daha fazla sayıda örnek vardır). Bununla birlikte, sınıflar dengelendiğinde yararlı bilgiler sağlar. Doğruluk skoru, algoritma çalıştırıldıktan sonra yapılan tahminlerin test verileri ile karşılaştırılmasıyla yapılan modelin performansını değerlendirmek için kullanılan bir yöntemdir. Bir tahminin gerçek değerlerle eşleşmesi için tahmin edilen tüm değerin oranına göre 0 ile 1 arasında bir değer üretilir [5]. Doğruluk hesaplaması ise Denklem 3.47'de gösterildiği şekilde yapılmaktadır.

$$Doğruluk = \frac{TP+TN}{TP+TN+FP+FN} \tag{3.47}$$

Kesinlik (p), x sınıfı olarak doğru bir şekilde sınıflandırılan öğelerin, x olarak sınıflandırılan tüm öğelere oranıdır ve Denklem 3.48'de verilmiştir.

$$p = \frac{TP}{TP+FP} \tag{3.48}$$

Duyarlılık (r), x olarak doğru şekilde sınıflandırılan öğelerin, gerçekte x sınıfı olan tüm öğelere oranıdır. Denklem 3.49'daki formülle hesaplanmaktadır.

$$r = \frac{TP}{TP+FN} \tag{3.49}$$

Negatif Öngörücü Değer (NPV), x değil olarak doğru şekilde sınıflandırılan öğelerin, x değil olarak sınıflandırılan tüm öğelere oranıdır ve Denklem 3.50'de verilmiştir.

$$NPV = \frac{TN}{TN+FN} \tag{3.50}$$

Özgüllük veya Gerçek Negatif Oran (TNR), x değil olarak doğru bir şekilde sınıflandırılan öğelerin, x sınıfı olmayan tüm öğelere oranıdır. Denklem 3.51'de verilmiştir.

$$TNR = \frac{TN}{TN+FP} \tag{3.51}$$

Yanlış Alarm Oranı (FAR) veya Yanlış Pozitif Oran (FPR), x sınıfı olarak yanlış sınıflandırılan öğelerin, x sınıfı olmayan tüm öğelere oranıdır. Denklem 3.52'deki formülle hesaplanmaktadır.

$$FPR = \frac{FP}{TN+FP} \tag{3.52}$$

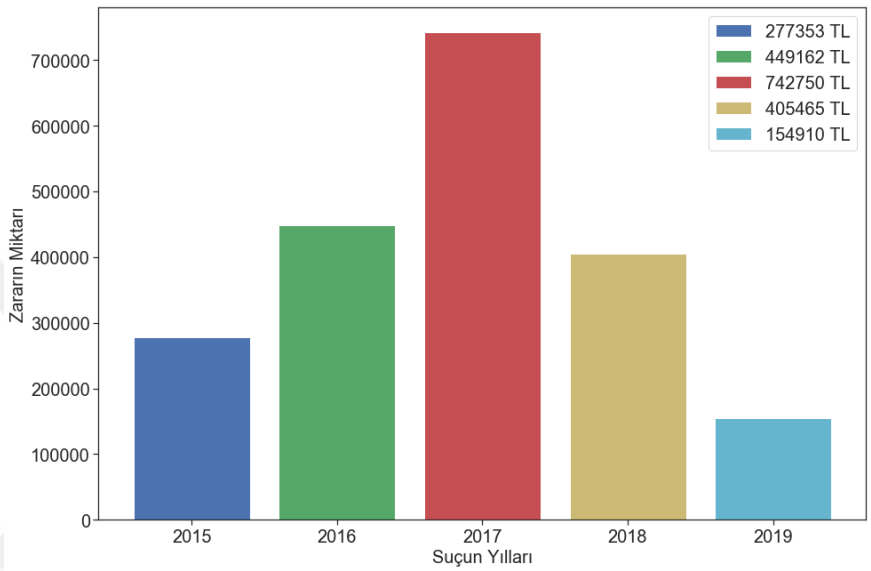
F1 Puanı (F1), hassasiyetin (p) ve gerçek pozitif oranın (duyarlılık) (r) harmonik ortalamasıdır. Denklem 3.53'te gösterildiği gibi üç farklı şekilde verilmiştir.

$$\mathrm{F1} = \frac{2TP}{2TP+FN+FP} = \frac{2}{\frac{1}{r}+\frac{1}{p}} = 2\frac{p*r}{p+r} \tag{3.53}$$

Bu, F-β fonksiyonunun kesinlik ve gerçek pozitif orana eşit önem verilen özel bir versiyonudur. Eğri altındaki alan (AUC), sınıflandırma eşiklerinin değiştirilmesiyle oluşturulan, yanlış pozitif orana karşı gerçek pozitif oranın bir grafiği olan bir alıcı çalışma özelliği (ROC) eğrisi altındaki alanın toplamıdır. Çok sınıflı problemler için doğruluk kolaylıkla hesaplanabilir ancak kesinlik, duyarlılık, FPR, F1 Puanı ve AUC gibi metrikler doğrudan hesaplanamamaktadır (ör. TP, TN, üç sınıflı problemler için mevcut değildir). Duyarlılık, geri çağırma vb. üç sınıflı bir problem için, problemi iki sınıflı bir probleme (yani, her sınıf için metriklerin hesaplandığı) bir soruna indirgeyerek belirlenmektedir. Genellikle, çok sınıflı problemler için yalnızca doğruluk kullanılmaktadır [5]. Önerilen model için diğer değerlendirme ölçütleri doğruluk, kesinlik, duyarlılık ve F1 skorudur.

4. BULGULAR VE TARTIŞMA

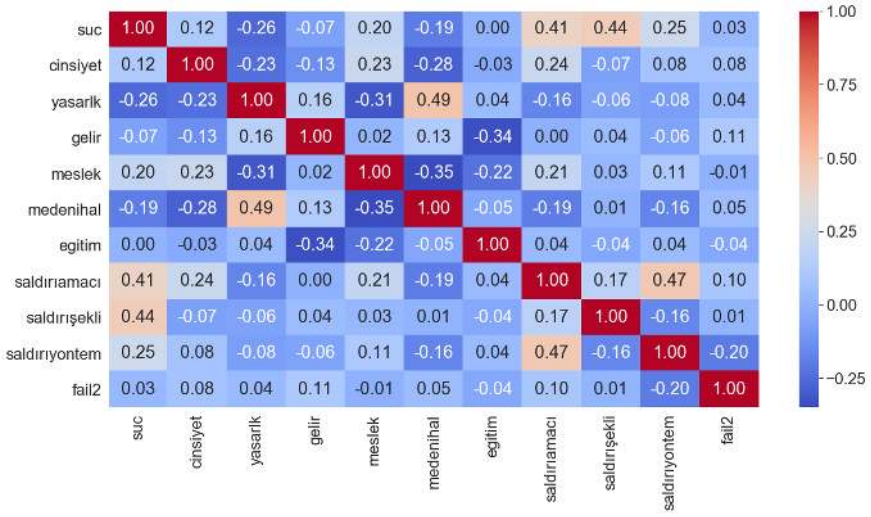
Elazığ ilinde yaşanan siber suçlarda mağdurların uğradıkları maddi zararlar olay istatistiklerinden ayrı ayrı elde edilmiştir. Veri setindeki mağdurların maddi kayıpları yıllara göre ayrılarak Şekil 4.1'deki gibi görselleştirilmiştir. Siber saldırılardan kaynaklanan maddi kayıplar, saldırı yöntemini bilmenin ve siber güvenliği sağlamanın önemini bir kez daha ortaya çıkarmıştır.



Şekil 4.1. Siber saldırı kaynaklı zararların yıllara göre miktarı

Bu bölümde Destek Vektör Makinesi (Doğrusal), Rastgele Orman, Lojistik Regresyon, eXtreme Gradyan Arttırma, Destek Vektör Makinesi (Kernel), Karar Ağacı, K-En Yakın Komşu, Naive Bayes, Yapay Sinir Ağları ve Derin öğrenme algoritmaları kullanılarak elde edilen sonuçlar verilmiştir. Bu veriler arasındaki Pearson korelasyon katsayısını Şekil 4.2'de gösterildiği gibi değerlendirilmiştir. Bu korelasyon matrisi, pratik olarak tüm değişken çiftleri arasında önemli korelasyonlar olduğunu göstermektedir. Modellerde uygulanan algoritmaları değerlendirme kriterleri açısından, tahmin edilen değerlerin test verileri ile karşılaştırılmasıyla doğruluk, kesinlik, duyarlılık ve F1 skor değerleri elde edilmiştir. Bu değerlerin yorumlamaları ayrı ayrı bölümlerde verilerek modelin güvenirliliği de test edilmiştir. Öncelikle verilerin demografik özellikleri grafiklerle verilmiş, daha sonra yapay zekâ yöntemleri ile elde edilen tahmin sonuçları kendi içerinde karşılaştırılmıştır. Literatürdeki çalışmalarla benzerlik ve farklılıkları ortaya koyulmuştur.

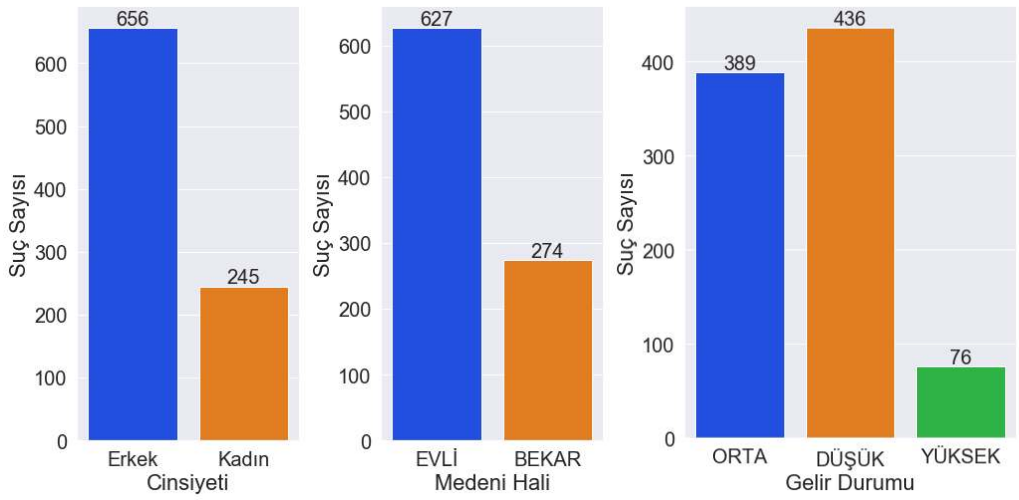
Veri setindeki özniteliklerden yola çıkarak, siber saldırıların hangi amaçla gerçekleştiği ve ne gibi zararı olduğunu gösteren "Siber saldırı amacı", saldırı faillerinin bilinip bilinmediğini gösteren "Siber saldırı faili" ve saldırıların hangi yöntem ve araçlarla gerçekleştirildiğini gösteren "Siber saldırı yöntemi" tahmin edilmiştir.



Şekil 4.2. Pearson korelasyonu

4.1. Verilerin Demografik Analizi

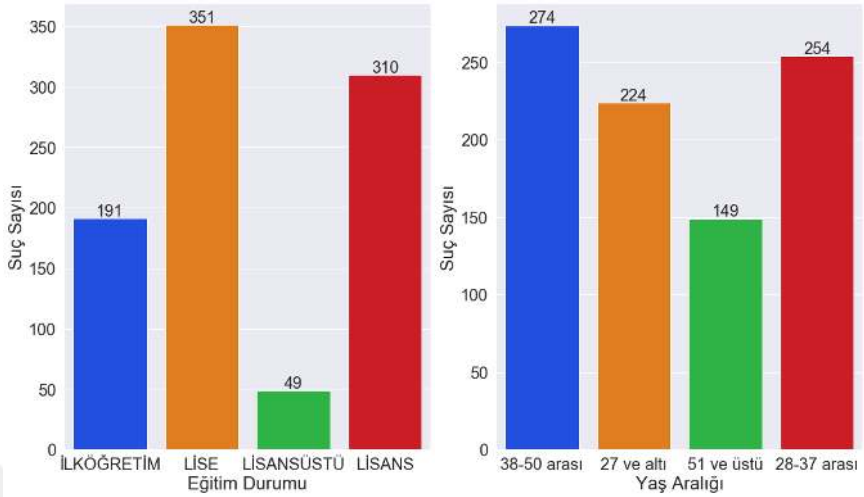
Veri setine göre Şekil 4.3'e bakıldığında siber suça maruz kalanların %72.81'i erkek, %27.19'u ise kadın, %69.59'u evli, %30.41'i bekarlardan oluşmaktadır. Gelir durumlarına bakıldığında %43.17'si orta gelirli, %48.39'u düşük gelirli ve %8.44'dünü ise yüksek gelirliler oluşturmaktadır. Erkek ve evli olan mağdurlar, kadın ve bekar olan mağdurlara göre yaklaşık iki kat daha fazla siber saldırıya maruz kalmıştır. Gelir durumu açısından yüksek gelirlilerin daha az saldırı mağduru olduğu söylenebilir.



Şekil 4.3. Mağdurların cinsiyet, medeni hal ve gelir durumu sayıları

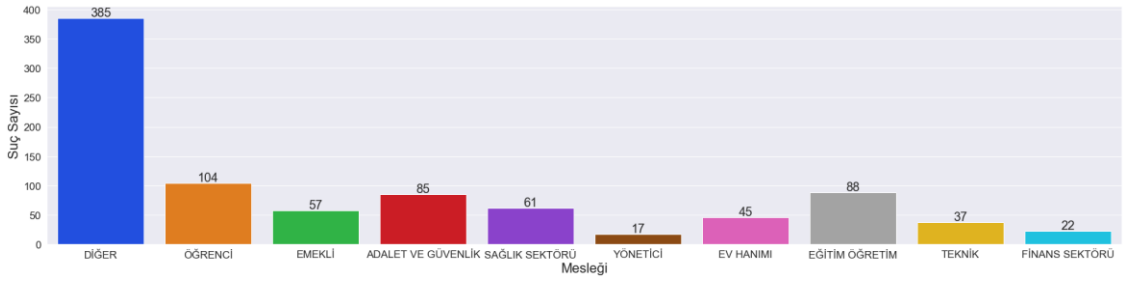
Suçtan zarar görenlerin %21.20'si ilköğretim, %38.96'i lise, %34.41'i lisans ve %5.43'ü ise lisansüstü mezunudur. Yaş dağılımlarına bakıldığında %24.86'sı 27 ve altında, %28.19'u 28 ile 37

yaş arlığında, %30.41'i 38 ile 50 yaş aralığında ve %16.54'ü ise 51 yaş ve üzerindedir. Tüm bu veriler Şekil 4.4'te gösterilmiştir.



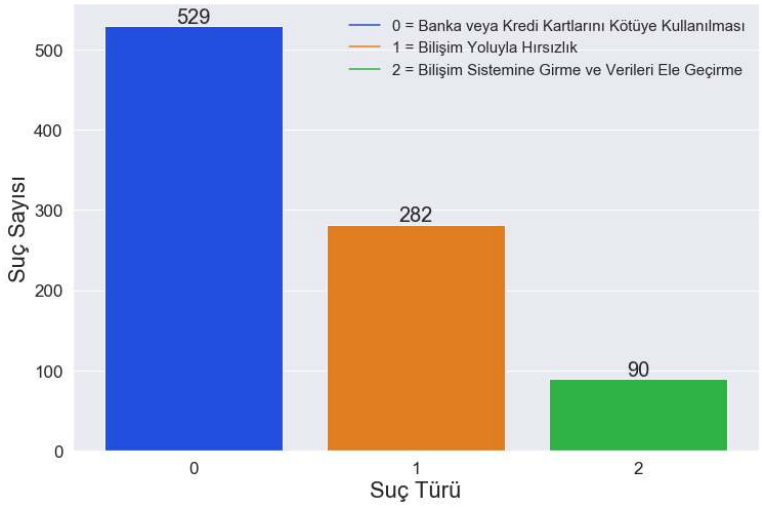
Şekil 4.4. Mağdurların eğitim ve yaş durumu sayıları

Siber saldırıya uğramış olan kişiler meslekleri açısından değerlendirildiğinde, en çok diğer en az ise yönetici meslek grubunun maruz kaldığı Şekil 4.5'ten ayrıntılı olarak görülmektedir.



Şekil 4.5. Mağdurların meslek gruplarına göre sayısı

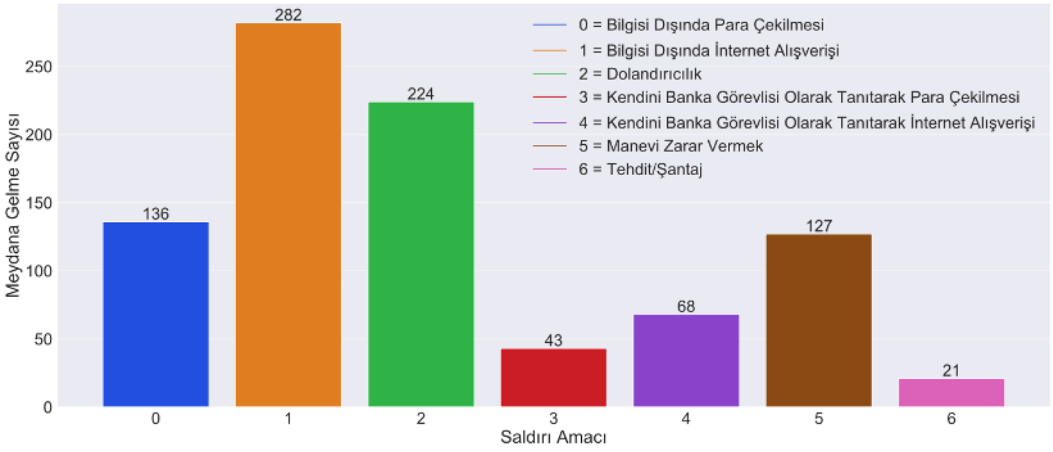
Şekil 4.6'daki suçlar incelendiğinde "Banka veya Kredi Kartlarının Kötüye Kullanılması" suçu en fazla meydana gelen suçlardandır. Kredi kartının başkaları tarafından ele geçirilerek alışveriş yapılması, bankamatiklere kurulan sistemlerce kopyalanması ve şifrelerin bir şekilde ele geçirilmesi bu kapsamda değerlendirilmektedir. Bir başkasına ait bilgisayara hangi yöntemle olursa olsun izinsiz girme, verileri değiştirme ve yok etme suçları "Bilişim Sistemine Girme ve Verileri Ele Geçirme" suçu kapsamında değerlendirilmektedir. Bu suça maruz kalan kişinin sistemindeki etkileri geri dönülemez olduğundan dolayı ciddi zaman ve maddi kayıplar oluşmaktadır. "Bilişim Yoluyla Hırsızlık" ise bilişim sistemleri üzerinden maddi menfaat sağlanan tüm suçları kapsamaktadır. Bu yöntemde mağdurlara genellikle maddi zararlar verilmektedir.



