**T.C.**

**HİTİT ÜNİVERSİTESİ**

**LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**

**ADLİ BİLİMLER ANABİLİM DALI**

# MAKİNE ÖĞRENİMİ ALGORİTMALARI İLE KREDİ KARTI İŞLEMLERİNDE DOLANDIRICILIK TESPİTİ

**YÜKSEK LİSANS TEZİ**

**Nuri KILIÇ**

**Çorum –2023**

# MAKİNE ÖĞRENİMİ ALGORİTMALARI İLE KREDİ KARTI İŞLEMLERİNDE DOLANDIRICILIK TESPİTİ

**Nuri KILIÇ**

**Lisansüstü Eğitim Enstitüsü**

**Adli Bilimler Anabilim Dalı**

**Yüksek Lisans Tezi**

**TEZ DANIŞMANI**

**Dr. Öğr. Üyesi Ömer Faruk AKMEŞE**

**ÇORUM 2023**

**KABUL ONAY SAYFASI**

Nuri KILIÇ tarafından hazırlanan "Makine Öğrenimi Algoritmaları ile Kredi Kartı İşlemlerinde Dolandırıcılık Tespiti" adlı tez çalışması 30/01/2023 tarihinde aşağıdaki jüri üyeleri tarafından oy birliği/oy çokluğu ile Hitit Üniversitesi Lisansüstü Eğitim Enstitüsü Adli Bilimler Anabilim Dalında Yüksek Lisans tezi olarak kabul edilmiştir.

(Dr. Öğr. Üyesi Ömer Faruk AKMEŞE)** ……

(Dr. Öğr Üyesi Hakan KÖR) ……

(Dr. Öğr. Üyesi Fahrettin HORASAN)* ……

Hitit Üniversitesi Lisansüstü Eğitim Enstitüsü Yönetim Kurulunun …/…/…… tarih ve …….. sayılı kararı ile Nuri KILIÇ'ın Adli Bilimler Anabilim Dalında Yüksek Lisans derecesi alması onanmıştır.

(İmza)

Prof. Dr. Muhammed Asıf YOLDAŞ

Lisansüstü Eğitim Enstitüsü Müdürü

**TEZ BEYANI**

Tez içindeki bütün bilgilerin etik davranış ve akademik kurallar çerçevesinde elde edilerek sunulduğunu, ayrıca tez yazım kurallarına uygun olarak hazırlanan bu çalışmada bana ait olmayan her türlü ifade ve bilginin kaynağına eksiksiz atıf yapıldığını beyan ederim.

Nuri KILIÇ

**MAKİNE ÖĞRENİMİ ALGORİTMALARI İLE KREDİ KARTI İŞLEMLERİNDE DOLANDIRICILIK TESPİTİ**

Nuri KILIÇ

ORCID: 0000-0003-1503-9592

HİTİT ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

Yüksek Lisans Tezi

Ocak 2023

**ÖZET**

Bilgisayar ve bilgisayarlı sistemler günümüzde büyük bir öneme sahiptir. Bir de bu sistemlerde Yapay Zekâ 'nın başlıklarından birisi olan Makine Öğrenmesi yöntemini kullandığımızda bir bilgisayarın insan olmadan neler yapabileceğini görebilmekteyiz. Bu sistemleri daha vazgeçilmez bir hale getiren internet kavramı vardır.

Hayatımızda büyük bir rol oynayan internetin, birde karanlık bir yüzü vardır. İnternet artık günümüzde vazgeçilmez bir konumdadır. Uzaktan her bilgiye erişebilme, her cihazı kontrol edebilme gibi birçok imkân sağlayan internet, her ne kadar işimizi kolaylaştırsa da internete bağlı olan cihazlar risk altında olabilir. Bu yüzden suça meyilli veya suçlu kişiler bu ortama yönelmektedir. İnternet ortamında işlenen suçlara Siber Suçlar denilmektedir.

Günümüzde insanlar alışverişlerini yaygın bir şekilde kredi kartları ile internetten yapmaktadırlar. Her ne kadar güvenilir web siteleri de olsa alışveriş yaptığımız bir cihaza bulaştırılmış olan bir Malware, kredi kartı bilgilerini çalma yöntemlerinden birisidir. Bu tez, kredi kartı ile yapılan alışverişlerin, çalınmış veya kopyalanmış bir kredi kartı kullanan dolandırıcılar tarafından mı yapıldığını, Yapay Zekâ ile tespit edilmesi üzerine yapılan bir uygulamayı içermektedir.

**Anahtar Kavramlar:** Yapay Zekâ, Yapay Öğrenme, Adli Bilişim

**Bilim Kodu:** 92432, 92431, 92401

**DETECTION OF FRAUD IN CREDIT CARD TRANSACTIONS WITH MACHINE LEARNING ALGORITHMS**

Nuri KILIC

ORCID: 0000-0003-1503-9592

HITIT UNIVERSITY

GRADUATE EDUCATION INSTITUTE

Master Thesis

January 2023

## ABSTRACT

Todays, computers and systems have a computer have great importance. Also, when we use the Machine Learning method, one of the titles of Artificial Intelligence, in these systems, we can see what a computer can do without a human. There is the concept of the Internet that makes these systems more indispensable.

The Internet, which plays a significant role in our lives, also has a dark side. The Internet is now indispensable today. Although the Internet, which provides many opportunities, such as accessing any information and controlling every device, makes our work easier, devices connected to the Internet may be at risk. Therefore, criminals tend to this environment. Crimes committed on the Internet are called Cyber Crimes.

Today, people do their shopping online with credit cards. Malware that has infected a device with which we shop is one of the methods of stealing credit card information, although there are reliable websites. This thesis includes an application made with Artificial Intelligence to determine whether the purchases made with a credit card are made by fraudsters using a stolen or copied credit card.