Şekil 4.6. Veri setindeki suç türleri ve sayısı

4.2. Saldırı Amacının Tahmin Edilmesi

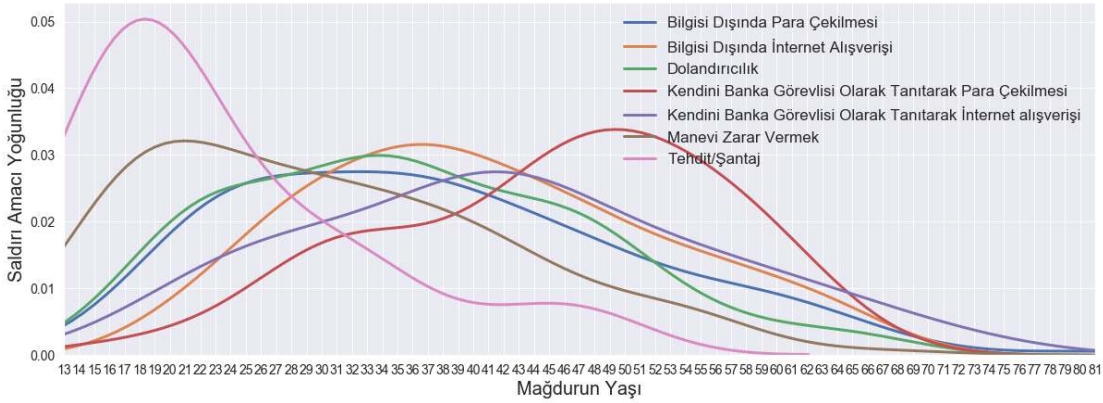
Veri setindeki saldırıların amacını tahmin etmeden önce saldırı amacı ile diğer veriler arasındaki ilişkiler ve gruplar arasında anlamlı farklar olup olmadığı çeşitli grafiklerle verilmiştir. Daha sonra yapılan doğruluk oranları ve algoritma karşılaştırmaları detaylı olarak anlatılmıştır. Şekil 4.7'de verilen saldırı amacının meydana gelme sayısı, "Bilgisi Dışında Para Çekilmesi", "Bilgisi Dışında İnternet Alışverişi", "Dolandırıcılık", "Kendini Banka Görevlisi Olarak Tanıtarak Para Çekilmesi", "Kendini Banka Görevlisi Olarak Tanıtarak İnternet alışverişi", "Manevi Zarar Vermek" ve "Tehdit/Şantaj" olarak karşımıza çıkmaktadır.



Şekil 4.7. Saldırı amacı gerçekleşme sayısı

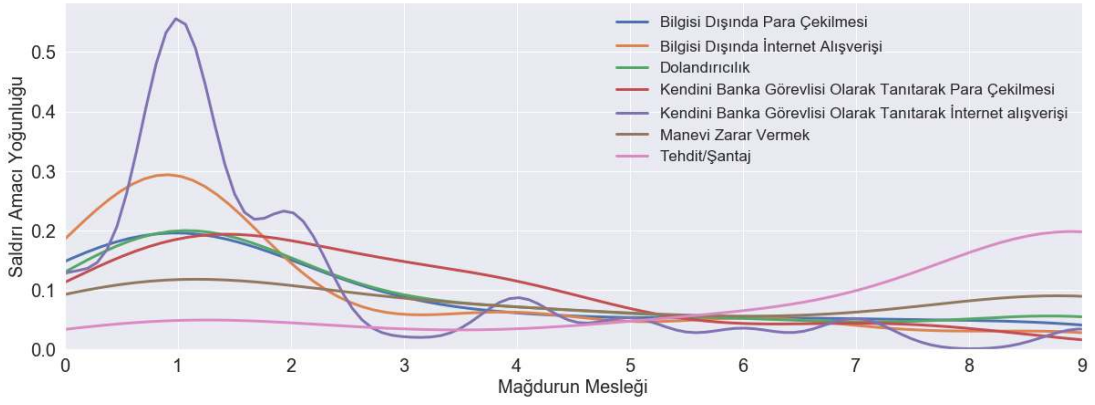
Şekil 4.8'de yaş ve saldırı amacı arasındaki ilişki yoğunluğu gösterilmiştir. Saldırı amacı, yaş gruplarıyla karşılaştırıldığında özellikle 20'li yaş civarındaki kişilerin en çok tehdit ve şantaj

amacıyla, ikinci olarak manevi zarar vermek amacıyla siber saldırıya maruz kaldığı görülmektedir. Dolandırıcılık ve bilgisi dışında internet alışverişi yapılan mağdurların çoğunun yaşları ise 32-40 yaş aralığındadır. Kendisini banka görevlisi olarak tanıtarak parası çekilen mağdurların büyük çoğunluğunun ise 50'li yaşlarda olduğu tespit edilmiştir.



Şekil 4.8. Saldırı amacının yaşa göre yoğunluğu

Veri setinde bulunan "Adalet ve Güvenlik (0)", "Diğer (1), "Emekli (2)", "Ev Hanımı (3)", "Eğitim Öğretim (4)", "Finans Sektörü (5)", "Sağlık Sektörü (6)", "Teknik (7)", "Yönetici (8)" ve "Öğrenci (9)" olanların maruz kaldığı saldırıların amacı Şekil 4.9'da verilmiştir. Özellikle emekli mağdurların manevi zarar vermek ve tehdit/şantaj amacıyla uğradığı saldırının az olduğu ve bu saldırıların genellikle öğrencilere yönelik olduğu görülmüştür. Diğer görevleri yapan meslek gruplarının veri setinde çok olması sebebiyle o grupta daha fazla saldırıya maruz kalmış gibi görülmektedir.



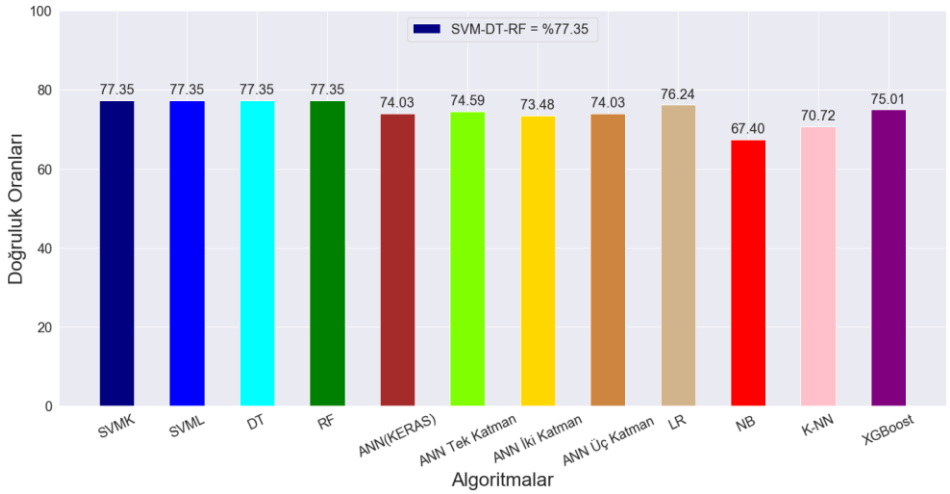
Şekil 4.9. Saldırı amacının mesleğe göre yoğunluğu

Saldırı amacı tahmin modelinde algoritmalar arasındaki doğruluk, kesinlik, geri çağırma ve F1-skor verileri Tablo 4.1'de ve tahminlerin doğruluk karşılaştırması Şekil 4.10'da

gösterilmektedir. Sonuçlar karşılaştırıldığında SVML, SVMK, DT ve RF algoritmaları %77.35 doğruluk oranıyla en yüksek performansı sergilemiştir ve NB ile KNN haricinde diğer sonuçlar birbirine yakındır. LR, XGB, ANN (Tek katman), ANN (keras), ANN (üç katman), ANN (iki katman), KNN ve NB algoritmalarından sırasıyla %1.11, %2.76, %2.76, %3.32, %3.32, %3.87, %7.18 ve %9.95 oranlarında daha iyi performans göstermiştir. En kötü performans, NB algoritmasıyla yapılan uygulamada elde edilmiştir. Algoritmaların kendi arasındaki ne kadar iyi performans gösterdiğini anlamak için F1-skor değerine bakıldığında SVML, SVMK, DT ve RF algoritmaları en iyileridir. Kesinlik açısından en iyi performansta, %78.66 oranıyla NB en iyisidir. Duyarlık açısından en iyi performansı yine SVML, SVMK, DT ve RF algoritmaları göstermiştir.

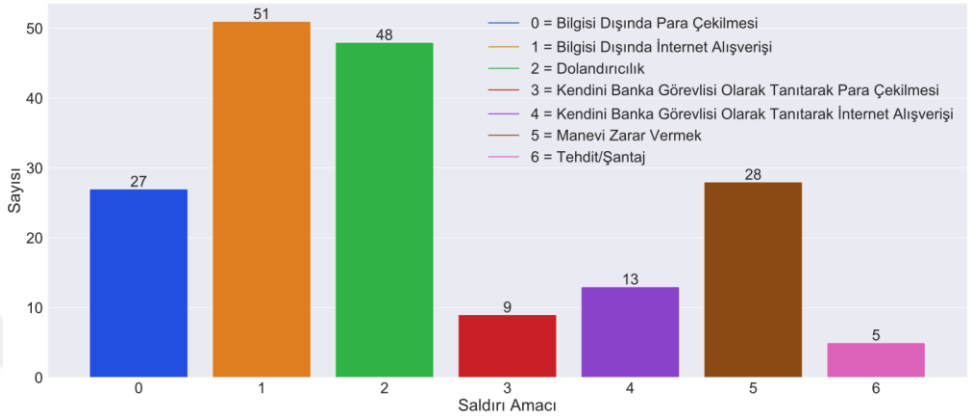
Tablo 4.1. Saldırı amacı tahmin performansı

Algoritmalar	Doğruluk %	Kesinlik %	Duyarlılık %	F1-score %
LR	76.24	74.22	76.24	74.42
KNN	70.17	68.05	70.17	68.45
SVML	77.35	75.02	77.35	75.17
SVMK	77.35	74.62	77.35	74.73
NB	67.40	78.66	67.40	66.41
DT	77.35	75.26	77.35	75.58
RF	77.35	75.35	77.35	75.20
XGB	74.59	73.57	74.59	73.65
ANN (Keras)	74.03	73.23	74.03	73.05
ANN (Tek Katman) Tensorflow	74.59	72.45	74.59	71.50
ANN (İki Katman) Tensorflow	73.48	70.00	73.48	71.36
ANN (Üç Katman) Tensorflow	74.03	71.50	74.03	72.54

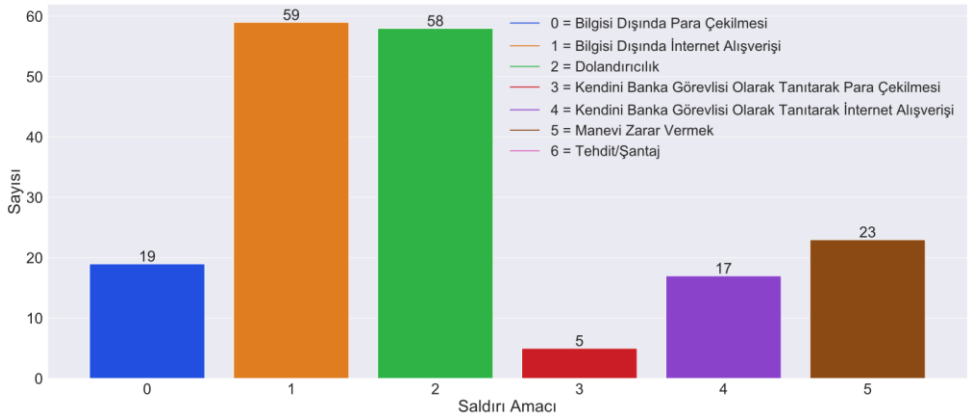


Şekil 4.10. Saldırı amacı tahmini doğruluk oranı karşılaştırması

Tahminlerle karşılaştırmak için %20'si test verisi olarak ayrılan gerçek verilerin dağılımı Şekil 4.11'de ayrıntılı olarak verilmiş ve SVML ile tahmin edilen verilerin grafiği de Şekil 4.12'de gösterilmiştir. İki grafik karşılaştırıldığında sayılardaki fark açıkça görülmektedir ve bu yanlış yapılan tahminlerin detayları hata matrisinde ayrıntılı olarak gösterilmiştir.

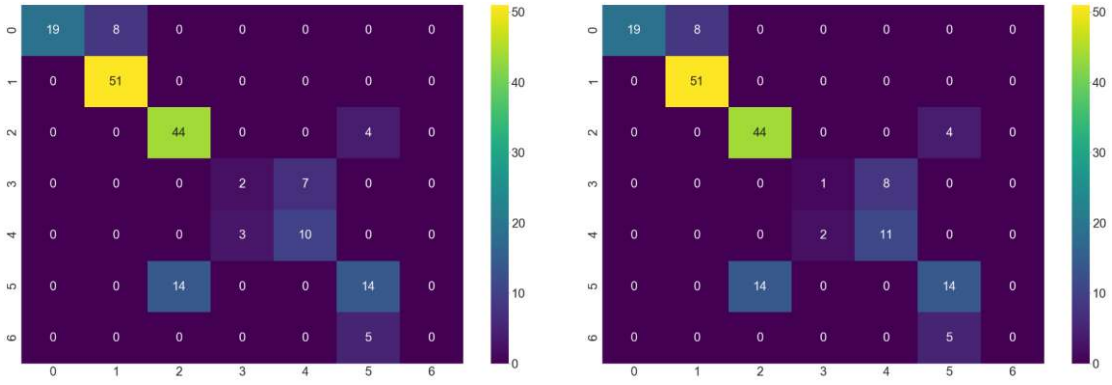


Şekil 4.11. Saldırı amacı test için ayrılan gerçek veriler

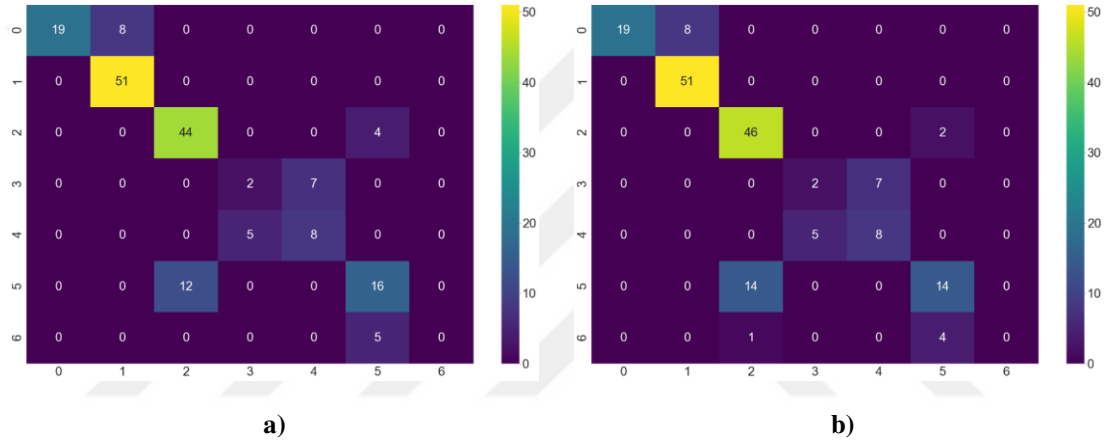


Şekil 4.12. Saldırı amacı SVML ile tahmin edilen veriler

Genel olarak, NB ve KNN hariç tüm algoritmalarda kriterler açısından %70'ten fazla performans elde edilmiştir. En iyi performansı sergileyen algoritmaların performansının değerlendirilmesi için tahminlerin ve gerçek değerlerin karşılaştırıldığı hata matrisleri Şekil 4.13 ve Şekil 4.14'te gösterilmiştir. Hata matrisleri incelendiğinde saldırı amacı tehdit/şantaj (6) olan suçu hiçbir algoritma doğru tahmin edememiştir. Algoritmaların, hata matrisinde "bilgisi dışında internet alışverişi" tahmininde 51 tanesini doğru tahmin ettiği, "dolandırıcılık" tahmininde SVML, SVMK ve DT'nin 44'ünü, RF'nin ise 46'sını doğru tahmin ettiği, "bilgisi dışında para çekilmesi" tahmininde 19 tanesini doğru tahmin ettiği görülmüştür. Diğer tahminler için 4 algoritmada da farklı olduğu görülmektedir. Ayrıca saldırı amacı tahmin karşılaştırma değerleri Ek-1'dedir.



Şekil 4.13. Saldırı amacı hata matrisleri. **a)** SVML sonuçları ve **b)** SVMK sonuçları

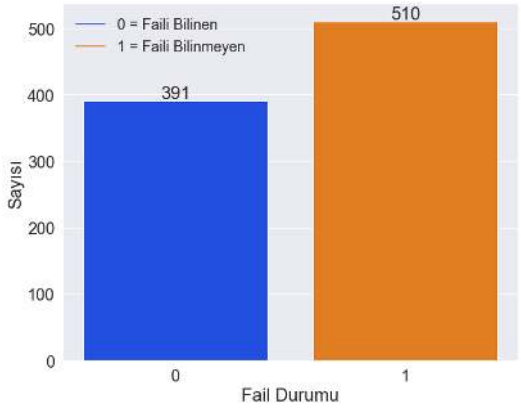


Şekil 4.14. Saldırı amacı hata matrisleri. **a)** DT sonuçları ve **b)** RF sonuçları

4.3. Saldırı Failinin Tahmin Edilmesi

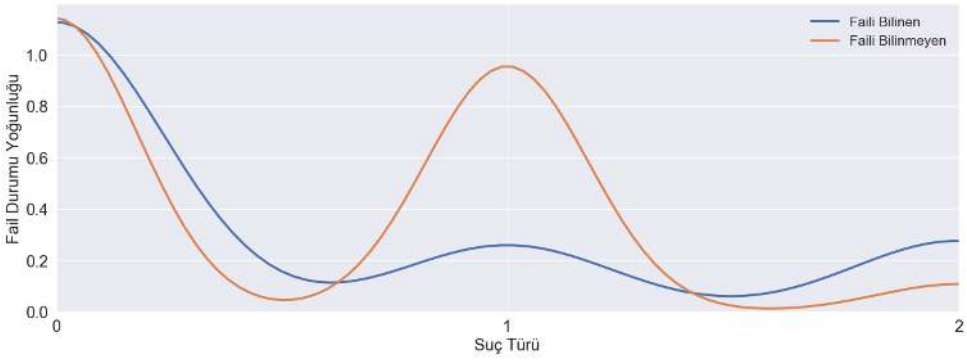
Saldırının failinin bilinip bilinmeyeceğinin tahmin edilmesindeki amaç bu tahminden yola çıkarak, bilinen faillerin öznitelikleri belirlendikten sonra, faili bilinmeyen saldırıları da gerçekleştirip gerçekleştirmediklerinin tespit edilmesidir. Fail ile mağdur öznitelikleri arasında bir ilişki olup olmadığını ortaya koymaktır. Saldırılardaki faillerin bilinme durumlarına bakıldığında, faili bilinmeyenlerin bilinenlerden yaklaşık %30 civarında fazla olduğu Şekil 4.15'te görülmektedir. Bu oran siber olaylarda birçok failin tespit edilemediğini göstermektedir.

Veri setindeki banka kredi kartlarının kötüye kullanılması, bilişim sistemine girme ve verileri ele geçirme ile bilişim yoluyla hırsızlık suçları açısından faillerinin bilinip bilinmediği ile ilgili yoğunluk grafiği Şekil 4.16'de ve suçlara göre bilinen-bilinmeyen faillerin sayıları Şekil 4.17'de verilmiştir. (Grafiklerde 0 = Banka veya Kredi Kartlarının Kötüye Kullanılması, 1 = Bilişim Sistemine Girme ve Verileri Ele Geçirme ve 2 = Bilişim Yoluyla Hırsızlık suçlarını ifade etmektedir) Grafikte banka kredi kartlarının kötüye kullanılması suçunu gerçekleştiren saldırganların bilinme ve bilinmeme durumları birbirine çok yakın durumdadır.

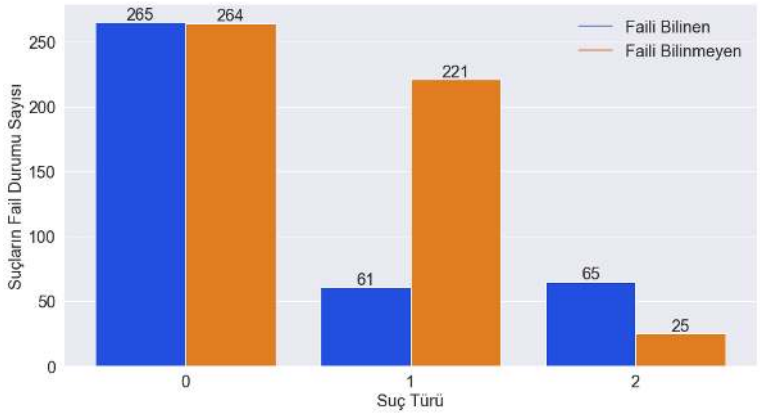


Şekil 4.15. Saldırganların bilinme ve bilinmeme sayısı

Bilişim sistemine girme ve verileri ele geçirme suçunda faili bilinmeyen olaylar bilinenlerden yaklaşık %30 daha fazla yoğunluğa sahiptir. Bilişim yoluyla hırsızlık suçunda ise failinin bilinme oranı bilinmeyenlere göre yaklaşık %260'tır. Bu veriler göstermektedir ki bilişim sistemine girme suçunda saldırganların kimliklerini gizlemede daha başarılı olduğunu, dolayısıyla birçok olayın failinin bilinmediğini göstermiştir.

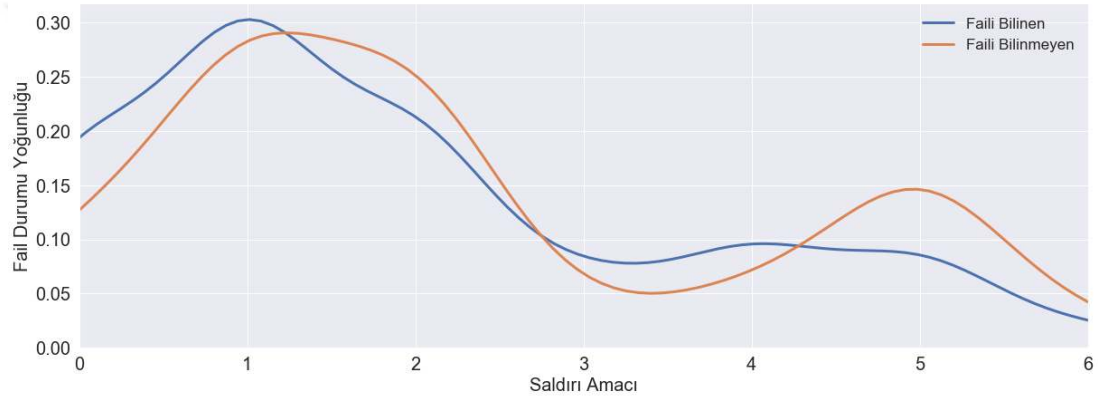


Şekil 4.16. Fail durumunun suç türüne göre yoğunluğu

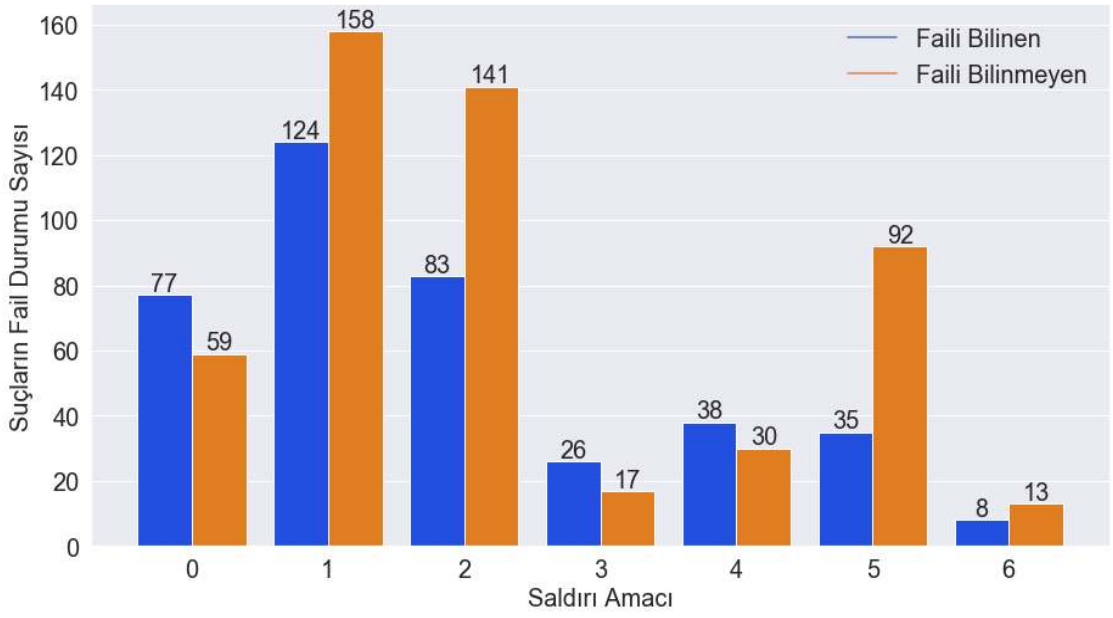


Şekil 4.17. Faillerin suç türüne göre bilinme ve bilinmeme sayıları

Meydana gelen saldırılarda bu saldırıları faillerin neden gerçekleştirdiğini ve bu faillerin bilinip bilinmediğini incelediğimizde, (burada “0 = Bilgisi Dışında Para Çekilmesi" ,"1 = Bilgisi Dışında İnternet Alışverişi", "2 = Dolandırıcılık", "3 = Kendini Banka Görevlisi Olarak Tanıtarak Para Çekilmesi", "4 = Kendini Banka Görevlisi Olarak Tanıtarak İnternet alışverişi", "5 = Manevi Zarar Vermek", "6 = Tehdit/Şantaj” olarak ifade edilmektedir) Şekil 4.18’da faillerin yoğunluk grafiği ve Şekil 4.19’de saldırıların amacına göre faillerin bilinme ve bilinmeme sayıları detaylı olarak gösterilmiştir. Bilgisi dışında para çekilmesi, kendisini banka görevlisi olarak tanıtılarak para çekilmesi ve internet alışverişi amacıyla yapılan saldırıların faillerinin bilinme oranının daha fazladır. Dolandırıcılık amacıyla işlenen suçlarda faillerin bilinmemesi durumu bilinmesinden yaklaşık %70 ve manevi zarar verme suçlarında ise %260 daha fazladır.

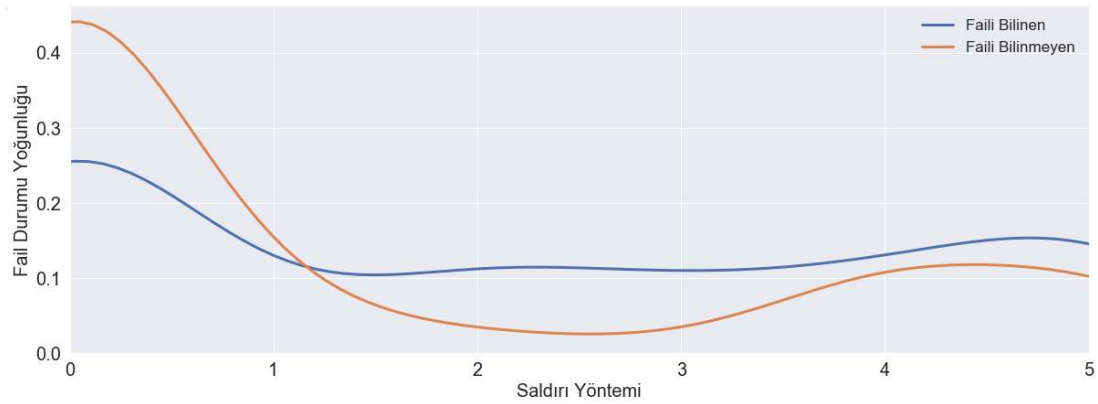


Şekil 4.18. Fail durumunun saldırı amacına göre yoğunluğu

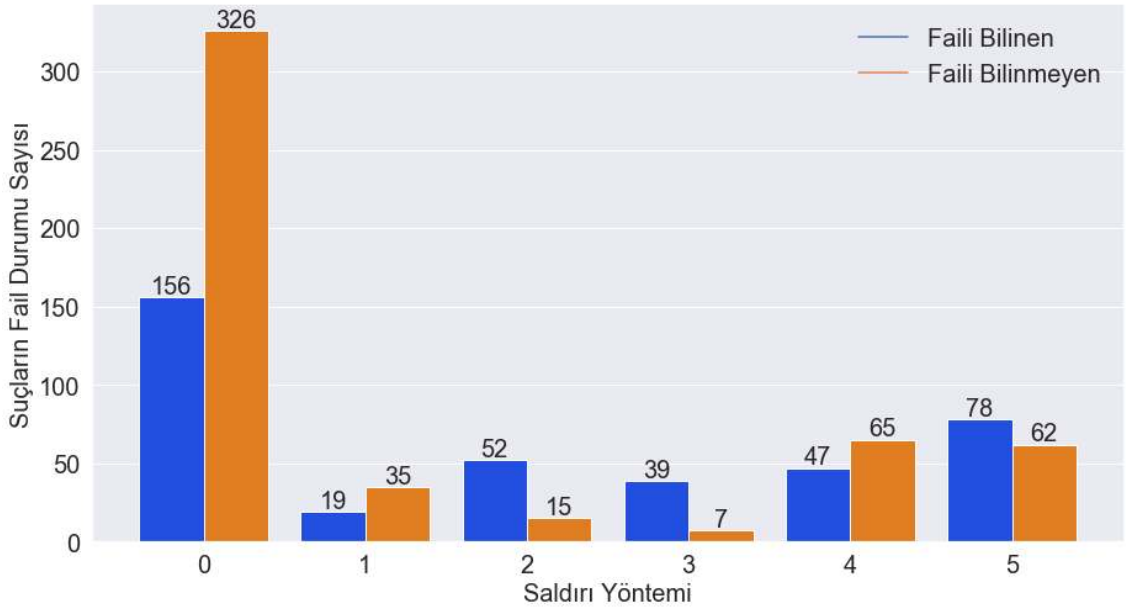


Şekil 4.19. Faillerin saldırı amacına göre bilinme ve bilinmeme sayıları

Faillerin saldırıları hangi yöntemle gerçekleştirdiğine bakıldığında, (burada ”0 = Hack Araçları veya Zararlı Yazılım Kullanarak”, “1 = Kart Kopyalama, Üretme Cihazlarını Kullanarak”, “2 = Phising (oltalama) Saldırısı Kullanarak”, “3= Sahte Alışveriş Sitesi Oluşturarak”, “4= Sosyal Medyadaki Herkese Açık Verilerini Alarak”, “5= Sosyal Mühendislik Kullanarak” ifade edilmektedir) Şekil 4.20 yoğunluk grafiği ve Şekil 4.21’da fail durumu sayıları detaylı olarak gösterilmiştir. Kart kopyalama ve sosyal medyadaki verileri alma suçlarının bilinmeme oranı yüksektir ve hack araçları ve zararlı yazılımlarla yapılan saldırı yönteminde failin bilinmeme oranı %208 civarında daha fazladır. Oltalama ve sosyal mühendislik saldırılarında fail bilinme oranı daha yüksektir. Sahte alışveriş sitesi oluşturarak işlenen suçlarda ise fail bilinme oranının yaklaşık %557 daha fazla olduğu görülmüştür.



Şekil 4.20. Fail durumunun saldırı yöntemine göre yoğunluğu



Şekil 4.21. Faillerin saldırı yöntemine göre bilinme ve bilinmeme sayıları

Saldırı faillerinin bilinip bilinemeyeceğinin tahmin edildiği modeldeki algoritmalar arasında tahmin doğruluğu, kesinlik, duyarlılık ve F1-skor verileri Tablo 4.2'de ve tahminlerin doğruluk karşılaştırması Şekil 4.22'de gösterilmektedir. Sonuçlar karşılaştırıldığında KNN algoritması %69.61 doğruluk oranıyla en yüksek performansı sergilemiştir ve ANN iki katmanlı modelde en kötü sonuç alınmış diğer algoritmaların sonuçları birbirine yakındır. KNN algoritması doğruluk açısından; RF, SVML, SVMK, ANN (keras), LR, NB, DT, ANN (Tek katman), ANN (üç katman), XGB ve ANN (iki katman) algoritmalarından sırasıyla %1.1, %2.21, %2.21, %2.21, %3.31, %3.31, %3.31, %3.31, %4.42, %6.63 ve %8.84 oranlarında daha iyi performans göstermiştir.

Algoritmalarımızın kendi arasındaki ne kadar iyi performans gösterdiğini anlamak için kesinlik, duyarlılık ve F1-skor değerine bakıldığında SVML, SVMK, DT ve RF algoritmaları en iyileridir.

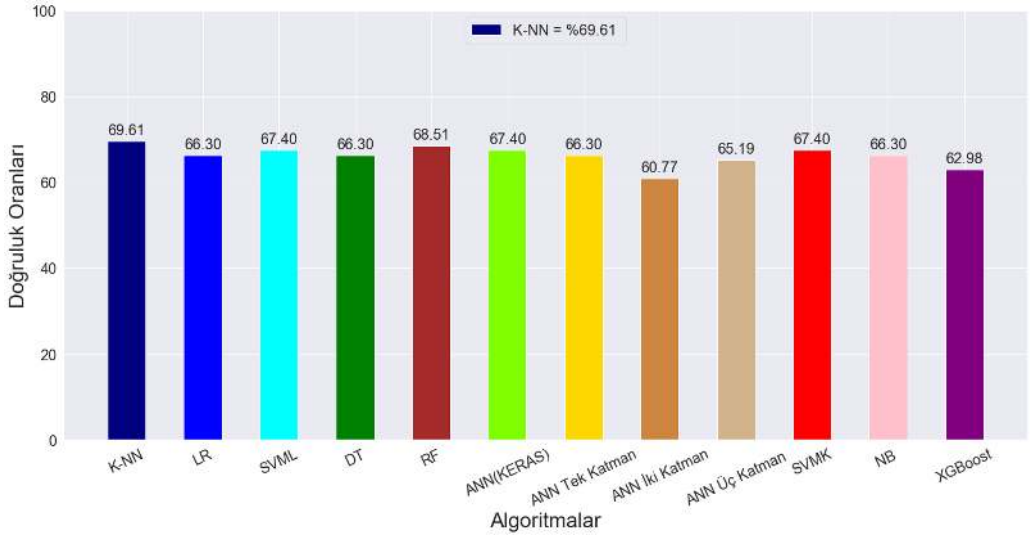
Kesinlik açısından genel olarak algoritmaların performansları %70 altında kalması hatalı tahminlerde olumsuz sonuçlar yaratabilecektir. Özellikle fail bilinme tahmininde yanlış kişilere yapılacak adli işlemin doğuracağı sonuçların geri dönülmesi olmayan zararları olacağından hatalı sonuçların çok olması bu modelin geliştirilmesi gerektiğini göstermiştir. Modelde yanlış kişiyi şüpheli olarak tespit etmenin kurumların güvenirliğine etki edeceğinden hata matrisinin önemini bir kez daha ortaya koymak gerekmektedir.

Duyarlılık açısından da yine performans değerleri beklenen sonuçları vermemiştir.

F1-skoru, kesinlik ve duyarlılık değerlerinin harmonik ortalamasını verdiğinden her iki durumda da sonuçların %70'in altında kalması ve hata oranlarının yüksek olması sebebiyle beklenen performansı gösterememiştir.