**Key Terms:** Artificial Intelligence, Machine Learning, Forensic Science

**Science Code:** 92432, 92431, 92401

**TEŞEKKÜR**

Çalışmalarımda büyük katkıları olan, desteğini, bilgisini ve vaktini hiç esirgemeyen danışman hocam Dr. Öğr. Üyesi Ömer Faruk AKMEŞE 'ye,

Tez Savunması'nda yer alan ve kıymetli görüşleriyle araştırmama katkıda bulunan değerli jüri üyesi hocalarım Dr. Öğr Üyesi Hakan KÖR ve Dr. Öğr. Üyesi Fahrettin HORASAN'a,

Her koşulda arkamda duran canım Ailem, Süleyman ve Aynur Kılıç'a ve şirinlikleriyle ilham veren canım yeğenim Esila Kaplan'a,

Yardımlarını hiç esirgemeyen çocukluk arkadaşım Çelebi Yaşar'a,

sonsuz teşekkür ederim.

Nuri KILIÇ

## İÇİNDEKİLER

**Sayfa**

**ÖZET** ..... **iv**

**ABSTRACT** ..... **v**

**TEŞEKKÜR** ..... **vi**

**İÇİNDEKİLER** ..... **vii**

**TABLOLAR DİZİNİ** ..... **ix**

**ŞEKİLLER DİZİNİ** ..... **x**

**GİRİŞ** ..... **1**

**1. BÖLÜM**

**SİBER İLE İLGİLİ TEMEL KAVRAMLAR**

1.1. Siber Güvenlik Nedir? ..... 4

1.2. Siber Güvenlikte Temel Prensipler ..... 4

1.3. Siber Güvenlik Tarihçesi ..... 5

1.4. Siber Saldırı Nedir? ..... 7

1.5. Siber Saldırı Türleri ..... 10

1.5.1. ZeroDay ..... 10

1.5.2. Malware ..... 10

1.5.3. Phishing (oltalama) ..... 16

1.5.4. Servis dışı bırakma saldırıları ..... 17

1.5.5. Spam ..... 17

1.6. Siber Suçlu ..... 17

**2. BÖLÜM**

**YAPAY ZEKÂ**

2.1. Yapay Zekâ Nedir? ..... 19

2.2. Yapay Zekanın Tarihçesi ..... 20

2.3. Yapay Zekâ Türleri ..... 20
2.3.1. Makine öğrenmesi (machine learning) ..... 20

**3. BÖLÜM**

**MATERYAL VE METOT**

3.1. Veri Seti Tasarımı ..... 25
3.2. Simülasyon Planlaması ..... 26
3.3. Simülasyon Planlanmasının Adımları ..... 26
3.3.1. Müşteri profil tablosu ..... 26
3.3.2. İşlem yeri tablosu ..... 27
3.3.3. Müşteri tablosunun işlem yeri tablosuna bağlanması ..... 27
3.3.4. Veri setinin üretimi ..... 29
3.3.5. Büyük veri seti üretimi ..... 31
3.3.6. Simülasyonun oluşturulması ..... 33
3.4. Makine Öğrenmesi Modeli ..... 35
3.5. Makine Öğrenmesi Algoritmaları ..... 36
3.5.1. Rastgele orman (random forest) ..... 36
3.5.2. Gradyan artırılmış ağaçlar (gradient boosting trees) ..... 37
3.5.3. Yapay sinir ağları (artificial neural network) ..... 38
3.5.4. Karar ağacı (decision tree) ..... 39