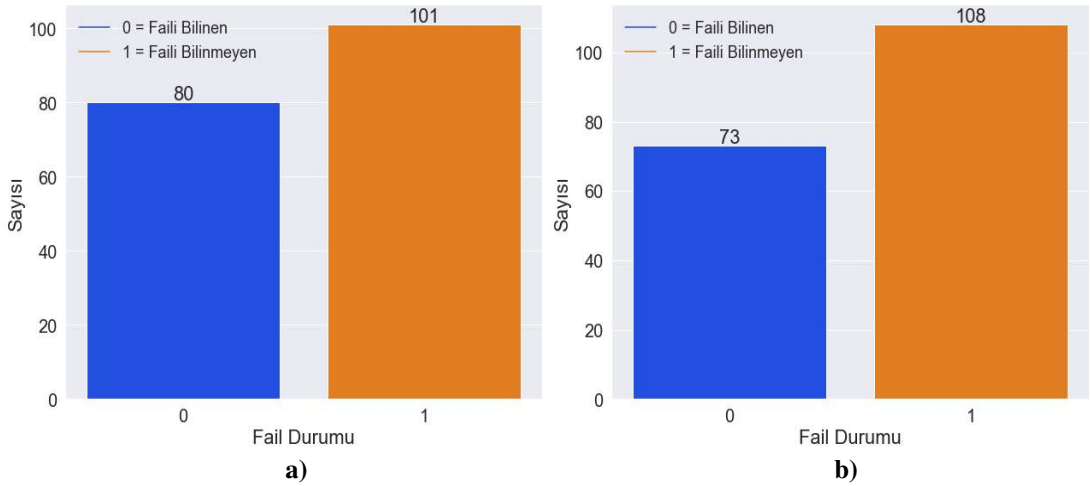
Tablo 4.2. Saldırı faili tahmin performansı

Algoritmalar	Doğruluk %	Kesinlik %	Duyarlılık %	F1-score %
LR	66.30	67.26	66.30	64.29
KNN	69.61	69.45	69.61	69.43
SVML	67.40	68.11	67.40	65.84
SVMK	67.40	68.11	67.40	65.84
NB	66.30	67.85	66.30	63.84
DT	66.30	69.64	66.30	62.77
RF	68.51	68.98	68.51	67.33
XGB	62.98	62.93	62.98	61.21
ANN (Keras)	67.40	67.19	67.40	67.04
ANN (Tek Katman) Tensorflow	66.30	67.26	66.30	64.29
ANN (İki Katman) Tensorflow	60.77	60.28	60.77	60.10
ANN (Üç Katman) Tensorflow	65.19	66.97	62.19	65.19



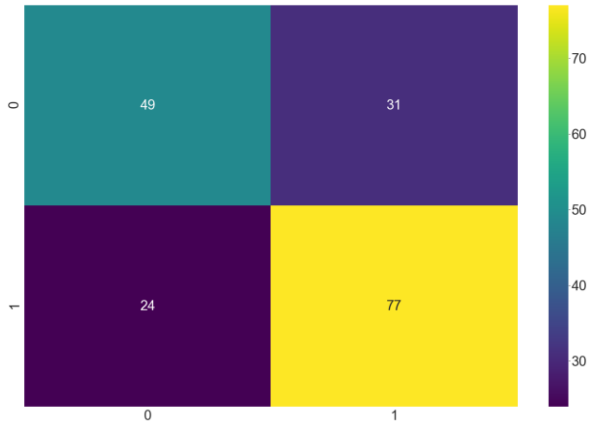
Şekil 4.22. Saldırı faili tahmini doğruluk oranı karşılaştırması

Modelde tahminlerle karşılaştırmak için %20'si test verisi olarak ayrılan gerçek verilerin dağılımı Şekil 4.23-a'da ayrıntılı olarak verilmiş ve tahmin edilen verilerin grafiği de Şekil 4.23-b'de gösterilmiştir. İki grafik karşılaştırıldığında sayılardaki fark ve yanlış yapılan tahminlerin detayları hata matrisinde ayrıntılı olarak gösterilmiştir.



Şekil 4.23. Saldırı faili **a)** Test için ayrılan gerçek veriler **b)** KNN ile tahmin edilen veriler

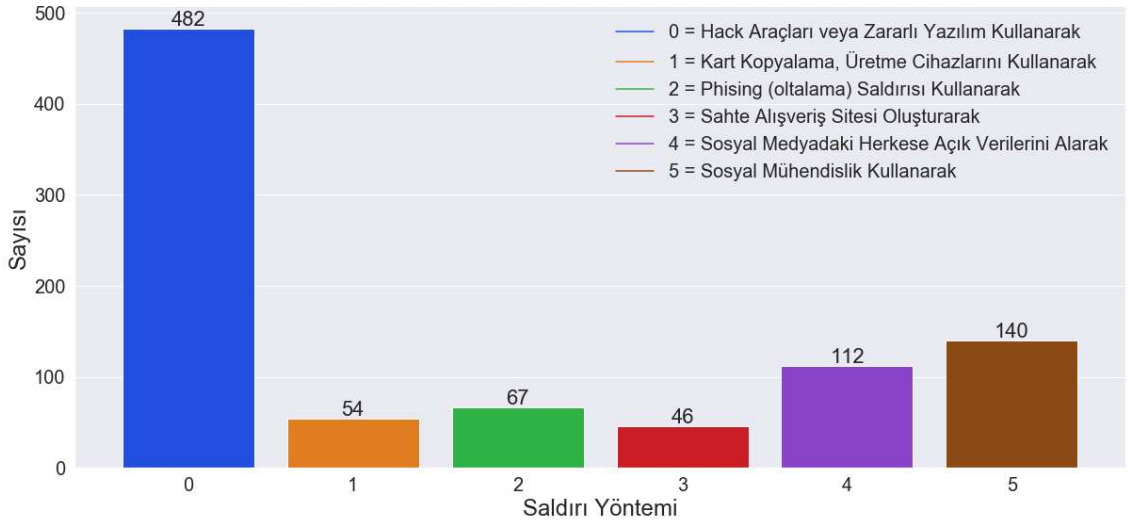
Genel olarak, tüm algoritmalar %60 ile %70 arasında performans göstermiştir. En iyi performansı sergileyen KNN algoritmasının tahminleri ve gerçek değerler karşılaştırıldığı hata matrisi Şekil 4.24'te gösterilmiştir. Hata matrisi incelendiğinde hata oranın yüksek olduğu görülmüştür. Fail tahmin modelinin performansının genel olarak düşük olduğu ve daha farklı modellerle tekrardan değerlendirilebilecektir. Ayrıca fail tahmin karşılaştırma değerleri Ek-2'de verilmiştir.



Şekil 4.24. Saldırı faili KNN algoritması hata matrisi

4.4. Saldırı Yönteminin Tahmin Edilmesi

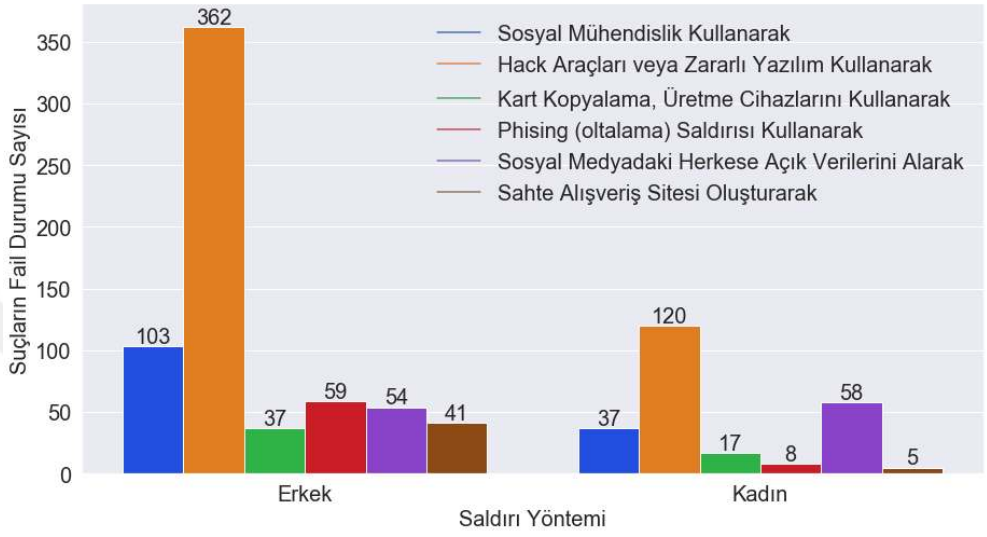
Meydana gelen saldırıların yöntemini tahmin edebilmek, gelecekte olabilecek olayların önlemini önceden almaya fayda sağlayabilir. Saldırıların yöntemini ile yaş, meslek, gelir gibi diğer özniteliklerle bağlantısını tespit etmek ve buna yönelik çözüm önerileri üretmek ciddi faydalar sağlayacaktır. Meydana gelen saldırıların yaklaşık %53.5'i hack araçları veya zararlı yazılım kullanılarak, %6'sı kart kopyalama cihazlarını kullanılarak, %7,5'i oltalama saldırısı kullanılarak, %5'i sahte alışveriş sitesi oluşturularak, %12,5'ü sosyal medyadaki herkese açık verilerin alınarak ve %15.5'i sosyal mühendislik kullanılarak gerçekleştirilmiştir. Şekil 4.25'te saldırı yöntemlerinin ayrıntılı sayıları gösterilmiştir.



Şekil 4.25. Saldırı yöntemi gerçekleşme sayısı

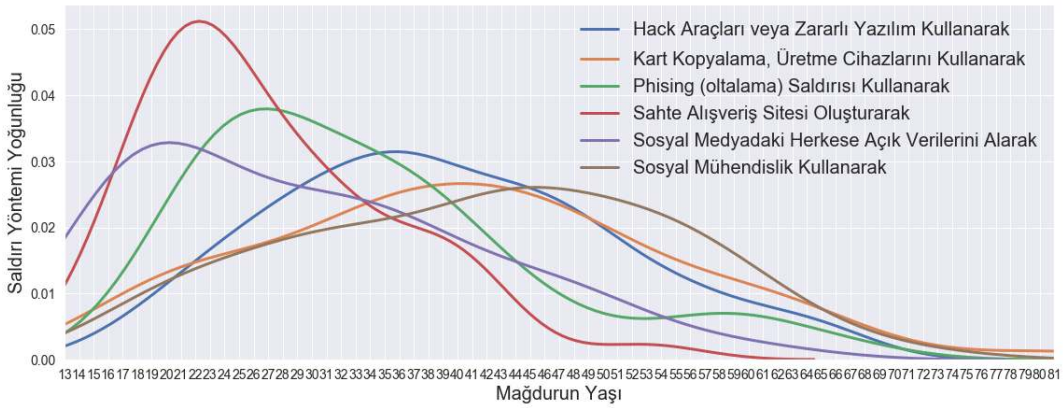
Saldırıları yöntem ve cinsiyet açısından değerlendirdiğimizde Şekil 4.26'deki grafikte kadın ve erkek sayıları verilmiştir. Erkek ve kadınların en çok hack ve zararlı yazlım araçları ile saldırıya

maruz kaldığı görülmektedir. Erkekler en az kart kopyalama ve üretme cihazları kullanılarak saldırıya maruz kalmaktadır. Kadınların sosyal medyadaki herkese açık verileri erkeklere göre daha fazla kullanılmıştır. Yine kadınların kendi içerisinde oltalama ve sahte alışveriş sitesiyle işlenen suçlara daha az maruz kaldığı tespit edilmiştir.



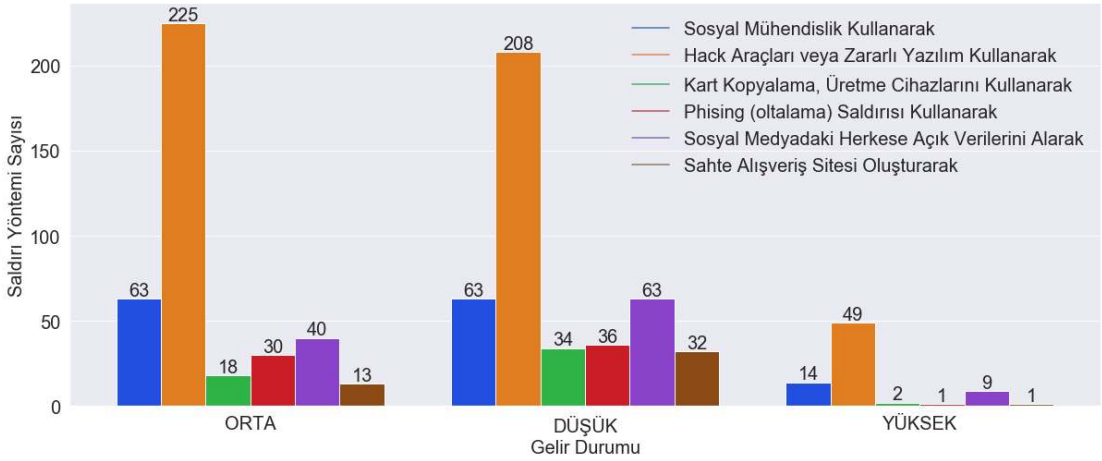
Şekil 4.26. Saldırı yönteminin cinsiyete göre karşılaştırmalı sayısı

Şekil 4.27'de saldırı yöntemi yaşa göre değerlendirildiğinde hack araçları veya zararlı yazılım kullanılarak yapılan saldırının yoğunlukla 30-40 yaş aralığındaki mağdurlara yöneliktir. Kart kopyalama ve üretme cihazlarını kullanılarak yapılan saldırıların 34-45 yaş aralığındakilere, oltalama saldırılarının 22-30 yaş arlığındaki mağdurlara, sahte alışveriş sitesi oluşturularak yapılan saldırıların 19-25 yaş aralığındaki mağdurlara, sosyal medyadaki herkese açık verilerin alınması saldırılarının 16-24 yaş aralığındaki mağdurlara ve sosyal mühendislik saldırıların 40-51 yaş aralığındaki mağdurlara yönelik daha yoğun meydana geldiği görülmektedir.



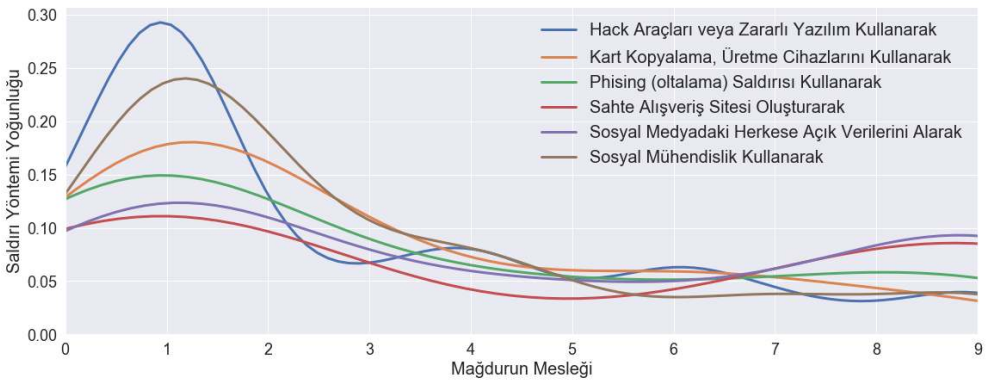
Şekil 4.27. Saldırı yönteminin yaşa göre yoğunluğu

Şekil 4.28'de saldırı yöntemi ile gelir arasındaki ilişki düşük, orta ve yüksek gelir açısından görselleştirilmiştir. Orta ve düşük gelirli mağdurlara yönelik saldırı yöntemlerinin birbirine yakın olduğu fakat kart kopyalama saldırısı ve sahte alışveriş sitesi yöntemlerinde %50 civarında fark olduğu görülmektedir. Yüksek gelirli mağdurlara yönelik saldırılar ile düşük ve orta gelir arasındaki farkın çok olduğu görülmektedir.



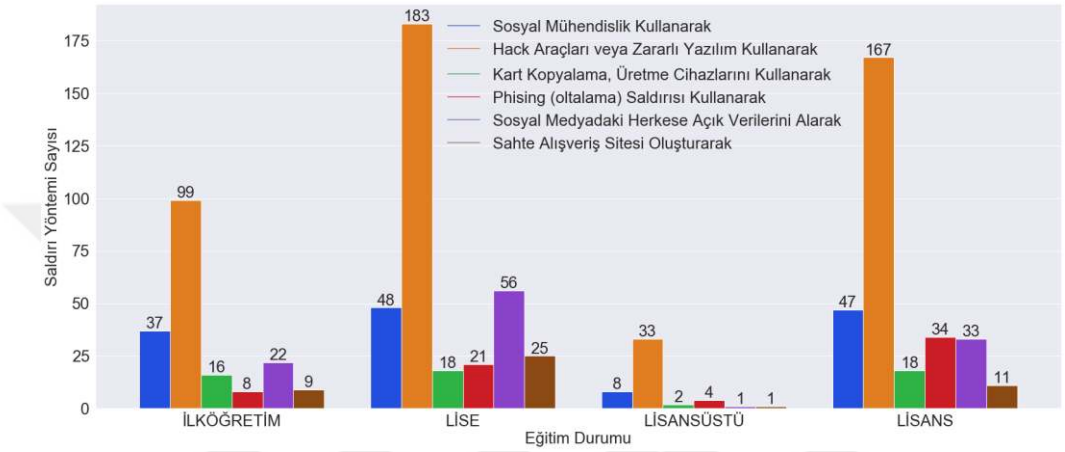
Şekil 4.28. Saldırı yönteminin gelire göre karşılaştırmalı sayısı

Şekil 4.29'da saldırı yöntemi ve meslek arasındaki yoğunluk grafiği verilmiştir. Şekilde "Adalet ve Güvenlik (0)", "Diğer (1), "Emekli (2)", "Ev Hanımı (3)", "Eğitim Öğretim (4)", "Finans Sektörü (5)", "Sağlık Sektörü (6)", "Teknik (7)", "Yönetici (8)" ve "Öğrenci (9)" şeklinde ifade edilmiştir. Grafikte öğrencilere karşı işlenen suçlarda sosyal medyadaki herkese açık verilerinin alınması ve sahte alışveriş sitesi oluşturarak dolandırılma suçları ön plana çıkmaktadır. Emeklilerin ise en yoğun şekilde sosyal mühendislik saldırısına ikinci sırada ise kartlarının kopyalanması saldırılarına maruz kaldığı görülmektedir. Eğitim öğretim ve finans sektöründe çalışanlar sahte alışveriş sitesi oluşturularak dolandırılma olaylarına daha az maruz kalmaktadır.



Şekil 4.29. Saldırı yönteminin mesleğe göre yoğunluğu

Şekil 4.30 da saldırı yöntemi ile eğitim durumu arasındaki ilişki ilköğretim, lise, lisans ve lisansüstü mezunları açısından değerlendirilerek gösterilmiştir. Lise mezunu mağdurların yoğun bir şekilde sahte alışveriş sitesi yöntemiyle dolandırıcılık ve sosyal medyadaki herkese açık verilerinin alınarak kullanılması suçlarına maruz kaldıkları tespit edilmiştir. Oltalama saldırısına maruz kalanların yarısı ise lise mezunudur. Genel olarak lisansüstü mezunlarının en az sayıda saldırı yöntemine maruz kaldığı görülmektedir.