**4. BÖLÜM**

**ARAŞTIRMA BULGULARI**

**SONUÇ VE ÖNERİLER ..... 47**

**KAYNAKÇA ..... 49**

# TABLOLAR DİZİNİ

**Tablo** **Sayfa**

**Tablo 3.1.** Veri Seti Örnek Tablo ..... **25**

**Tablo 3.2.** Örnek Müşteri Tablosu ..... **27**

**Tablo 3.3.** Örnek İşlem Yeri Tablosu ..... **27**

**Tablo 3.4.** Terminallerle ilişkilendirilmesi Biten Müşteri Profil Tablosu ..... **29**

**Tablo 3.5.** Bir Müşteri 'nin 10 Günlük İşlemi ..... **29**

**Tablo 3.6.** Üretilen Müşteri Profili 'nin Kontrolü ..... **30**

**Tablo 3.7.** Küçük Bir Veri Seti Oluşturma ..... **31**

**Tablo 3.8.** Üretilen Veri Setinin İlk ve Son 5 Verisi ..... **32**

**Tablo 3.9.** Veri Setinin İlk 5 Verisi ..... **34**

**Tablo 3.10.** Data Frame 'in İlk 5 Verisi ..... **35**

**Tablo 4.1.** Algoritmaların Test Sonuçları ..... **41**

**Tablo 4.2.** Gradyan Artırma (GBM) Algoritması Karışıklık Matrisi ..... **43**

**Tablo 4.3.** CatBoost Algoritması Karışıklık Matrisi ..... **44**

**Tablo 4.4.** Sınıflandırma ve Regresyon Ağacı (CART) Algoritması Karışıklık Matrisi ..... **45**

## ŞEKİLLER DİZİNİ

**Şekil** **Sayfa**

**Şekil 1.1.** Bilgi Güvenlik Üçlüsü CIA (Singh ve diğerleri, 2014) **4**

**Şekil 1.2.** IBM İlk Antivirüs (Kaspersky) **6**

**Şekil 1.3.** DOS Saldırısı (Keycdn, 2018) **7**

**Şekil 1.4.** Siber Saldırı Finansal Hasar (Yan ve diğerleri, 2021). **8**

**Şekil 1.5.** Siber Saldırı Tespit Sistem Araçları Firewall Örneği (Ayvaz, 2022) **8**

**Şekil 1.6.** Siber Ölüm Zinciri (Netrusion) **9**

**Şekil 1.7.** Zero Day Tanımları (Technopat, 2021) **10**

**Şekil 1.8.** Virüs Bulaşmış svchost.exe (Technopat, 2022) **11**

**Şekil 1.9.** Worm için Alınabilecek Önlem (Thetartan, 2022) **12**

**Şekil 1.10.** Truva Atı Örnek Banka İşlemi (Security Affairs, 2022) **12**

**Şekil 1.11.** RootKit (WM Aracı, 2022) **13**

**Şekil 1.12.** Spyware (WM Aracı, 2022) **14**

**Şekil 1.13.** RAT (Remote Access Trojan, 2022) **15**

**Şekil 1.14.** Adware Sürekli Açılan Reklamlar (Abbott ve diğerleri, 2015) **16**

**Şekil 1.15.** Oltalama (Jampen ve diğerleri, 2020) **16**

**Şekil 2.1.** Temel Öğrenme Öğeleri (Keleş ve Ocak, 2007) **19**

**Şekil 2.2.** Makine Öğrenmesi (Chapman ve diğerleri, 2000, s. 16) **21**

**Şekil 2.3.** Makine Öğrenmesi Türleri (Shewan, 2017) **22**

**Şekil 2.4.** Denetimli Öğrenme (Shewan, 2017) **22**

**Şekil 2.5.** Denetimsiz Öğrenme (Shewan, 2017) **23**

**Şekil 2.6.** Denetimli- Denetimsiz Karşılaştırma (Shewan, 2017) **24**

**Şekil 2.7.** Makine Öğrenmesi Özet (Akdağlı, 2021) **24**

**Şekil 3.1.** Simülasyon Planlama Adımları **26**

**Şekil 3.2.** Müşteri- Terminal İlişkilendirme **28**

**Şekil 3.3.** Üretilen Toplam Veri Sayısı ..... **31**

**Şekil 3.4.** İşlem Miktarlarının Dağılım Grafiği ..... **32**

**Şekil 3.5.** İşlem Sürelerinin Dağılım Grafiği ..... **33**

**Şekil 3.6.** Dolandırıcılık İşlemleri Ekleme ..... **34**

**Şekil 3.7.** Dolandırıcılık İşlemlerinin Yüzdesi ..... **34**

**Şekil 3.8.** Dolandırıcılık İşlemlerinin Sayısı ..... **34**

**Şekil 3.9.** Artırma ve Torbalama Yöntemleri (Breiman, 2001) ..... **36**

**Şekil 3.10.** Tahmin Üreten Gradyan Artırma Fonksiyonu ..... **37**

**Şekil 3.11.** Tahminler ve Hedef Arasındaki Fark ($h_1$) ..... **37**

**Şekil 3.12.** Oluşan Yeni Ağaç ..... **37**

**Şekil 3.13.** Farkın Yeniden Hesaplanması ..... **38**

**Şekil 3.14.** Gradyan Artırma Yöntemi ..... **38**

**Şekil 3.15.** Basit Yapay Sinir Ağ Modeli ..... **39**

**Şekil 3.16.** Gerekli Bilgi Miktarı ..... **39**

**Şekil 4.1.** Doğruluk Oranı Hesaplaması ..... **42**

**Şekil 4.2.** F1 Score Hesaplaması ..... **42**

**Şekil 4.3.** Seçilen Algoritmaların Doğruluk Oranları ..... **42**

**Şekil 4.4.** Kesinlik Hesaplaması ..... **43**

**Şekil 4.5.** Duyarlılık Hesaplaması ..... **43**

**Şekil 4.6.** Gradyan Artırma Algoritması Tahmin Sayıları ..... **44**

**Şekil 4.7.** CatBoost Algoritması Tahmin Sayıları ..... **45**

**Şekil 4.8.** Sınıflandırma ve Regresyon Ağacı (CART) Algoritması Tahmin Sayıları ..... **46**

# GİRİŞ

Günümüzde, bilgisayar, tablet ve akıllı telefonlar gibi bilişim sistemlerinin kullanımının artmasıyla, veri güvenliği, kişisel verilerin gizliliği, yapay zekâ ve siber güvenlik gibi disiplin konuları da artmıştır. Bilişim sistemlerin veri güvenliğini üçüncü kişilerden korunmasını sağlamak için devletler ve şirketler siber güvenliği sağlamak amacıyla yüksek oranlarda bütçeler ayırmak zorunda kalmıştır. Her ne kadar tüm olanaklar kullanılarak önlemler alınsa da yine sızma ve ihlaller olabilmektedir. Finansal ve kritik altyapıya sahip kurum ve kuruluşlara yapılan siber saldırılar gün geçtikçe artmaya devam etmektedir. Kullanılan herhangi bir bilişim sistemi ile internete erişim arttıkça siber saldırıya maruz kalma ihtimali de artmaktadır. Bu durum saldırılan sistemde maddi ve manevi zararlar ortaya çıkarmaktadır. Birçok kişisel verinin internet ortamında bulunması bazı kolaylıklar sağlasa da beraberinde getirdiği sorunların başında kişisel verilerin güvenliği gelmektedir.

Günümüzde insanlar birtakım sorunları kendileri çözmek yerine makineleri eğiterek sonuç üretmeye başlamışlardır. Örnek vermek gerekirse bilişim sistemi üzerinden internet aracılığıyla gelişmiş bir e ticaret sitesinden alışveriş yaptığınızda girdiğiniz verilerden, gezindiğiniz ürünlere kadar bilgi toplayarak alabileceğiniz öneriler göstermektedir. Bilişim sisteminin güvenliğini sağlamak amacıyla önceden yapılan siber saldırıların yöntemlerinden çıkarılan birçok özellik ile elde edilen verileri, yapay zekâ algoritmaları kullanılarak insan gücünden daha etkili bir şekilde siber saldırılara karşı savunma daha etkili olmuştur.

Zeka kelimesi sözlükte, düşünerek ve akıl yürüterek, bir gerçeği anlama ve yargılayarak sonuç çıkarma işlemi olarak yer almaktadır (Topçuoğlu, 2001, s. 39). Yapay zeka kavramının ortak bir karara varılmış bir tanımı yoktur. Bu sebepten dolayı yapay zeka kavramının tanımı, farklı yaklaşımlara karşı farklı anlamlarda olmaktadır. Yapay Zeka Ansiklopedisi 'nde "Bilgisayar biliminin, akıl yürüterek problem çözebilme gibi yeteneklere sahip olacak bilişim sistemleri tasarlayan bölümüdür" şeklinde tanımlanmıştır (Barr ve Feigenbaum, 1981, s. 3).

Yapay zeka, insana benzer şekilde düşünerek bilgisayara girilen işlemleri akıl yürüterek sonuç çıkarma işlemi olarak tanımlanır. Sonuç olarak kodlanmış bir bilgisayarın düşünmeye çalışmasıdır. Daha da detay verecek olursak, insan beynine özgü zeka, öğrenme, düşünme ve bu girişlere göre karar verme özelliklerine sahip bilişim sistemleridir. 1956 yılında, on bilim adamı tarafından, ABD 'nin Dartmouth şehrinde vermis olduğu bir konferansta ilk defa Yapay Zeka kelimesini tanımlandı. Yapay Zeka fikrini ortaya atan bu bilim adamları, zekanın programlarla bütünleştiği kodların geliştirilmesini önerdiler. Bu on bilim adamından birisi olan M. Minksy, insanlar tarafından tasarlanıp geliştirilen ve zeki davranışlarda bulunan makinalar için Yapay Zeka tanımlamasını 1995 yılında yapmıştır (Kocamaz, 2012).

Makine öğrenmesinde algoritmalar ile eldeki verileri analiz edilerek eğitim yapılır. Bilgisayarın kendisine kazandırdığı bu yetenek sayesinde makineye sorulan işlemlerin sonucu için bir

tahminde bulunulur. Literatürde 26 adet makine öğrenme algoritması mevcuttur. Bu algoritmalar, eldeki verilere göre başarı sonucu değişmektedir. Bu sebeple belirli bir yöntem ile tüm veriler ile işlem yapmak doğru değildir (Kocamaz, 2012)

Makine Öğrenmesi, dışarıdan herhangi bir insan müdahalesine maruz kalmadan yetkilendirildiği veriler ile bilgi toplayarak kendisini eğitmesi ve sonuca varması için tasarlanmış yapay zekâ dalıdır. İnsanın baş edemeyeceği büyüklükte bir veri seti mevcut olduğunda makine öğrenmesi yöntemlerine başvurulabilir (Aslan).

Siber güvenlik alanında bilgisi olan birisinin herhangi bir kötü niyetli durumu anlaması, bilgisi olmayan kişilere göre daha iyidir. Telefon, bilgisayar, tablet vb. bilişim sistemlerine ağ üzerinden yetkisiz bir girişin tespiti siber güvenlik alanında bilgi sahibi olmayan birçok kullanıcının endişelendiği bir konu olmuştur (Uslu, 2021). Siber saldırıları etkili bir şekilde savunabilmek zor ve kritik bir süreçtir. Saldırıları otonom bir şekilde tespit edebilmek ve siber suçluları tekrar saldırı yapamayacak şekilde etkisiz hale getirmek amacıyla makine öğrenmesi ile eğitilen bir sistem gerekmektedir (Coşar, 2019). Siber saldırılar için geliştirilmiş otonom savunma sistemleri yapay zeka algoritmalarına entegre edilerek siber saldırılara karşı daha etkin savunma durumu üzerine çalışılmıştır. Kalıplaşmış siber savunma metotları incelenerek, siber savunma metotlarını geliştirmek için yapay zekâ kullanılan farklı teknikler araştırılmıştır (Bilgisayar Sistemleri).

Siber saldırı yöntemlerinin içeriğindeki zararlı uygulamalar sürekli değişiklik gösterdiğinden, siber saldırılarla mücadele etmek için ulusal tedbirlerin önemi daha da artmıştır. Siber suç örgüteri, internet ortamındaki dolandırıcılıklarını yapabilmek için siber savaşçıları bünyelerinde çalıştırmak amacıyla işe almaktadır. Siber örgütlerle savaşabilmek için kendini geliştirmiş siber savaşçılar ve ulusal olarak savunma konusunda ilerlemek zorunlu olmuştur (Andress ve Winterfeld, 2013). Artan siber saldırılar, kurumların da maddi zarar ve veri kayıplarını artırmaktadır. Yapılan test ve anketlere göre, katılımcıların yaklaşık %70 'i en az bir tane de olsa saldırıya uğradıklarını bildirmişlerdir (Jang-Jaccard ve Nepal, 2014).

Siber güvenlik, son zamanlarda artan siber saldırılardan dolayı, üzerinde geniş çalışmalar yürütülülen bir konu olmuştur. Parasal anlamda güçlü olan veya kazanç arayan kişiler siyah şapkalı hackerlar aracılığıyla, kişi ve kurumlar tehdit altına alabilir. Genellikle kurum çalışanları kullanılarak sunucular ve sistemlere yasadışı giriş yapılmasıyla bu tehdit oluşmaktadır (Bıçakçı ve diğerleri; Ünver, 2016)

Günümüzde kredi kartı ve banka kartları ile yapılan online veya herhangi bir mağaza, market, petrol vb. yerlerden yapılan alışveriş çok önemli boyutlara ulaşmaktadır. Bu kadar maddi işlemin yapılabildiği bir ortamda haliyle suçluların da ilgisi büyük olacaktır.