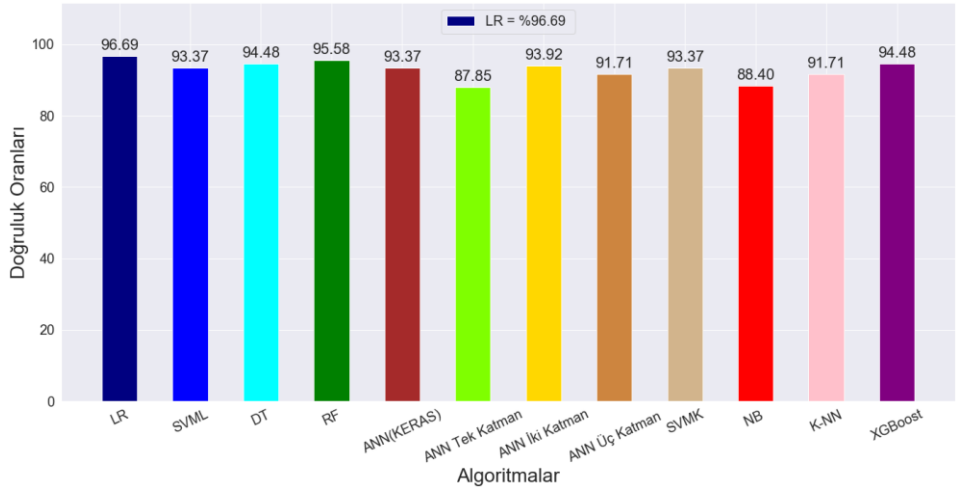
Şekil 4.30. Saldırı yönteminin eğitime göre karşılaştırmalı sayısı

Günümüzde siber saldırılarla mücadele etmek, engellemek ve zararı en aza indirgemek kurum, kuruluş ve şahısların önem verdiği konuların başında geldiğinden dolayı saldırı çeşitlerini bilmek bu konuda fayda sağlayacaktır. Geçmişteki siber olayların nasıl ve ne şekilde gerçekleştiğini bilmek, önlem almak için gelecekte bize ışık tutacaktır. Özellikle önerilen modelde en yüksek performanslar saldırı yöntem tahmininde elde edilmiştir. Yöntem tahmininin doğruluk, kesinlik, duyarlılık ve F1-skor sonuçları Tablo 4.3'te ve doğruluk performans karşılaştırmaları Şekil 4.31'de verilmiştir.

Tahmin doğruluğu açısından en iyi performans LR algoritmasında elde edilmiş ve en kötü performansı ANN (Tek Katman) algoritması sergilemiştir. LR algoritması ile elde edilen tahmin, RF, DT, XGB, ANN (İki Katman), SVML, SVMK, ANN (Keras), KNN, ANN (Üç Katman), NB ve ANN (Tek Katman) algoritmalarından sırasıyla %1.11, %2.21, %2.21, %2.77, %3.32, %3.32, %3.32, %4.98, %4.98, %8.29 ve %8.84 oranında daha yüksek sonuç vermiştir. Kesinlik, duyarlılık ve f1 skor açısından değerlendirildiğinde tüm sonuçlar %86 ile %97 arasında sonuç sergilemiştir. En kötü performansı kesinlik, duyarlılık ve F1-skor açısından ANN (Tek Katman) sergilemiştir. Tahmin doğruluğunda en iyi performansı veren LR algoritması kesinlik duyarlılık ve F1-skor açısından da en yüksek sonuçları vermiş ve hata oranı %4'ün altındadır. Bu sonuçlar önerilen modelin başarısını bir kez daha göstermiştir.

Tablo 4.3. Saldırı yöntemi tahmin performansı

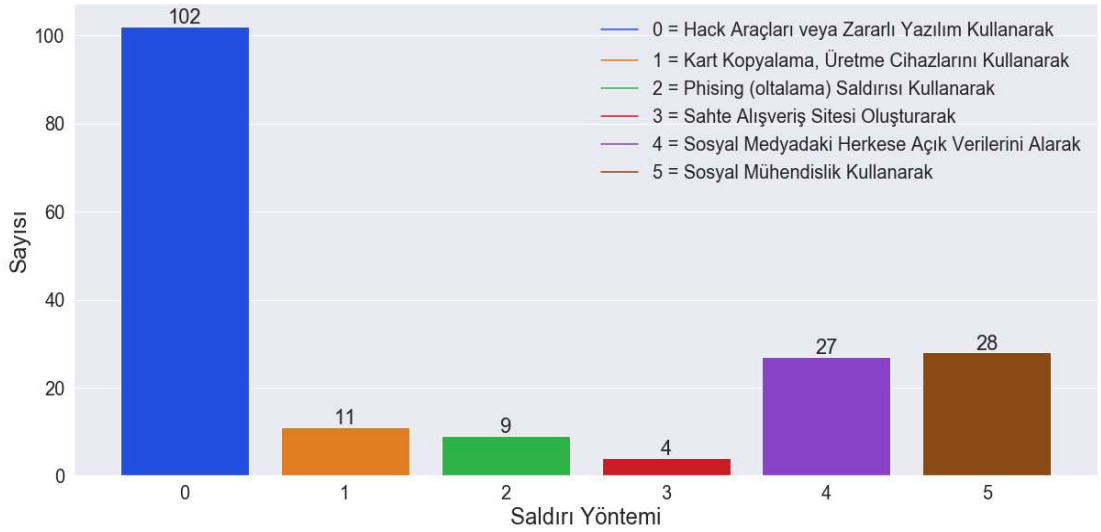
Algoritmalar	Doğruluk %	Kesinlik %	Duyarlılık %	F1-score %
LR	96.69	96.96	96.69	96.77
KNN	91.71	92.54	91.71	91.99
SVML	93.37	95.18	93.37	93.95
SVMK	93.37	95.18	93.37	93.95
NB	88.40	89.56	88.40	88.11
DT	94.48	95.64	94.48	94.80
RF	95.58	96.07	95.58	95.73
XGB	94.48	95.23	94.48	94.71
ANN (Keras)	93.37	93.68	93.37	93.49
ANN (Tek Katman) Tensorflow	87.85	89.14	87.85	86.07
ANN (İki Katman) Tensorflow	93.92	94.29	93.92	94.03
ANN (Üç Katman) Tensorflow	91.71	92.12	91.71	91.80



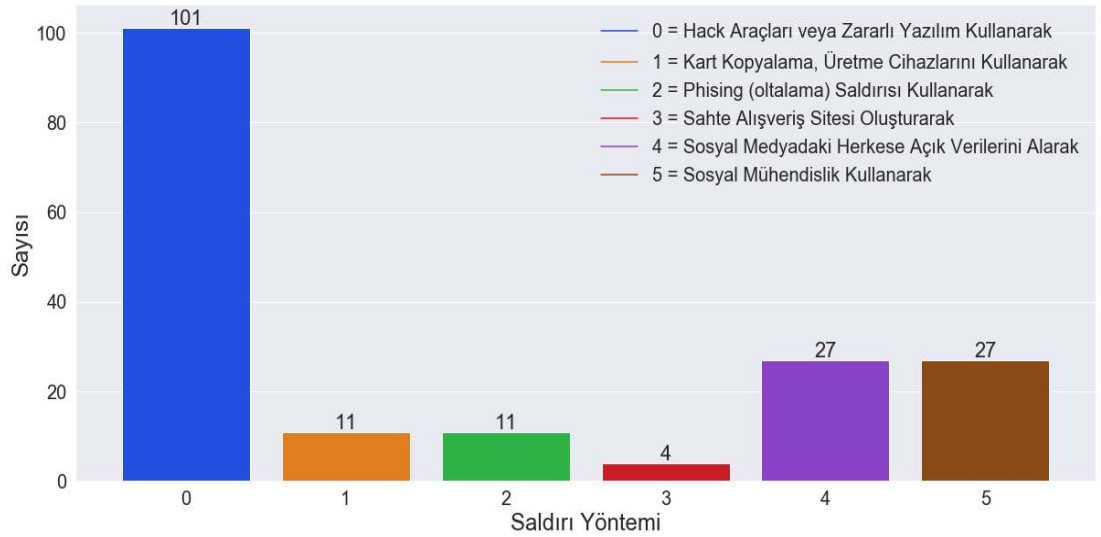
Şekil 4.31. Saldırı yöntemi tahmini doğruluk oranı karşılaştırması

Python yazılımı, sklearn kütüphanesinde bulunan, çapraz doğrulama ve veri kümesini bölmek için kullanılan model_selection modülü ile rastgele yöntem ile veriler ayrılmıştır. Ayrılan bu veri tahminleri karşılaştırmak ve sonuçları değerlendirmek için kullanılmıştır. Şekil 4.32'de saldırı yöntemi modelinde test için ayrılan gerçek verilerin dağılımı ve Şekil 4.33'te en iyi performansı sergileyen LR algoritması ile elde edilen tahmin verilerin dağılımı gösterilmiştir.

İki grafik karşılaştırıldığında sayılardaki fark ve yanlış yapılan tahminlerin detayları hata matrisinde ayrıntılı olarak gösterilmiştir. Ayrıca saldırı yöntemi en yüksek performansı sergileyen LR algoritması tahmin karşılaştırma değerleri Ek-3'te verilmiştir.



Şekil 4.32. Saldırı yöntemi test için ayrılan gerçek veriler



Şekil 4.33. Saldırı yöntemi LR ile tahmin edilen veriler

Yukarıda dağılımları verilen gerçek test verisi ile en iyi performansı sergileyen LR algoritması tahmin verileri hata matrisi Şekil 4.34'te gösterilmiştir. Hata matrisi incelendiğinde;

"Hack araçları veya zararlı yazılım kullanılması (0)" tahmininde 102 gerçek veriden 99 tanesi doğru pozitif (TP) olarak tahmin edilmiştir. 2 tanesi de yanlış negatif (FN) yani gerçek veride "Hack araçları veya zararlı yazılım kullanarak" iken, "Sosyal mühendislik saldırısı" olarak tahmin edilmiştir. 3 tanesi yanlış pozitif (FP) yani gerçek veride "Oltalama saldırısı kullanarak" iken, "Hack araçları veya zararlı yazılım kullanarak" olarak tahmin edilmiştir.

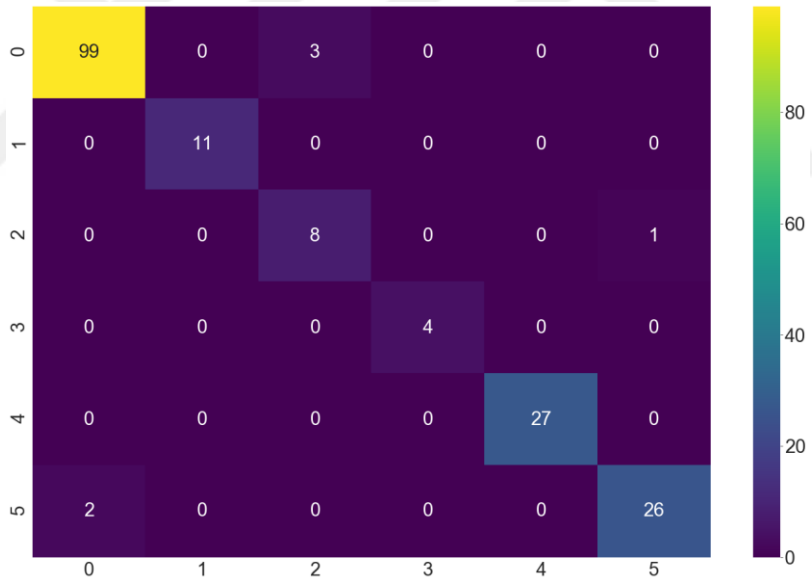
"Kart kopyala ve üretme cihazlarını kullanma (1)" yönteminde 11 gerçek veriden 11 tanesi de doğru pozitif (TP) olarak tahmin edilmiştir.

“Oltalama saldırısı kullanma (2)” yönteminde 11 gerçek veriden 8 tanesi doğru pozitif (TP) olarak tahmin edilmiştir. 3 tanesi de yanlış negatif (FN) yani gerçek veride “Oltalama saldırısı” iken, “Hack araçları veya zararlı yazılım kullanarak” olarak tahmin edilmiştir. 1 tanesi ise yanlış negatif (FN) gerçek veride “Sosyal mühendislik kullanılarak” iken, “Oltalama saldırısı” olarak tahmin edilmiştir.

“Sahte alışveriş sitesi oluşturularak (3)” saldırı yönteminde, 4’üde doğru pozitif (TP) olarak tahmin edilmiştir.

“Sosyal medyadaki herkese açık verilerin alınarak (4)” saldırı yönteminde 27’side doğru pozitif (TP) olarak tahmin edilmiştir.

“Sosyal mühendislik kullanılarak (5)” saldırı yönteminde, 27 gerçek veriden 26 tanesi doğru pozitif (TP) olarak tahmin edilmiştir. 1 tanesi de yanlış negatif (FN) yani gerçek veride “Sosyal mühendislik kullanılarak” iken, “Oltalama saldırısı kullanarak” olarak tahmin edilmiştir. 2 tanesi ise yanlış negatif (FN) gerçek veride “Hack araçları veya zararlı yazılım kullanarak” iken, “Sosyal mühendislik kullanılarak” olarak tahmin edilmiştir.



Şekil 4.34. Saldırı yöntemi LR algoritması hata matrisi

Literatürde suçları ve saldırı yöntemlerini tahmin edebilmek için birçok yöntem kullanılmıştır. Genellikle suç tahmininde genel suç türleri tahmin edilmiş ve siber suç verileriyle daha az çalışılmıştır. Diğer tahminlerde ise suçun türü, zamanı, yeri, bölgesi, saldırgan, demografik ve bölgesel faktörler gibi verilerin daha çok öznitelik olarak kullanıldığı görülmektedir.

Bazı çalışmalar, önerilen model ile karşılaştırıldıktan sonra bazı özniteliklerin benzer olduğu ortaya çıkmıştır. Özellikle siber suç odaklı olanlarda, veri setleri ve parametreler birbirinden farklıdır.

Yapılan çalışmanın sınırları, veri setinin gerçek verilerden oluşması nedeniyle veri setinin miktarıdır. Zamansal veriler olayların gelecekte ne zaman meydana geleceğini zaman serileriyle tahmin edilmektedir. Olayların hangi günde ve hangi saatte meydana geldiği verisinin mevcut olmaması da bir başka sınırlılık olarak görülmektedir. Aynı şekilde, saldırıların teknik detayları polis kayıtlarında mevcut olsaydı, detaylı tahminler ile failin yakalanması kolaylaştırabilirdi.

Pek çok çalışmada, gelecekte suçların nerede ve ne zaman işleneceğini tahmin edilmiştir. Ancak birçoğu suçların yöntemine, suçların nasıl önlenebileceğine ve failin özelliklerinin neler olduğuna değinmemiştir. Bu çalışmanın en önemli avantajlarından biri gerçek verileri kullanarak saldırıya uğrayan kurbanlarla benzer özelliklere sahip kişiler için profil oluşturmaya yönelik bir ön adımdır. Önerilen çalışmanın bir diğer avantajı da siber saldırı yönteminin ne olacağı ve failinin tespit edilip edilemeyeceğinin tahmin edilmesidir. Böylece bu suçlarla ilgili olan alanlardaki kurum ve kuruluşların alacağı siber önlemlerle ilgili ön bilgi sağlanmış olacak ve bu sayede maddi kayıpların önüne geçilebilecektir. Yine kişiye bağlı öznitelikler (yaş, cinsiyet, eğitim, meslek vd.) sisteme girildiğinde hangi saldırı yöntemi ile karşı karşıya kalabileceği tahmin edilmektedir.

5. SONUÇLAR

Siber güvenlik alanında, yapay zekâ teknolojilerinin kullanılmasının kullanıcılara kolaylıklar sağlamasının yansıra insanoğlunun bu alandaki düşünme ve problem çözme yetisine de ciddi katkıları olmuştur. Özellikle makine öğrenmesi ve derin öğrenme çözümleri, insanoğlunun yerine karar alıp uyguladığından bu tehditlerin üstesinden gelmede önde gelen yöntemlerdendir. Yapay zekanın gücünü kullanarak siber tehditleri algılamak ve siber alemde suç ve suçlularla mücadele etmek insanoğlunun problem çözme becerilerini bir adım daha ileriye taşımıştır.

Bu tezde; siber suç verileri kullanılarak makine öğrenimi ve derin öğrenme gibi akıllı öğrenme algoritmaları ile siber saldırıları öngören ve tespit eden bir sistem geliştirilmiştir. Bu modeller, suç saldırısının amacını, suçun failini ve suçun hangi yöntemle işlendiğini tahmin eden üç ana modeldir.

Saldırı amacını tahmin modelinde;

- SVML, SVMK, DT ve RF algoritmalarıyla en iyi sonuçlar elde edilmiştir. Sonuçlar yeterince tatmin edici olmakla birlikte hibrit yöntemlerle geliştirilebilir.
- Tehdit ve şantaj amacıyla işlenen suçlar genellikle 25 yaşın altındaki mağdurlara yöneliktir. Öğrenciler yoğun şekilde manevi zarar amacıyla saldırıya uğramaktadır. 50 yaşının üzerindeki mağdurların hiçbiri tehdit-şantaj amacıyla işlenen suça maruz kalmamıştır. 60 yaşın üzerindeki hiçbir mağdurda manevi zarar amacıyla işlenen suçun mağduru olmamıştır.
- Lisansüstü mezunlarına yönelik saldırılardaki amaç, çoğunlukla bilgisi dışında para veya internet alışverişi içindir.
- Bilgisi dışında para çekilmesi ve internet alışverişi ile dolandırıcılık amacıyla gerçekleşen suçlara erkekler daha fazla maruz kalmaktadır.

Saldırı faili tahmin modelinde;

- KNN algoritmasıyla en iyi performans elde edilmiştir. Fakat tüm algoritmalarla elde edilen sonuçlar yeterince iyi olmamakla birlikte modelin geliştirilmesi gereklidir. Özellikle hatalı tahminlerin yaratacağı sonuçlardan dolayı modelin bu haliyle kullanılmaması kanaatine varılmıştır. Veri setindeki bilgilerden sadece failin bilinip bilinemeyeceği tahmin edildiğinden, diğer bilgilere ulaşılamayacağı kanaatine varılmıştır.
- Bilişim sistemine girme ve verileri ele geçirme suçunda faillerin bilinme oranı bu suçun ortaya çıkma ihtimalinin daha yüksek olduğunu göstermiştir.
- Oltalama saldırısı kullanılarak ve sahte alışveriş sitesi oluşturularak ürün satışı yapılan saldırıların failinin bilinme oranı çok daha yüksektir.

Saldırı yöntemi tahmin modelinde;

- LR algoritması ile en iyi tahmin performansı elde edilmiştir. Bu saldırı yöntemi tahmin modeli 3 model arasındaki en başarılı model olduğundan umut vaat edicidir. Saldırı yöntemini tahmin ettiğimiz sonuçlardaki hata oranı oldukça düşüktür ve doğru pozitif tahminleri oldukça yüksektir.
- Özellikle kötü amaçlı yazılım ve sosyal mühendislik saldırılarına dikkat çekmek gerektiği sonucuna varılmıştır.
- Siber saldırıların maddi ve manevi zararlarının kişi ve kurumları olumsuz etkilediğinden saldırı yöntemlerini önceden tahmin etmenin çok önemli olduğu sonucuna varılmıştır.

Modelde saldırıya uğrayabilecek kişilerin özellikleri ile hangi saldırı yöntemlerine maruz kalabilecekleri ortaya çıkarılmıştır.

Makine öğrenme yöntemlerinin yeterince başarılı olduğu görülmüştür.

Mağdurun eğitim ve gelir düzeyi ne kadar yüksekse, siber saldırı olasılığının o kadar az olduğu bulunmuştur.

Bu sonuçlar, yaklaşımların siber saldırı yöntemini tahmin etmek için başarıyla uygulanabileceğini göstermiştir.

Geliştirilen sistem, saldırıların hangi yöntemle işlendiğini önceden tahmin ederek siber suçla mücadele eden kolluk kuvvetlerine çözüm sunmaktadır. Böylece suç ve suçluları tespit etmede ve alınması gereken önlemler hakkında daha hızlı daha etkili olunarak mağduriyetler en aza indirilecektir.

ÖNERİLER

Bu alanda yapılacak sonraki çalışmalar için; suç, suçlu ve mağdur profilleme ile siber saldırılar, makine ve derin öğrenme algoritmaları kullanılarak tahmin edilebilir ve sonuçlar karşılaştırılabilir.

Suç veri tabanına sahip diğer yetkili birimlerle yapılan görüşmelerden yola çıkarak diğer şehirlere ait siber suç verileri de bu çalışma ve benzer çalışmalarla karşılaştırmak için elde edilebilir. Suç oranlarını düşürmek için suç ve suçlularla mücadelede kolluk kuvvetlerine faydalı olabilecek akıllı suç-mağdur tespit sistemleri oluşturulabilir.

Siber suçun hangi mağdura hangi yöntemle işlendiği ve faille ilgili diğer anonim verilerin elde edilmesi halinde failin bilinip bilinmediği bilgisinden yola çıkarak aynı failin siber saldırıyı gerçekleştirip gerçekleştirmediği akıllı yöntemlerle geliştirilecek yeni bir modelle tespit edilebilecektir.

Analiz çalışmamızda ortaya çıkan saldırı mağdurlarının özelliklerinin değerlendirilmesi ile benzer özelliklere sahip kişiler için yeni eğitim ve uyarı sistemleri oluşturulabilir.

Önerilen sistem ve yöntemlerle, olaylarda toplanan verilerin doğru analiz edilmesiyle birlikte suçun önceden önlenebilmesi ve faillerin yakalanması sağlanabilir.

Elde edilen bu sonuçlar suçla mücadele eden kolluk kuvvetlerinin yürüttüğü soruşturmalarda kullanılarak gizli kalan taraflar ortaya çıkarılabilir.

KAYNAKLAR

[1] Ben-Asher, N., & Gonzalez, C. (2015). Effects of cyber security knowledge on attack detection. Computers in Human Behavior, C. 48, 51-61.

[2] Sayan, C. M. (2017). An Intelligent Security Assistant for Cyber Security Operations. In 2017 IEEE 2nd International Workshops on Foundations and Applications of Self* Systems, 375-376.

[3] Crawford, J. (2017). The Impact of Artificial Intelligence on Autonomous Cyber Defense, Doctoral dissertation, Utica College.

[4] Qiu, J., Wu, Q., Ding, G., Xu, Y., & Feng, S. (2016). A survey of machine learning for big data processing. EURASIP Journal on Advances in Signal Processing, C. 1, 67.

[5] Berman, D. S., Buczak, A. L., Chavis, J. S., & Corbett, C. L. (2019). A survey of deep learning methods for cyber security. Information, C. 10(4), 122.

[6] Chen, Y. C., Li, Y. J., Tseng, A., & Lin, T. (2017). Deep learning for malicious flow detection. In 2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications, 1-7.

[7] Teoh, T. T., Chiew, G., Franco, E. J., Ng, P. C., Benjamin, M. P., & Goh, Y. J. (2018). Anomaly detection in cyber security attacks on networks using MLP deep learning, 1-5.

[8] Apruzzese, G., Colajanni, M., Ferretti, L., Guido, A., & Marchetti, M. (2018). On the effectiveness of machine and deep learning for cyber security. In 2018 10th International Conference on Cyber Conflict, 371-390.

[9] Jan, B., Farman, H., Khan, M., Imran, M., Islam, I. U., Ahmad, A., ... & Jeon, G. (2019). Deep learning in big data analytics: a comparative study. Computers & Electrical Engineering, C. 75, 275-287.

[10] Abeshu, A., & Chilamkurti, N. (2018). Deep learning: The frontier for distributed attack detection in fog-to-things computing. IEEE Communications Magazine, C. 56(2), 169-175.

[11] Yan, R., Xiao, X., Hu, G., Peng, S., & Jiang, Y. (2018). New deep learning method to detect code injection attacks on hybrid applications. Journal of Systems and Software, C. 137, 67-77.

[12] Kang, B., & Choo, H. (2016). A deep-learning-based emergency alert system. ICT Express, C. 2(2), 67-70.

[13] Wang, H., Ruan, J., Wang, G., Zhou, B., Liu, Y., Fu, X., & Peng, J. (2018). Deep learning-based interval state estimation of AC smart grids against sparse cyber attacks. IEEE Transactions on Industrial Informatics, C. 14(11), 4766-4778.

[14] Kim, S., Joshi, P., Kalsi, P. S., & Taheri, P. (2018, November). Crime Analysis Through Machine Learning. In 2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference, 415-420.

[15] Duan, L., Hu, T., Cheng, E., Zhu, J., & Gao, C. (2017). Deep convolutional neural networks for spatiotemporal crime prediction. In Proceedings of the International Conference on Information and Knowledge Engineering, 61-67.

[16] Chandrasekar, A., Raj, A. S., & Kumar, P. (2015). Crime prediction and classification in San Francisco City. URL http://cs229. stanford. edu/proj2015/228 {_} report. pdf.

[17] Ivan, N., Ahishakiye, E., Omulo, E. O., & Taremwa, D. (2017). Crime Prediction Using Decision Tree (J48) Classification Algorithm.

[18] Bharati, A., RA.K, S. (2018). Crime Prediction and Analysis Using Machine Learning, International Research Journal of Engineering and Technology, 1037-1042.

[19] Feng, M., Zheng, J., Han, Y., Ren, J., & Liu, Q. (2018, July). Big data analytics and mining for crime data analysis, visualization and prediction. In International Conference on Brain Inspired Cognitive Systems, Springer, Cham, 605-614.

[20] Gerber, M. S. (2014). Predicting crime using Twitter and kernel density estimation. Decision Support Systems, C. 61, 115-125.

[21] Krishnan, A., Sarguru, A., & Sheela, A. S. (2018). Predicitve analysis of crime data using deep learning. International Journal of Pure and Applied Mathematics, C. 118(20), 4023-4031.

[22] Stec, A., & Klabjan, D. (2018). Forecasting Crime with Deep Learning. arXiv preprint arXiv:1806.01486.

[23] Chackravarthy, S., Schmitt, S., & Yang, L. (2018). Intelligent Crime Anomaly Detection in Smart Cities Using Deep Learning. In 2018 IEEE 4th International Conference on Collaboration and Internet Computing, 399-404.

[24] Kang, H. W., & Kang, H. B. (2017). Prediction of crime occurrence from multi-modal data using deep learning. PloS one, C. 12(4).

[25] Shermila, A. M., Bellarmine, A. B., & Santiago, N. (2018, May). Crime data analysis and prediction of perpetrator identity using machine learning approach. In 2018 2nd International Conference on Trends in Electronics and Informatics, 107-114).

[26] Chintal, P. L., Gaikwad, R. J., & Deshmukh, R. R. (2018). Cyber Crime Analysis of Maharashtra State using Gradient Descent Approach with Linear Regression. International Journal of Pure and Applied Mathematics, C. 119(16), 3537-3542.

[27] Arora, T., Sharma, M., & Khatri, S. K. (2019). Detection of Cyber Crime on Social Media using Random Forest Algorithm. In 2019 2nd International Conference on Power Energy, Environment and Intelligent Control, 47-51.

[28] Ch, R., Gadekallu, T. R., Abidi, M. H., & Al-Ahmari, A. (2020). Computational system to classify cyber crime offenses using machine learning. Sustainability, C. 12(10), 4087.

[29] Sağıroğlu, Ş., Alkan, M. (2018). Siber Güvenlik ve Savunma Farkındalık ve Caydırıcılık, BGD Yayınları, Ankara.

[30] Bayuk, J. L., Healey, J., Rohmeyer, P., Sachs, M. H., Schmidt, J., & Weiss, J. (2012). Cyber security policy guidebook, 3-4, Hoboken: Wiley.

[31] Adalı, E. (2016). Bilgisayar ve Bilgi Güvenliği, İTÜ Yayınları, İstanbul.

[32] https://www.hack2secure.com/blogs/cyber-security-vs-information-security, Erişim: 4 Mart 2020.

[33] Von Solms, R., & Van Niekerk, J. (2013). From information security to cyber security. computers & security, C. 38, 97-102.

[34] Bayindir, R., Colak, I., Fulli, G., & Demirtas, K. (2016). Smart grid technologies and applications. Renewable and Sustainable Energy Reviews, C. 66, 499-516.

[35] Abbott, R. G., McClain, J., Anderson, B., Nauer, K., Silva, A., & Forsythe, C. (2015). Log analysis of cyber security training exercises. Procedia Manufacturing, C. 3, 5088-5094.

[36] Yin, C., Zhu, Y., Fei, J., & He, X. (2017). A deep learning approach for intrusion detection using recurrent neural networks. Ieee Access, C. 5, 21954-21961.

[37] Andress, J., & Winterfeld, S. (2013). Cyber warfare: techniques, tactics and tools for security practitioners, Elsevier.

[38] Jang-Jaccard, J., & Nepal, S. (2014). A survey of emerging threats in cybersecurity. Journal of Computer and System Sciences, C. 80(5), 973-993.

[39] Check Point Güvenlik Raporu (2020), https://research.checkpoint.com/.

[40] Samonas, S., & Coss, D. (2014). The Cia Strikes Back: Redefining Confidentiality, Integrity And Availability In Security. Journal of Information System Security, 10(3).

[41] Biju, J. M., Gopal, N., & Prakash, A. J. (2019). Cyber Attacks And Its Dıfferent Types, International Research Journal of Engineering and Technology, 2395-0056.

[42] Sinha, P., kumar Rai, A., & Bhushan, B. (2019). Information Security threats and attacks with conceivable counteraction. In 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies, C. 1, 1208-1213.

[43] Akbari Roumani, M., Fung, C. C., Rai, S., & Xie, H. (2016). Value analysis of cyber security based on attack types. ITMSOC: Transactions on Innovation and Business Engineering, C. 1, 34-39.

[44] Nguyen, K. K., Hoang, D. T., Niyato, D., Wang, P., Nguyen, D., & Dutkiewicz, E. (2018, April). Cyberattack detection in mobile cloud computing: A deep learning approach. In 2018 IEEE Wireless Communications and Networking Conference, 1-6.

[45] Probst, C. W., Hunker, J., Gollmann, D., & Bishop, M. (2010). Aspects of insider threats. In Insider Threats in Cyber Security, 1-15, Springer, Boston, MA.

[46] Kaur Chahal, J., Bhandari, A., & Behal, S. (2019). Distributed Denial of Service Attacks: A Threat or Challenge. New Review of Information Networking, 24(1), 31-103.

[47] Sahi, A., Lai, D., Li, Y., & Diykh, M. (2017). An efficient DDoS TCP flood attack detection and prevention system in a cloud environment. IEEE Access, C. 5, 6036-6048.

[48] Utomo, S. P., Pramudiono, B., & Muharram, A. (2019). Method to Uncover IP Spoofing Attack On Network Forensics Using NFAT And IP Correlation As Combined Approach. In 2019 International Conference on Information and Communications Technology.

[49] Yu, J., Kim, E., Kim, H., & Huh, J. H. (2020). Design of a Framework to Detect Device Spoofing Attacks Using Network Characteristics. IEEE Consumer Electronics Magazine, C. 9(2), 34-40.

[50] Scott, B., Xu, J., Zhang, J., Brown, A., Clark, E., Yuan, X., ... & Williams, K. (2017, October). An interactive visualization tool for teaching ARP spoofing attack. In 2017 IEEE Frontiers in Education Conference, 1-5.

[51] Batista, L. O., de Silva, G. A., Araújo, V. S., Araújo, V. J. S., Rezende, T. S., Guimarães, A. J., & Souza, P. V. D. C. (2019). Fuzzy neural networks to create an expert system for detecting attacks by sql injection. arXiv preprint arXiv:1901.02868.

[52] Sodagudi, S., Kotha, S. K., & Raju, M. D. (2019). Novel Approaches to Identify and Prevent Cyber Attacks in Web. In 2019 3rd International Conference on Computing Methodologies and Communication, 833-839.

[53] Das, S., Kim, A., Tingle, Z., & Nippert-Eng, C. (2019). All about phishing: Exploring user research through a systematic literature review. arXiv preprint arXiv:1908.05897.

[54] Sahingoz, O. K., Buber, E., Demir, O., & Diri, B. (2019). Machine learning based phishing detection from URLs. Expert Systems with Applications, C. 117, 345-357.

[55] Song, J., Shimamura, J., Eto, M., Inoue, D., & Nakao, K. (2011). Correlation analysis between spamming botnets and malware infected hosts. In 2011 IEEE/IPSJ International Symposium on Applications and the Internet, 372-375.

[56] Breda, F., Barbosa, H., & Morais, T. (2017, March). Social engineering and cyber security. In em Conference: International Technology, Education and Development Conference.

[57] Aldawood, H., & Skinner, G. (2018, December). Educating and raising awareness on cyber security social engineering: A literature review. In 2018 IEEE International Conference on Teaching, Assessment, and Learning for Engineering, 62-68.

[58] Conteh, N. Y., & Schmick, P. J. (2016). Cybersecurity: risks, vulnerabilities and countermeasures to prevent social engineering attacks. International Journal of Advanced Computer Research, C. 6(23), 31.

[59] Ghafir, I., Saleem, J., Hammoudeh, M., Faour, H., Prenosil, V., Jaf, S., ... & Baker, T. (2018). Security threats to critical infrastructure: the human factor. The Journal of Supercomputing, C. 74(10), 4986-5002.

[60] Rehak, D., Senovsky, P., Hromada, M., & Lovecek, T. (2019). Complex approach to assessing resilience of critical infrastructure elements. International Journal of Critical Infrastructure Protection, C. 25, 125-138.

[61] Kovacevic, A., & Nikolic, D. (2015). Cyber attacks on critical infrastructure: Review and challenges. In Handbook of Research on Digital Crime, Cyberspace Security, and Information Assurance, 1-18, IGI Global.

[62] Mikhalevich, I. F., & Trapeznikov, V. A. (2019). Critical infrastructure security: alignment of views. In 2019 Systems of Signals Generating and Processing in the Field of on Board Communications, 1-5.

[63] Kritik altyapılar, https://www.cisa.gov/critical-infrastructure-sectors, Erişim:28.09.2020.

[64] Zhong, C., Lin, T., Liu, P., Yen, J., & Chen, K. (2018). A cyber security data triage operation retrieval system. Computers & Security, C. 76, 12-31 .

[65] Thakur, K., Qiu, M., Gai, K., & Ali, M. L. (2015). An investigation on cyber security threats and security models. In 2015 IEEE 2nd International Conference on Cyber Security and Cloud Computing, 307-311.

[66] Baldassarre, M. T., Santa Barletta, V., Caivano, D., Raguseo, D., & Scalera, M. (2019). Teaching Cyber Security: The HACK-SPACE Integrated Model. In ITASEC.

[67] Sengupta, S., Chowdhary, A., Sabur, A., Alshamrani, A., Huang, D., & Kambhampati, S. (2020). A survey of moving target defenses for network security. IEEE Communications Surveys & Tutorials.

[68] Enoch, S. Y., Hong, J. B., Ge, M., & Kim, D. S. (2020). Composite metrics for network security analysis. arXiv preprint arXiv:2007.03486.

[69] Lam, W. M. W. (2016). Attack-prevention and damage-control investments in cybersecurity. Information Economics and Policy, C. 37, 42-51.

[70] Sani, A. S., Yuan, D., Jin, J., Gao, L., Yu, S., & Dong, Z. Y. (2019). Cyber security framework for Internet of Things-based Energy Internet. Future Generation Computer Systems, C. 93, 849-859.

[71] Pan, L., Zheng, X., Chen, H. X., Luan, T., Bootwala, H., & Batten, L. (2017). Cyber security attacks to modern vehicular systems. Journal of information security and applications, C. 36, 90-100.

[72] Bhardwaj, A., Avasthi, V., & Goundar, S. (2019). Cyber security attacks on robotic platforms. Network Security, C. 10, 13-19.

[73] Spremić, M., & Šimunic, A. (2018). Cyber security challenges in digital economy. In Proceedings of the World Congress on Engineering, C. 1, 341-346.

[74] Elnagdy, S. A., Qiu, M., & Gai, K. (2016). Cyber incident classifications using ontology-based knowledge representation for cybersecurity insurance in financial industry, 301-306.

[75] Elnagdy, S. A., Qiu, M., & Gai, K. (2016). Understanding taxonomy of cyber risks for cybersecurity insurance of financial industry in cloud computing. In 2016 IEEE 3rd International Conference on Cyber Security and Cloud Computing, 295-300.

[76] Singh, S., & Singh, N. (2016, December). Blockchain: Future of financial and cyber security. In 2016 2nd international conference on contemporary computing and informatics, 463-467.

[77] Voeller, J. G. (Ed.). (2014). Cyber Security. John Wiley & Sons.

[78] Lehto, M. (2015). Phenomena in the Cyber World. In Cyber Security: Analytics, Technology and Automation (pp. 3-29). Springer, Cham.

[79] Sedgewick, A. (2014). Framework for improving critical infrastructure cybersecurity, version 1.0 (No. NIST-Cybersecurity Framework).

[80] Sun, C. C., Hahn, A., & Liu, C. C. (2018). Cyber security of a power grid: State-of-the-art. International Journal of Electrical Power & Energy Systems, C. 99, 45-56.

[81] Maglaras, L. A., Kim, K. H., Janicke, H., Ferrag, M. A., Rallis, S., Fragkou, P., & Cruz, T. J. (2018). Cyber security of critical infrastructures. Ict Express, C. 4(1), 42-45.

[82] Park, J., Suh, Y., & Park, C. (2016). Implementation of cyber security for safety systems of nuclear facilities. Progress in Nuclear Energy, C. 88, 88-94.

[83] Moreno, V. C., Reniers, G., Salzano, E., & Cozzani, V. (2018). Analysis of physical and cyber security-related events in the chemical and process industry. Process Safety and Environmental Protection, C. 116, 621-631.

[84] Wang, D., Wang, X., Zhang, Y., & Jin, L. (2019). Detection of power grid disturbances and cyber-attacks based on machine learning. Journal of Information Security and Applications, C. 46, 42-52.

[85] Ficco, M., Choraś, M., & Kozik, R. (2017). Simulation platform for cyber-security and vulnerability analysis of critical infrastructures. Journal of computational science, C. 22, 179-186.