İnternet kullanımı arttıkça kişilerin ve kurumların güvenliklerini artırmaları gerekir. Kimlik doğrulama, güvenlik denetimi ve yama yönetimi gibi tedbirler alınıp kritik altyapılarda kullanılan yazılım ve donanımların güçlü ve zayıf yönleri tespit edilmelidir. Sisteme gelebilecek

saldırılardan korunma ve uygulamalarını iyileştirmeyi amaçlayan analiz çalışmalarını artırmalılardır (Bayindir ve diğerleri, 2016). Siber güvenliği sağlamada yeni tedbirler almanın yansıra, geniş bir ağ operasyonu yapılmalı ve bilgi güvenliği bilgisine sahip profesyonel siber güvenlik uzmanlarına görev verilmelidir. Siber güvenlik uzmanları sayesinde insan ile makine etkileşimi üst düzeye çıkarılarak, siber saldırılara müdahalede daha etkin bir hal alacaktır (Abbott ve diğerleri, 2015).

Kredi kartlarının çalınması veya kopyalanması, kredi kartı sahibinin fark etmesine kadar geçen sürede büyük tehlike arz etmektedir. Bankalar ve kolluk kuvvetleri ne kadar bunun önüne geçmek için ellerinden gelen her şeyi yapsalar da kredi kartı dolandırıcılık vakaları devam etmektedir.

Bu tez, Adli Bilişim ve Siber Güvenlik Uzmanı olan birisinin bilirkişilik yapabileceği kredi kartı dolandırıcılık tespitini, Makine Öğrenmesi yöntemleri ile kredi kartının alışveriş işlemi yapılan konumu, alışveriş tutarı gibi verileri analiz ederek dolandırıcılık tespiti üzerine geliştirilen yazılım uygulamasını içermektedir.

# 1. BÖLÜM

## SİBER İLE İLGİLİ TEMEL KAVRAMLAR

### 1.1. Siber Güvenlik Nedir?

Siber Güvenlik Kavramı, genel bir ifadeyle bir bilişim sisteminin erişebilirliğinin sağlanması, gizliliği ve bütünlüğünün korunması için herhangi bir saldırının tespit edilip savunulmasıdır (Sağıroğlu ve Alkan, 2018). Bir bilişim sisteminde mevcut olan donanım ve yazılımların ürettiği veya içerdiği verilerin değiştirilmemesi veya bozulmamasını sağlamak amacıyla oluşturulmuş güvenlik sistemidir (Adalı, 2016). Verilerin, yetkisiz olarak erişilememesi, değiştirilememesi ve silinememesini dolayısıyla çeşitli programlar yardımıyla güvenli şekilde depolanmasını garanti etmektedir (Herman, 2022).

Siber saldırılarla ilgili çok sayıda vaka bulunmaktadır. Checkpoint firmasının 2020 yılında yayınladığı güvenlik raporunda, dünyanın uçak üretici devlerinden olan Airbus firmasının bilişim sistemlerine saldırılarak personellerin kişisel bilgilerinin çalınması örnek olarak verilebilir. Mcafee gibi birkaç Anti Virüs firmasına ait sunuculara sızılarak uzun süre uzaktan erişime açılmış ve lisans anahtarları çalınmıştır. Yine bir siber suçlunun Yeni Zelanda 'da 1 milyona yakın kişinin tıbbi bilgilerini çalarak satmakla tehdit etmiştir (Check Point, 2020).

### 1.2. Siber Güvenlikte Temel Prensipler

Siber Güvenlik kavramının ana malzemesi hedef kişi için toplanılmış verilerdir. Bu ifadeyi dikkate aldığımızda, siber alemde güvende kalabilmemiz için verilerimizin; Bütünlük, Gizlilik ve Erişebilirlik (CIA) prensiplerini sağlaması gerekmektedir (Goodrich ve Tamassio, 2010).


**Şekil 1.1.** Bilgi Güvenlik Üçlüsü CIA (Singh ve diğerleri, 2014)

Gizlilik, sunucunun veya herhangi bir bilişim sisteminin yetkisiz erişilmesini ve verilerin açığa çıkarılmasını önlemektir. Sisteme izinsiz girmiş kişi, verileri görebilir, değiştirebilir ve kendi çıkarı doğrultusunda kullanabilir. Sistemden çalınan verilerle yapılan analizler sonucu, bilişim sisteminde kullanılan özel yazılımların kullanımı da mümkündür (Samonas ve Coss, 2014). Sonuç olarak gizlilik, bilişim sistemindeki verilere sadece yetkili kişilerin erişebilmesi, yetkisiz kişilerin verileri çalamamasıdır (Singh ve diğerleri, 2014).

Bütünlük, sunucu veya bilişim sistemine yetkisiz şekilde giriş yapan kişinin, mevcut veriyi değiştirmemesi veya silmemesidir. Sistemdeki verilerin hedef adreslere iletiminde verinin değiştirilmediği tam olarak doğrulanmalıdır. Veri gönderilmeden önce belirli algoritmalardan geçirilerek özet veri oluşturulmalı ve karşıya gönderilen verinin bozulup bozulmadığı karşılaştırılmalıdır (Samonas ve Coss, 2014). Sonuç olarak, bilişim sisteminde depolanan verinin eksiksiz, silinmemiş ve değiştirilmemiş olarak saklanması gerekir. Veriyi herhangi bir bozulma yaşanmadan orjinal şekilde tutmak asıl amaçtır (Singh ve diğerleri, 2014).

Erişilebilirlik, verilerin depolandığı sunucu ve bilişim sistemlerine istenildiği zaman kesintisiz ve eksiksiz şekilde yetkili olarak erişilebilmesi anlamına gelir. Saldırı durumunda, verilere erişebilecek olan yetkili kullanıcı, saldırıyı yapan hacker tarafından engellenebilmektedir (Samonas ve Coss, 2014). Sonuç olarak erişilebilirlik, yetkili kullanıcıların verilere her an ulaşıp kullanabilir durumda olmasını içermektedir (Singh ve diğerleri, 2014).