[86] Limba, T., Plėta, T., Agafonov, K., & Damkus, M. (2019). Cyber security management model for critical infrastructure.

[87] Wang, W., & Lu, Z. (2013). Cyber security in the smart grid: Survey and challenges. Computer networks, C. 57(5), 1344-1371.

[88] Mezzour, G., Carley, L., & Carley, K. M. (2014). Global mapping of cyber attacks. Available at SSRN 2729302.

[89] Srinivas, J., Das, A. K., & Kumar, N. (2019). Government regulations in cyber security: Framework, standards and recommendations. Future Generation Computer Systems, C. 92, 178-188.

[90] Wilk, A. (2016, June). Cyber security education and law. In 2016 IEEE International Conference on Software Science, Technology and Engineering, 94-103.

[91] https://ec.europa.eu/digital-single-market/en/eu-cybersecurity-act, Erişim: 03 Aralık 2020.

[92] https://www.dhs.gov/sites/default/files/publications/DHS-Cybersecurity-Strategy_1.pdf, Erişim: 04.12.2020.

[93] https://www.ssi.gouv.fr/en/cybersecurity-in-france/, Erişim: 04.12.2020.

[94] https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data /file/937702/6.6788_CO_National-Cyber-Security-Strategy-2016-2021_WEB3.pdf, Erişim: 04.12.2020.

[95] https://assets.kpmg/content/dam/kpmg/cn/pdf/en/2017/02/overview-of-cybersecurity-law.pdf, Erişim: 06.12.2020 .

[96] https://www.mid.ru/en/foreign_policy/official_documents/-/asset_publisher/CptICkB6BZ29/content/id/2563163, Erişim: 06.12.2020.

[97] https://www.ab.gov.tr/siteimages/birimler/kpb/ulusal_eylem_plani-_yeni.pdf, Erişim: 06.12.2020.

[98] https://www.egm.gov.tr/siber/, Erişim: 06.12.2020.

[99] https://www.mevzuat.gov.tr/MevzuatMetin/1.5.5237.pdf, Erişim: 06.12.2020.

[100] Hulick, K., (2015) Artificial Intelligence, ABDO Publishing Company, ProQuest Ebook Central.

[101] Chowdhury, M. M. U., Hammond, F., Konowicz, G., Xin, C., Wu, H., & Li, J. (2017, October). A few-shot deep learning approach for improved intrusion detection. In 2017 IEEE 8th Annual UEMCON, 456-462.

[102] Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., Al-Nemrat, A., & Venkatraman, S. (2019). Deep learning approach for intelligent intrusion detection system. IEEE Access, C. 7, 41525-41550.

[103] Almseidin, M., Alzubi, M., Kovacs, S., & Alkasassbeh, M. (2017). Evaluation of machine learning algorithms for intrusion detection system. In 2017 IEEE 15th International Symposium on Intelligent Systems and Informatics, 000277-00028.

[104] Fang, X., Xu, M., Xu, S., & Zhao, P. (2019). A deep learning framework for predicting cyber attacks rates. EURASIP Journal on Information Security, C. 1, 5.

[105] Noor, U., Anwar, Z., Amjad, T., & Choo, K. K. R. (2019). A machine learning-based FinTech cyber threat attribution framework using high-level indicators of compromise. Future Generation Computer Systems, C. 96, 227-242.

[106] Feng, F., Liu, X., Yong, B., Zhou, R., & Zhou, Q. (2019). Anomaly detection in ad-hoc networks based on deep learning model: A plug and play device. Ad Hoc Networks, C. 84, 82-89.

[107] He, Y., Mendis, G. J., & Wei, J. (2017). Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism. IEEE Transactions on Smart Grid, C. 8(5), 2505-2516.

[108] Wang, Y., Cai, W. D., & Wei, P. C. (2016). A deep learning approach for detecting malicious JavaScript code. Security and Communication Networks, C. 9(11), 1520-1534.

[109] Cui, Z., Xue, F., Cai, X., Cao, Y., Wang, G. G., & Chen, J. (2018). Detection of malicious code variants based on deep learning. IEEE Transactions on Industrial Informatics, C. 14(7), 3187-3196.

[110] Shahini, M., Farhanian, R., & Ellis, M. (2019). Machine Learning to Predict the Likelihood of a Personal Computer to Be Infected with Malware. SMU Data Science Review, C. 2(2), 9.

[111] Rege, M., & Mbah, R. B. K. (2018). Machine learning for cyber defense and attack. DATA ANALYTICS, 83.

[112] Wu, W., Alvarez, J., Liu, C., & Sun, H. M. (2018). Bot detection using unsupervised machine learning. Microsystem Technologies, C. 24(1), 209-217.

[113] Najjar, A., Kaneko, S. I., & Miyanaga, Y. (2018). Crime Mapping from Satellite Imagery via Deep Learning. arXiv preprint arXiv:1812.06764.

[114] Li, S., Zhang, H., Ye, L., Su, S., Guo, X., Yu, H., & Fang, B. (2020). Prison term prediction on criminal case description with deep learning. Computers, Materials & Continua, C. 62(3), 1217-1231.

[115] Wang, S., & Yuan, K. (2019, July). Spatiotemporal Analysis and Prediction of Crime Events in Atlanta Using Deep Learning. In 2019 IEEE 4th International Conference on Image, Vision and Computing, 346-350.

[116] Yamini, M. P. C. (2019). A violent crime analysis using fuzzy c-means clustering approach. ICTACT Journal on Soft Computing, C. 9(3), 1939-1944.

[117] McClendon, L., & Meghanathan, N. (2015). Using machine learning algorithms to analyze crime data. Machine Learning and Applications: An International Journal, C. 2(1), 1-12.

[118] Back, S., LaPrade, J., Shehadeh, L., & Kim, M. (2019, June). Youth hackers and adult hackers in South Korea: An application of cybercriminal profiling. In 2019 IEEE European Symposium on Security and Privacy Workshops, 410-413.

[119] Mehta, Y., Majumder, N., Gelbukh, A., & Cambria, E. (2019). Recent trends in deep learning based personality detection. Artificial Intelligence Review, 1-27.

[120] Alpaydin, E. (2020). Introduction to machine learning. MIT press.

[121] Alpaydin, E. (2016). Machine learning: the new AI. MIT press.

[122] Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Applied logistic regression, C. 398, John Wiley & Sons.

[123] Pradhan, B. (2010). Landslide susceptibility mapping of a catchment area using frequency ratio, fuzzy logic and multivariate logistic regression approaches. Journal of the Indian Society of Remote Sensing, C. 38(2), 301-320.

[124] https://sebastianraschka.com/faq/docs/logisticregr-neuralnet.html. erişim:08.03.2021.

[125] Leung, K. M. (2007). k-Nearest Neighbor algorithm for classification. Polytechnic University Department of Computer Science/Finance and Risk Engineering.

[126] Schölkopf, B., Smola, A. J., & Bach, F. (2002). Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT press.

[127] Wang, Y., Meng, X., & Zhu, L. (2018). Cell group recognition method based on adaptive mutation PSO-SVM. Cells, C. 7(9), 135.

[128] Wu, X., Kumar, V., Quinlan, J. R., Ghosh, J., Yang, Q., Motoda, H., ... & Steinberg, D. (2008). Top 10 algorithms in data mining. Knowledge and information systems, C. 14(1), 1-37.

[129] Lin, W., Wu, Z., Lin, L., Wen, A., & Li, J. (2017). An ensemble random forest algorithm for insurance big data analysis. Ieee access, C. 5, 16568-16575.

[130] Schonlau, M., & Zou, R. Y. (2020). The random forest algorithm for statistical learning. The Stata Journal, C. 20(1), 3-29.

[131] Jin, C., De-Lin, L., & Fen-Xiang, M. (2009, July). An improved ID3 decision tree algorithm. In 2009 4th International Conference on Computer Science & Education, 127-130.

[132] https://towardsdatascience.com/decision-trees-in-machine-learning-641b9c4e8052, Erişim.23.02.2021.

[133] Mitchell, R., & Frank, E. (2017). Accelerating the XGBoost algorithm using GPU computing. PeerJ Computer Science, C. 3, e127.

[134] Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining, 785-794.

[135] Kim, K. G. (2016). Book review: Deep learning. Healthcare informatics research, C. 22(4), 351.

[136] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. nature, C. 521(7553), 436-444.

[137] Ravì, D., Wong, C., Deligianni, F., Berthelot, M., Andreu-Perez, J., Lo, B., & Yang, G. Z. (2016). Deep learning for health informatics. IEEE journal of biomedical and health informatics, C. 21(1), 4-21.

[138] Kyprianidis, K., & Dahlquist, E. (2021). AI and Learning Systems-Industrial Applications and Future Directions.

[139] Haykin, S. (1994). Neural networks a comprehensive foundation. New York, MacMillan.

[140] Efendigil, T., Önüt, S., & Kahraman, C. (2009). A decision support system for demand forecasting with artificial neural networks and neuro-fuzzy models: A comparative analysis. Expert Systems with Applications, C. 36(3), 6697-6707.

[141] https://www.math.univ-toulouse.fr/~besse/Wikistat/pdf/st-m-hdstat-rnn-deep-learning.pdf, Erişim:10.03.2021.

[142] Xiong, N. (2018). Deep learning, FUDIPO Course, Mälardalen University.

EKLER

EK- 1: SALDIRI AMACI GERÇEK DEĞERLERİ VE TAHMİN DEĞERLERİ

BDPÇ	Bilgisi Dışında Para Çekilmesi
BDİA	Bilgisi Dışında İnternet Alışverişi
D	Dolandırıcılık
KBGOTPÇ	Kendini Banka Görevlisi Olarak Tanıtarak Para Çekilmesi
KBGOTİÇ	Kendini Banka Görevlisi Olarak Tanıtarak İnternet alışverişi
MZV	Manevi Zarar Vermek
TŞ	Tehdit/Şantaj

	Gerçek Değer	**SVML**	**SVMK**	**DT**	**RF**
1	**KBGOTİÇ**	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ
2	**BDİA**	BDİA	BDİA	BDİA	BDİA
3	**D**	D	D	D	D
4	**D**	D	D	D	D
5	**BDİA**	BDİA	BDİA	BDİA	BDİA
6	**BDİA**	BDİA	BDİA	BDİA	BDİA
7	**BDİA**	BDİA	BDİA	BDİA	BDİA
8	**BDİA**	BDİA	BDİA	BDİA	BDİA
9	**BDİA**	BDİA	BDİA	BDİA	BDİA
10	**KBGOTİÇ**	KBGOTİÇ	KBGOTİÇ	KBGOTPÇ	KBGOTPÇ
11	**MZV**	MZV	MZV	MZV	MZV
12	**MZV**	MZV	MZV	MZV	MZV
13	**BDİA**	BDİA	BDİA	BDİA	BDİA
14	**BDİA**	BDİA	BDİA	BDİA	BDİA
15	**D**	D	D	D	D
16	**D**	D	D	D	D
17	**MZV**	D	D	D	D
18	**KBGOTİÇ**	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ
19	**D**	D	D	D	D
20	**MZV**	D	D	D	D
21	**BDİA**	BDİA	BDİA	BDİA	BDİA
22	**BDİA**	BDİA	BDİA	BDİA	BDİA
23	**D**	D	D	D	D
24	**D**	D	D	D	D
25	**KBGOTİÇ**	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ
26	**KBGOTİÇ**	KBGOTPÇ	KBGOTPÇ	KBGOTPÇ	KBGOTPÇ
27	**BDPÇ**	BDPÇ	BDPÇ	BDPÇ	BDPÇ
28	**KBGOTİÇ**	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ
29	**D**	D	D	D	D
30	**BDPÇ**	BDİA	BDİA	BDİA	BDİA
31	**D**	D	D	D	D
32	**KBGOTPÇ**	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ

33	**D**	D	D	D	D
34	**D**	MZV	MZV	MZV	D
35	**D**	D	D	D	D
36	**BDİA**	BDİA	BDİA	BDİA	BDİA
37	**BDPÇ**	BDPÇ	BDPÇ	BDPÇ	BDPÇ
38	**D**	D	D	D	D
39	**BDPÇ**	BDPÇ	BDPÇ	BDPÇ	BDPÇ
40	**BDPÇ**	BDPÇ	BDPÇ	BDPÇ	BDPÇ
41	**KBGOTİÇ**	KBGOTPÇ	KBGOTPÇ	KBGOTPÇ	KBGOTPÇ
42	**MZV**	MZV	MZV	MZV	MZV
43	**BDPÇ**	BDİA	BDİA	BDİA	BDİA
44	**BDPÇ**	BDPÇ	BDPÇ	BDPÇ	BDPÇ
45	**BDİA**	BDİA	BDİA	BDİA	BDİA
46	**D**	D	D	D	D
47	**D**	D	D	D	D
48	**BDİA**	BDİA	BDİA	BDİA	BDİA
49	**BDPÇ**	BDPÇ	BDPÇ	BDPÇ	BDPÇ
50	**BDİA**	BDİA	BDİA	BDİA	BDİA
51	**BDİA**	BDİA	BDİA	BDİA	BDİA
52	**D**	D	D	D	D
53	**BDİA**	BDİA	BDİA	BDİA	BDİA
54	**MZV**	D	D	D	D
55	**BDİA**	BDİA	BDİA	BDİA	BDİA
56	**BDPÇ**	BDPÇ	BDPÇ	BDPÇ	BDPÇ
57	**D**	D	D	D	D
58	**D**	D	D	D	D
59	**D**	D	D	D	D
60	**BDİA**	BDİA	BDİA	BDİA	BDİA
61	**BDPÇ**	BDİA	BDİA	BDİA	BDİA
62	**MZV**	MZV	MZV	MZV	MZV
63	**MZV**	D	D	MZV	MZV
64	**BDİA**	BDİA	BDİA	BDİA	BDİA
65	**BDİA**	BDİA	BDİA	BDİA	BDİA
66	**TŞ**	MZV	MZV	MZV	MZV
67	**BDPÇ**	BDPÇ	BDPÇ	BDPÇ	BDPÇ
68	**D**	MZV	MZV	MZV	D
69	**KBGOTİÇ**	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ
70	**MZV**	MZV	MZV	MZV	MZV
71	**TŞ**	MZV	MZV	MZV	MZV
72	**KBGOTPÇ**	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ
73	**KBGOTPÇ**	KBGOTPÇ	KBGOTPÇ	KBGOTPÇ	KBGOTPÇ
74	**MZV**	D	D	D	D
75	**BDİA**	BDİA	BDİA	BDİA	BDİA
76	**D**	D	D	D	D
77	**BDPÇ**	BDİA	BDİA	BDİA	BDİA
78	**MZV**	MZV	MZV	MZV	D
79	**D**	D	D	D	D

80	**D**	D	D	D	D
81	**KBGOTİÇ**	KBGOTİÇ	KBGOTİÇ	KBGOTPÇ	KBGOTPÇ
82	**MZV**	MZV	MZV	D	MZV
83	**D**	D	D	D	MZV
84	**BDİA**	BDİA	BDİA	BDİA	BDİA
85	**D**	D	D	D	D
86	**BDİA**	BDİA	BDİA	BDİA	BDİA
87	**D**	D	D	D	D
88	**D**	D	D	D	D
89	**BDİA**	BDİA	BDİA	BDİA	BDİA
90	**MZV**	D	D	D	D
91	**BDİA**	BDİA	BDİA	BDİA	BDİA
92	**BDİA**	BDİA	BDİA	BDİA	BDİA
93	**MZV**	D	D	MZV	MZV
94	**BDİA**	BDİA	BDİA	BDİA	BDİA
95	**BDPÇ**	BDPÇ	BDPÇ	BDPÇ	BDPÇ
96	**MZV**	D	D	D	D
97	**BDİA**	BDİA	BDİA	BDİA	BDİA
98	**D**	MZV	MZV	MZV	D
99	**D**	D	D	D	D
100	**BDİA**	BDİA	BDİA	BDİA	BDİA
101	**MZV**	MZV	MZV	MZV	D
102	**BDİA**	BDİA	BDİA	BDİA	BDİA
103	**D**	D	D	D	D
104	**KBGOTPÇ**	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ
105	**BDPÇ**	BDPÇ	BDPÇ	BDPÇ	BDPÇ
106	**TŞ**	MZV	MZV	MZV	D
107	**KBGOTİÇ**	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ
108	**BDPÇ**	BDPÇ	BDPÇ	BDPÇ	BDPÇ
109	**BDİA**	BDİA	BDİA	BDİA	BDİA
110	**MZV**	MZV	MZV	MZV	MZV
111	**BDİA**	BDİA	BDİA	BDİA	BDİA
112	**BDİA**	BDİA	BDİA	BDİA	BDİA
113	**D**	D	D	D	D
114	**BDPÇ**	BDPÇ	BDPÇ	BDPÇ	BDPÇ
115	**D**	D	D	D	D
116	**BDPÇ**	BDPÇ	BDPÇ	BDPÇ	BDPÇ
117	**BDİA**	BDİA	BDİA	BDİA	BDİA
118	**MZV**	D	D	D	D
119	**D**	D	D	D	D
120	**MZV**	D	D	D	D
121	**D**	D	D	D	D
122	**BDPÇ**	BDPÇ	BDPÇ	BDPÇ	BDPÇ
123	**BDİA**	BDİA	BDİA	BDİA	BDİA
124	**BDPÇ**	BDPÇ	BDPÇ	BDPÇ	BDPÇ
125	**D**	D	D	D	D
126	**BDİA**	BDİA	BDİA	BDİA	BDİA

127	**BDİA**	BDİA	BDİA	BDİA	BDİA
128	**MZV**	MZV	MZV	MZV	MZV
129	**D**	D	D	D	D
130	**D**	D	D	D	D
131	**MZV**	D	D	D	D
132	**BDPÇ**	BDPÇ	BDPÇ	BDPÇ	BDPÇ
133	**TŞ**	MZV	MZV	MZV	MZV
134	**D**	D	D	D	D
135	**MZV**	D	D	MZV	MZV
136	**D**	D	D	D	D
137	**KBGOTPÇ**	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ
138	**KBGOTPÇ**	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ
139	**BDİA**	BDİA	BDİA	BDİA	BDİA
140	**BDİA**	BDİA	BDİA	BDİA	BDİA
141	**MZV**	D	D	D	D
142	**MZV**	MZV	MZV	MZV	MZV
143	**BDPÇ**	BDPÇ	BDPÇ	BDPÇ	BDPÇ
144	**D**	D	D	D	D
145	**BDİA**	BDİA	BDİA	BDİA	BDİA
146	**TŞ**	MZV	MZV	MZV	MZV
147	**BDPÇ**	BDİA	BDİA	BDİA	BDİA
148	**MZV**	D	D	D	D
149	**BDİA**	BDİA	BDİA	BDİA	BDİA
150	**BDPÇ**	BDİA	BDİA	BDİA	BDİA
151	**D**	D	D	D	D
152	**KBGOTİÇ**	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ
153	**BDİA**	BDİA	BDİA	BDİA	BDİA
154	**MZV**	MZV	MZV	MZV	MZV
155	**BDİA**	BDİA	BDİA	BDİA	BDİA
156	**D**	D	D	D	D
157	**D**	MZV	MZV	MZV	D
158	**KBGOTPÇ**	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ
159	**BDPÇ**	BDİA	BDİA	BDİA	BDİA
160	**BDPÇ**	BDPÇ	BDPÇ	BDPÇ	BDPÇ
161	**MZV**	MZV	MZV	MZV	MZV
162	**KBGOTİÇ**	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ
163	**D**	D	D	D	MZV
164	**BDİA**	BDİA	BDİA	BDİA	BDİA
165	**BDPÇ**	BDİA	BDİA	BDİA	BDİA
166	**D**	D	D	D	D
167	**KBGOTİÇ**	KBGOTPÇ	KBGOTİÇ	KBGOTPÇ	KBGOTPÇ
168	**D**	D	D	D	D
169	**BDİA**	BDİA	BDİA	BDİA	BDİA
170	**BDİA**	BDİA	BDİA	BDİA	BDİA
171	**MZV**	MZV	MZV	MZV	D
172	**BDİA**	BDİA	BDİA	BDİA	BDİA
173	**BDİA**	BDİA	BDİA	BDİA	BDİA

174	**D**	D	D	D	D
175	**BDPÇ**	BDPÇ	BDPÇ	BDPÇ	BDPÇ
176	**BDİA**	BDİA	BDİA	BDİA	BDİA
177	**BDİA**	BDİA	BDİA	BDİA	BDİA
178	**KBGOTPÇ**	KBGOTPÇ	KBGOTİÇ	KBGOTPÇ	KBGOTPÇ
179	**BDİA**	BDİA	BDİA	BDİA	BDİA
180	**KBGOTPÇ**	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ	KBGOTİÇ
181	**D**	D	D	D	D

EK- 2: FAİL GERÇEK DEĞERLERİ VE TAHMİN DEĞERLERİ

	Tahmin (KNN)	Gerçek Değer		Tahmin (KNN)	Gerçek Değer
1	Faili Bilinen	**Faili Bilinen**	92	Faili Bilinmeyen	**Faili Bilinen**
2	Faili Bilinmeyen	**Faili Bilinen**	93	Faili Bilinmeyen	**Faili Bilinmeyen**
3	Faili Bilinmeyen	**Faili Bilinmeyen**	94	Faili Bilinmeyen	**Faili Bilinen**
4	Faili Bilinmeyen	**Faili Bilinmeyen**	95	Faili Bilinen	**Faili Bilinen**
5	Faili Bilinmeyen	**Faili Bilinmeyen**	96	Faili Bilinmeyen	**Faili Bilinmeyen**
6	Faili Bilinmeyen	**Faili Bilinmeyen**	97	Faili Bilinen	**Faili Bilinen**
7	Faili Bilinmeyen	**Faili Bilinmeyen**	98	Faili Bilinmeyen	**Faili Bilinen**
8	Faili Bilinmeyen	**Faili Bilinmeyen**	99	Faili Bilinmeyen	**Faili Bilinmeyen**
9	Faili Bilinen	**Faili Bilinmeyen**	100	Faili Bilinmeyen	**Faili Bilinmeyen**
10	Faili Bilinen	**Faili Bilinen**	101	Faili Bilinen	**Faili Bilinen**
11	Faili Bilinmeyen	**Faili Bilinmeyen**	102	Faili Bilinmeyen	**Faili Bilinen**
12	Faili Bilinmeyen	**Faili Bilinen**	103	Faili Bilinmeyen	**Faili Bilinen**
13	Faili Bilinmeyen	**Faili Bilinmeyen**	104	Faili Bilinen	**Faili Bilinen**
14	Faili Bilinen	**Faili Bilinen**	105	Faili Bilinen	**Faili Bilinen**
15	Faili Bilinen	**Faili Bilinen**	106	Faili Bilinmeyen	**Faili Bilinen**
16	Faili Bilinmeyen	**Faili Bilinmeyen**	107	Faili Bilinen	**Faili Bilinen**
17	Faili Bilinmeyen	**Faili Bilinmeyen**	108	Faili Bilinen	**Faili Bilinen**
18	Faili Bilinen	**Faili Bilinen**	109	Faili Bilinen	**Faili Bilinmeyen**
19	Faili Bilinen	**Faili Bilinen**	110	Faili Bilinen	**Faili Bilinen**
20	Faili Bilinmeyen	**Faili Bilinen**	111	Faili Bilinmeyen	**Faili Bilinen**
21	Faili Bilinen	**Faili Bilinen**	112	Faili Bilinmeyen	**Faili Bilinmeyen**
22	Faili Bilinen	**Faili Bilinen**	113	Faili Bilinmeyen	**Faili Bilinmeyen**
23	Faili Bilinmeyen	**Faili Bilinmeyen**	114	Faili Bilinmeyen	**Faili Bilinen**
24	Faili Bilinen	**Faili Bilinen**	115	Faili Bilinen	**Faili Bilinen**
25	Faili Bilinen	**Faili Bilinmeyen**	116	Faili Bilinmeyen	**Faili Bilinmeyen**
26	Faili Bilinen	**Faili Bilinmeyen**	117	Faili Bilinmeyen	**Faili Bilinen**
27	Faili Bilinmeyen	**Faili Bilinen**	118	Faili Bilinmeyen	**Faili Bilinmeyen**
28	Faili Bilinen	**Faili Bilinmeyen**	119	Faili Bilinmeyen	**Faili Bilinmeyen**
29	Faili Bilinmeyen	**Faili Bilinmeyen**	120	Faili Bilinmeyen	**Faili Bilinmeyen**
30	Faili Bilinmeyen	**Faili Bilinen**	121	Faili Bilinmeyen	**Faili Bilinmeyen**
31	Faili Bilinmeyen	**Faili Bilinmeyen**	122	Faili Bilinmeyen	**Faili Bilinmeyen**
32	Faili Bilinen	**Faili Bilinen**	123	Faili Bilinmeyen	**Faili Bilinmeyen**
33	Faili Bilinmeyen	**Faili Bilinen**	124	Faili Bilinen	**Faili Bilinen**
34	Faili Bilinen	**Faili Bilinmeyen**	125	Faili Bilinmeyen	**Faili Bilinmeyen**
35	Faili Bilinen	**Faili Bilinen**	126	Faili Bilinmeyen	**Faili Bilinmeyen**
36	Faili Bilinmeyen	**Faili Bilinmeyen**	127	Faili Bilinen	**Faili Bilinmeyen**
37	Faili Bilinmeyen	**Faili Bilinmeyen**	128	Faili Bilinen	**Faili Bilinmeyen**
38	Faili Bilinmeyen	**Faili Bilinmeyen**	129	Faili Bilinmeyen	**Faili Bilinen**
39	Faili Bilinmeyen	**Faili Bilinen**	130	Faili Bilinmeyen	**Faili Bilinmeyen**
40	Faili Bilinen	**Faili Bilinmeyen**	131	Faili Bilinmeyen	**Faili Bilinmeyen**
41	Faili Bilinen	**Faili Bilinmeyen**	132	Faili Bilinmeyen	**Faili Bilinen**
42	Faili Bilinmeyen	**Faili Bilinen**	133	Faili Bilinen	**Faili Bilinmeyen**
43	Faili Bilinen	**Faili Bilinmeyen**	134	Faili Bilinen	**Faili Bilinen**
44	Faili Bilinen	**Faili Bilinen**	135	Faili Bilinmeyen	**Faili Bilinen**
45	Faili Bilinen	**Faili Bilinmeyen**	136	Faili Bilinmeyen	**Faili Bilinmeyen**
46	Faili Bilinmeyen	**Faili Bilinmeyen**	137	Faili Bilinen	**Faili Bilinen**

47	Faili Bilinmeyen	**Faili Bilinmeyen**
48	Faili Bilinmeyen	**Faili Bilinmeyen**
49	Faili Bilinmeyen	**Faili Bilinmeyen**
50	Faili Bilinmeyen	**Faili Bilinmeyen**
51	Faili Bilinmeyen	**Faili Bilinen**
52	Faili Bilinmeyen	**Faili Bilinmeyen**
53	Faili Bilinen	**Faili Bilinen**
54	Faili Bilinmeyen	**Faili Bilinmeyen**
55	Faili Bilinmeyen	**Faili Bilinmeyen**
56	Faili Bilinen	**Faili Bilinen**
57	Faili Bilinmeyen	**Faili Bilinmeyen**
58	Faili Bilinmeyen	**Faili Bilinmeyen**
59	Faili Bilinen	**Faili Bilinmeyen**
60	Faili Bilinmeyen	**Faili Bilinmeyen**
61	Faili Bilinmeyen	**Faili Bilinmeyen**
62	Faili Bilinmeyen	**Faili Bilinmeyen**
63	Faili Bilinmeyen	**Faili Bilinmeyen**
64	Faili Bilinmeyen	**Faili Bilinmeyen**
65	Faili Bilinmeyen	**Faili Bilinmeyen**
66	Faili Bilinmeyen	**Faili Bilinmeyen**
67	Faili Bilinmeyen	**Faili Bilinen**
68	Faili Bilinen	**Faili Bilinmeyen**
69	Faili Bilinen	**Faili Bilinen**
70	Faili Bilinen	**Faili Bilinmeyen**
71	Faili Bilinmeyen	**Faili Bilinmeyen**
72	Faili Bilinen	**Faili Bilinen**
73	Faili Bilinen	**Faili Bilinen**
74	Faili Bilinmeyen	**Faili Bilinmeyen**
75	Faili Bilinen	**Faili Bilinen**
76	Faili Bilinmeyen	**Faili Bilinmeyen**
77	Faili Bilinen	**Faili Bilinen**
78	Faili Bilinmeyen	**Faili Bilinmeyen**
79	Faili Bilinmeyen	**Faili Bilinmeyen**
80	Faili Bilinmeyen	**Faili Bilinmeyen**
81	Faili Bilinmeyen	**Faili Bilinmeyen**
82	Faili Bilinmeyen	**Faili Bilinmeyen**
83	Faili Bilinmeyen	**Faili Bilinmeyen**
84	Faili Bilinen	**Faili Bilinen**
85	Faili Bilinmeyen	**Faili Bilinmeyen**
86	Faili Bilinen	**Faili Bilinen**
87	Faili Bilinen	**Faili Bilinen**
88	Faili Bilinmeyen	**Faili Bilinmeyen**
89	Faili Bilinmeyen	**Faili Bilinen**
90	Faili Bilinmeyen	**Faili Bilinmeyen**
91	Faili Bilinmeyen	**Faili Bilinmeyen**
138	Faili Bilinen	**Faili Bilinen**
139	Faili Bilinmeyen	**Faili Bilinen**
140	Faili Bilinen	**Faili Bilinen**
141	Faili Bilinmeyen	**Faili Bilinmeyen**
142	Faili Bilinen	**Faili Bilinen**
143	Faili Bilinmeyen	**Faili Bilinen**
144	Faili Bilinmeyen	**Faili Bilinmeyen**
145	Faili Bilinmeyen	**Faili Bilinmeyen**
146	Faili Bilinmeyen	**Faili Bilinen**
147	Faili Bilinmeyen	**Faili Bilinmeyen**
148	Faili Bilinmeyen	**Faili Bilinmeyen**
149	Faili Bilinmeyen	**Faili Bilinmeyen**
150	Faili Bilinen	**Faili Bilinen**
151	Faili Bilinmeyen	**Faili Bilinmeyen**
152	Faili Bilinen	**Faili Bilinen**
153	Faili Bilinmeyen	**Faili Bilinmeyen**
154	Faili Bilinen	**Faili Bilinen**
155	Faili Bilinen	**Faili Bilinmeyen**
156	Faili Bilinen	**Faili Bilinen**
157	Faili Bilinmeyen	**Faili Bilinmeyen**
158	Faili Bilinen	**Faili Bilinmeyen**
159	Faili Bilinen	**Faili Bilinmeyen**
160	Faili Bilinen	**Faili Bilinen**
161	Faili Bilinen	**Faili Bilinmeyen**
162	Faili Bilinen	**Faili Bilinmeyen**
163	Faili Bilinmeyen	**Faili Bilinmeyen**
164	Faili Bilinen	**Faili Bilinen**
165	Faili Bilinen	**Faili Bilinmeyen**
166	Faili Bilinmeyen	**Faili Bilinen**
167	Faili Bilinen	**Faili Bilinmeyen**
168	Faili Bilinmeyen	**Faili Bilinmeyen**
169	Faili Bilinen	**Faili Bilinen**
170	Faili Bilinen	**Faili Bilinen**
171	Faili Bilinmeyen	**Faili Bilinen**
172	Faili Bilinmeyen	**Faili Bilinmeyen**
173	Faili Bilinmeyen	**Faili Bilinen**
174	Faili Bilinmeyen	**Faili Bilinmeyen**
175	Faili Bilinen	**Faili Bilinen**
176	Faili Bilinen	**Faili Bilinmeyen**
177	Faili Bilinmeyen	**Faili Bilinen**
178	Faili Bilinen	**Faili Bilinen**
179	Faili Bilinen	**Faili Bilinen**
180	Faili Bilinen	**Faili Bilinen**
181	Faili Bilinmeyen	**Faili Bilinen**

EK- 3: SALDIRI YÖNTEMİ GERÇEK DEĞERLERİ VE TAHMİN DEĞERLERİ

	Gerçek Değer	Tahmin LR
1	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
2	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
3	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
4	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
5	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
6	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
7	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
8	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
9	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
10	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
11	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
12	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
13	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
14	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
15	Sahte Alışveriş Sitesi Oluşturarak	Sahte Alışveriş Sitesi Oluşturarak
16	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
17	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
18	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
19	Sahte Alışveriş Sitesi Oluşturarak	Sahte Alışveriş Sitesi Oluşturarak
20	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
21	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
22	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
23	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
24	Sahte Alışveriş Sitesi Oluşturarak	Sahte Alışveriş Sitesi Oluşturarak
25	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
26	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
27	Kart Kopyalama, Üretme Cihazlarını Kullanarak	Kart Kopyalama, Üretme Cihazlarını Kullanarak
28	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
29	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
30	Hack Araçları veya Zararlı Yazılım Kullanarak	Phising (oltalama) Saldırısı Kullanarak
31	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
32	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
33	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
34	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
35	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
36	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
37	Kart Kopyalama, Üretme Cihazlarını Kullanarak	Kart Kopyalama, Üretme Cihazlarını Kullanarak
38	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
39	Kart Kopyalama, Üretme Cihazlarını Kullanarak	Kart Kopyalama, Üretme Cihazlarını Kullanarak
40	Kart Kopyalama, Üretme Cihazlarını Kullanarak	Kart Kopyalama, Üretme Cihazlarını Kullanarak
41	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
42	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
43	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
44	Phising (oltalama) Saldırısı Kullanarak	Phising (oltalama) Saldırısı Kullanarak

45	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
46	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
47	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
48	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
49	Kart Kopyalama, Üretme Cihazlarını Kullanarak	Kart Kopyalama, Üretme Cihazlarını Kullanarak
50	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
51	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
52	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
53	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
54	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
55	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
56	Phising (oltalama) Saldırısı Kullanarak	Phising (oltalama) Saldırısı Kullanarak
57	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
58	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
59	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
60	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
61	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
62	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
63	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
64	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
65	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
66	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
67	Kart Kopyalama, Üretme Cihazlarını Kullanarak	Kart Kopyalama, Üretme Cihazlarını Kullanarak
68	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
69	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
70	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
71	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
72	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
73	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
74	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
75	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
76	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
77	Hack Araçları veya Zararlı Yazılım Kullanarak	Phising (oltalama) Saldırısı Kullanarak
78	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
79	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
80	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
81	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
82	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
83	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
84	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
85	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
86	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
87	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
88	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
89	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
90	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
91	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak

92	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
93	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
94	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
95	Phising (oltalama) Saldırısı Kullanarak	Phising (oltalama) Saldırısı Kullanarak
96	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
97	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
98	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
99	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
100	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
101	Sosyal Mühendislik Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
102	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
103	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
104	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
105	Phising (oltalama) Saldırısı Kullanarak	Phising (oltalama) Saldırısı Kullanarak
106	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
107	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
108	Phising (oltalama) Saldırısı Kullanarak	Phising (oltalama) Saldırısı Kullanarak
109	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
110	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
111	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
112	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
113	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
114	Kart Kopyalama, Üretme Cihazlarını Kullanarak	Kart Kopyalama, Üretme Cihazlarını Kullanarak
115	Sahte Alışveriş Sitesi Oluşturarak	Sahte Alışveriş Sitesi Oluşturarak
116	Kart Kopyalama, Üretme Cihazlarını Kullanarak	Kart Kopyalama, Üretme Cihazlarını Kullanarak
117	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
118	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
119	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
120	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
121	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
122	Kart Kopyalama, Üretme Cihazlarını Kullanarak	Kart Kopyalama, Üretme Cihazlarını Kullanarak
123	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
124	Phising (oltalama) Saldırısı Kullanarak	Phising (oltalama) Saldırısı Kullanarak
125	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
126	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
127	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
128	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
129	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
130	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
131	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
132	Kart Kopyalama, Üretme Cihazlarını Kullanarak	Kart Kopyalama, Üretme Cihazlarını Kullanarak
133	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
134	Phising (oltalama) Saldırısı Kullanarak	Sosyal Mühendislik Kullanarak
135	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
136	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
137	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
138	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak

139	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
140	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
141	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
142	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
143	Kart Kopyalama, Üretme Cihazlarını Kullanarak	Kart Kopyalama, Üretme Cihazlarını Kullanarak
144	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
145	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
146	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
147	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
148	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
149	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
150	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
151	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
152	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
153	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
154	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
155	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
156	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
157	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
158	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
159	Hack Araçları veya Zararlı Yazılım Kullanarak	Phising (oltalama) Saldırısı Kullanarak
160	Phising (oltalama) Saldırısı Kullanarak	Phising (oltalama) Saldırısı Kullanarak
161	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
162	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
163	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
164	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
165	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
166	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak
167	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
168	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
169	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
170	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
171	Sosyal Mühendislik Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
172	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
173	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
174	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
175	Phising (oltalama) Saldırısı Kullanarak	Phising (oltalama) Saldırısı Kullanarak
176	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
177	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
178	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
179	Hack Araçları veya Zararlı Yazılım Kullanarak	Hack Araçları veya Zararlı Yazılım Kullanarak
180	Sosyal Mühendislik Kullanarak	Sosyal Mühendislik Kullanarak
181	Sosyal Medyadaki Herkese Açık Verilerini Alarak	Sosyal Medyadaki Herkese Açık Verilerini Alarak