Bu kavramlar CIA incelenirken, aralarındaki ters ilişki dikkate alınmalıdır. Verinin, gizlilik ve bütünlük kriterlerini artırmak için yapılacak bir hamle, erişebilirlik kriterini olumsuz olarak etkileyebilir ve aynı şekilde erişebilirlik kriterini artırmak için yapılacak bir hamle, gizlilik ve bütünlük kriterlerini olumsuz etkileyebilmektedir. Bu yüzden bu 3 prensip arasında bir denge kurulabilmesi, bahsedilen ters ilişki nedeni ile zordur. Sonuç olarak veri, siber dünyanın temel unsurudur. Siber güvenlik, veri ve veri depolayan sunucu ve bilişim sistemlerinin korunmasıdır. Kişi ve kurumlar, belirli periyotlarla siber güvenlik politikalarını ve kontrollerini Siber Güvenlik Uzmanları ile sürdürmelidirler (Singh ve diğerleri, 2014).

### 1.3. Siber Güvenlik Tarihçesi

1980 'li yıllarda artış göstermeye başlayan kişisel bilgisayar kullanımları ile birlikte veri depolanması da artmıştır. Bu yüzden kişi ve kurumların kişisel bilgisayarlarına siber saldırılar artmıştır. Bu olaylar üzerine dönemin ABD Hükümeti siber saldırıları engelleyebilmek için saldırı yapan hackerlara hapis cezası kararı uygulamıştır. Böylece siber saldırganlara yaptırımlar için ilk adım atılmış oldu. Sonraki yıllarda siber saldırı tehditi tüm dünyada artmış ve bu tehditlere karşı mücadele edebilmek için yeni yöntemler geliştirilmiştir. Hackerlar worm

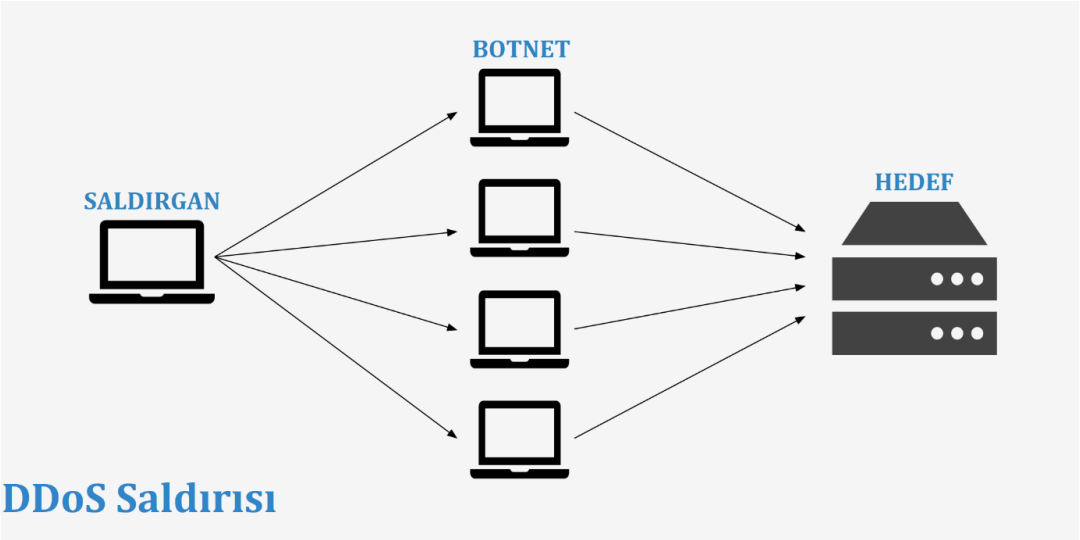
ve trojan gibi malware kullanarak çok sayıda bilişim sistemine zarar vermişlerdir (Coşar, 2019).

1989 yılında, siber saldırıların artmaya devam etmesi ile IBM, ilk anti virus programını geliştirmiştir (Coşar, 2019). Malwareler için anti virus yazılımı ve güvenlik duvarı kullanımına geçilmiştir.

1990 'lı yıllarda ortaya çıkan web tarayıcılarının kullanılmasıyla, hackerların internet aracılığıyla kullanıcılarla etkileşim kurabileceği bir ortam doğmuştur. Bilinçli kullanımın olmadığı internet ortamında, kullanıcıların kişisel verilerini paylaşması da bir çok sıkıntıyı yanında getirmiştir. Microsoft 'un Windows işletim sisteminin güvenlik duvarını barındıran Win 98 sürümünü piyasada yayınlamasıyla birlikte birçok güvenlik açığının kapatılması sağlanmış ve siber suçluların bilişim sistemlerine kolay erişimleri engellenmiştir. Siber suçlular bu dönemlerde web sitelerine DDoS atakları gerçekleştirdiler (Coşar, 2019). Günümüzde çok gelişmiş anti virus programları ve güvenlik tedbirlerinin olmasına rağmen bu saldırılar hala devam etmektedir.


**Şekil 1.2.** IBM İlk Antivirüs (Kaspersky)

Saldırgan, saldırıdan önce ele geçirmiş olduğu BOTNET adı verilen ve kullanıcılarından habersiz arka planda saldırı yapabilecek bilgisayar topluluğuna hedefteki kişinin IP adresini tanımlayıp saldırıyı başlattığında, hedefteki kişi saldırıya göre saniye megabitlerce hatta terabit büyüklüğünde paketler almaktadır. Bu durumda çok fazla yüklenen sistem kilitlenir ve hedefi esas kullanması gereken kullanıcılar sistemden bir yanıt alamazlar ve o an yapması gereken işlemleri yapamazlar.




**Şekil 1.3.** DOS Saldırısı (Keycdn, 2018)

Ülkemizde Siber Güvenlik uygulamaları, 90 'lı yılların sonuyla 2000 'li yılların başında sağlanmıştır. 1997 'de TÜBİTAK 'ın BİLGEM enstitüsü, verinin yalnızca kripto yöntemler ile güvenli olmayacağını, ağ üzerinden bilişim sistemlerine yapılacak saldırıların da ciddi bir konu olduğunu ve Siber ile ilgili uzman bir birimin olması gerektiğini belirtmiş ve bu kararları yürürlüğe koymuştur (Bilgisayar Sistemleri). Bunun sonucunda da ülkemizde, Windows ve Unix işletim sistemlerinde ki E-Posta ve veri tabanı gibi uygulamaların güvenliği için ağ cihazları ve saldırıyı tespit eden sistemler kullanılmıştır. 2001 yılında Oktat Kriter Test Merkezi (OKTEM) projesi Genel Kurmay Bakanlığı desteği ile hayata geçirilmiştir (Bilgisayar Sistemleri). Güvenlik yazılımları geliştirilmiş ve TÜBİTAK bünyesinde siber savunma ekipleri kurulmuştur (Bilgisayar Sistemleri).

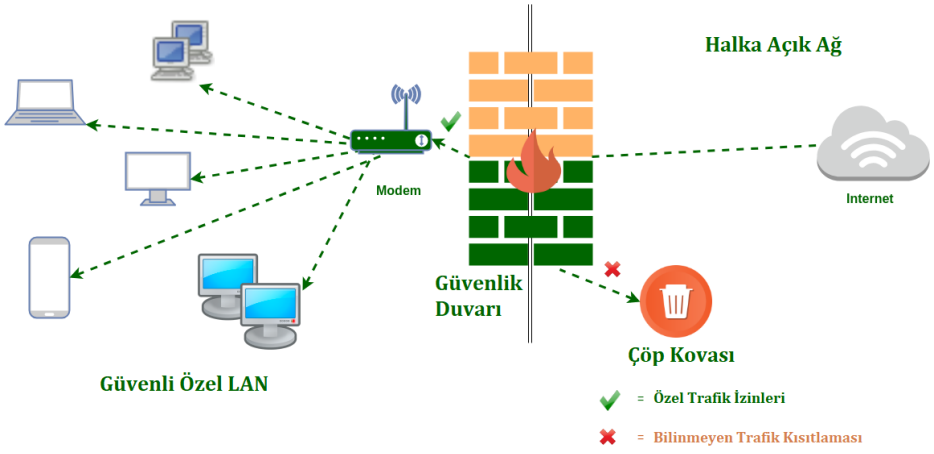
## 1.4. Siber Saldırı Nedir?

İnternet aracılığıyla farklı tür bilişim sistemleri üzerinde kazanç sağlamak veya zarar vermek gibi amaçlar doğrultusunda siber suçluların kişi ve kurumları hedef aldıkları saldırılara Siber Saldırı denir (Singh ve diğerleri, 2014). 2009 yılında ABD 'nin yürüttüğü bir çalışmada siber saldırıların tanımını "Ağ, bilişim sistemi veya veriyi bozmak, değiştirmek ve silmek için bilerek yapılan eylemler" olarak açıklamıştır (Singh ve diğerleri, 2014). Günümüzde gelişmiş ve daha da karmaşık hale gelmiş siber saldırıların artması, siber saldırıların verdiği önemli finansal ve manevi zararlar önlenmesi gereken bir risktir (City of Vancouver, 2016). Siber saldırı, hedef alınmış yerin ağ ve bilişim sistemlerine yapılan planlanmış bir saldırıdır (Alkan, 2012). Bir siber saldırının hedefi, bilişim sistemlerinde depolanan veriyi değiştirmek, silmek veya çıkarı

doğrultusunda kullanmak amacıyla bahsi geçen bilişim sistemlerinin mevcut olduğu ağlardır (Biju ve diğerleri, 2019).

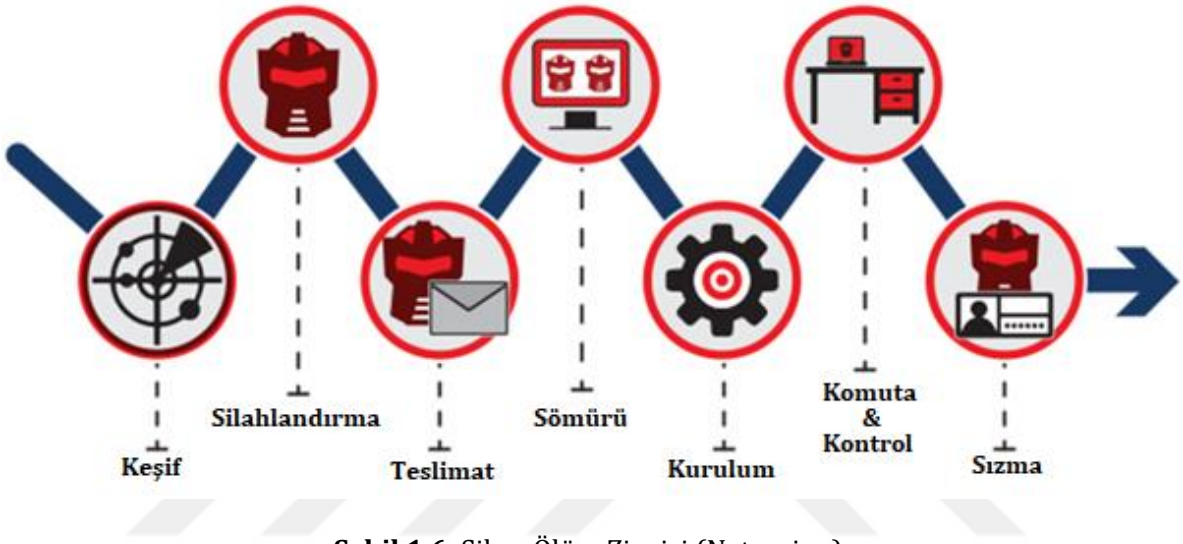

**Şekil 1.4.** Siber Saldırı Finansal Hasar (Yan ve diğerleri, 2021).

Siber suçlu, saldırısına başlamadan önce saldırıdan elde edeceği kazancını belirler. Ardından hedefteki bilişim sisteminin güvenlik açıklarını tespit eder. Bulduğu açıklara yönelik farklı saldırı metotları geliştirir (Sinha ve diğerleri, 2019). Bu saldırılar hedefe ağır hasarlar verebilir. Verilecek hasarı en aza indirmek veya tamamen engellemek için siber saldırıları saldırıdan önce tespit etmek büyük bir önem taşımaktadır. Siber saldırıların tespitinde kullanılan yazılım ve ağ cihazları, bilişim sistemlerini içeren ağdaki anormal hareketliliği tespit edemeyebilir. Siber saldırıların tespit edilme teknikleri genellikle, anti-virus, güvenlik duvarı (firewall), anti-spyware ve anti-malwarelar olduğu ve bu teknolojilerin kaliteli bir siber saldırıyı önleyemeyeceği ve daha kesin bir çözümün geliştirilmesi gerektiği araştırmalar sonucu belirlenmiştir (Pingree ve diğerleri, 2015).




**Şekil 1.5.** Siber Saldırı Tespit Sistem Araçları Firewall Örneği (Ayvaz, 2022)

Güvenlik duvarına SSH servisi için bir engelleme kuralı eklendiğinde, yapılan siber saldırılarla ilgili hiç bir trafiğin geçemeyeceğini söyleseler de, böyle bir durum kesinlik kaliteli bir siber saldırı için geçerli değildir. Büyük bir siber saldırıyı engellemenin bir takım taktikleri vardır. Siber suçluyu bilişim sistemleri için yaptığı keşif sırasında yavaşlatmak ve bilişim sistemine erişimini karmaşık ve zor bir hale getirmek için siber ölüm zinciri adı verilen gelişmiş bir model kullanılmalıdır. Saldırının tespit edilmesiyle birlikte karşı saldırı da siber suçluyu yavaşlatabilir.




**Şekil 1.6.** Siber Ölüm Zinciri (Netrusion)

Siber saldırılar için farklı yöntemler vardır. Yapay zeka son yıllarda saldırı tespiti için kullanılmıştır. Yapay zekanın alt başlıklarından olan derin öğrenme yöntemi, içeriğinde barındırdığı algoritmalar sayesinde mevcut olarak kullanılan siber güvenlik sistemlerinden daha başarılı olmuştur (Nguyen ve diğerleri, 2018). Bilişim sistemi içerisine sızdırılacak bir tehdit, tehditlerin en önemlilerinden birisidir. Daha önceden bilişim sistemine yapılmış siber saldırıların karakter analizlerini yapmak, bu tehditleri önleyebilmenin yollarından birisidir (Probst ve diğerleri, 2010, s. 1-15).

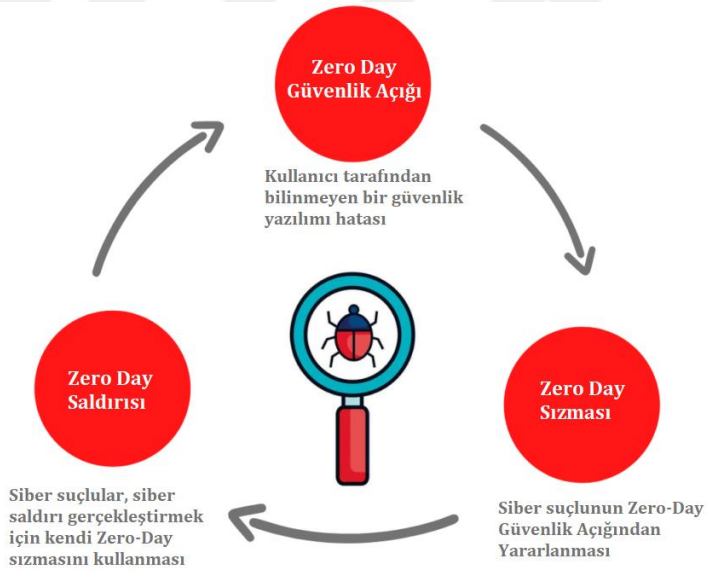
Günümüzde ev ve iş yerlerinde kullanılan kişisel bilgisayarlar kadar cebimizde taşıdığımız akıllı telefonlar da günlük yaşantının vazgeçilmez bir unsuru olmuştur. Kişilerin veya kurumların bilişim sistemlerindeki siber güvenlik açıkları, günümüzün önemli risklerinden birisi olmuştur. Gün geçtikçe yaygınlaşan siber saldırılar, hedef yerde oluşturduğu ekonomik veya itibar kaybı yaşatmasından dolayı önlem alınması gereken önemli bir tehdittir. Kurumlar veya kişiler, siber bir saldırıya maruz kalmadan önce gerekli önlemleri almalı ve güvenlik sistemlerinin analizlerini yapmalılardır.

## 1.5. Siber Saldırı Türleri

Siber dünyanın elbette çok sayıda saldırı içeriği vardır. Bu dünyada en çok kullanılan saldırı türleri; ZeroDay, Malware, Ransomware, Adware, Phising, Network Atakları ve Spam oldu söylenebilir. Bu saldırı türlerini kullanan siber suçluların ortak amaçları, bilişim sistemlerine yetkisiz bir şekilde erişim sağlayarak, sistemdeki verileri ele geçirmek, verileri yok etmek veya bilişim sistemini servis dışı bırakmaktır.

### 1.5.1. ZeroDay

ZeroDay saldırısı, siber suçluların saldıracağı bilişim sisteminin güvenlik açıkları için kullanılacak geniş kapsamlı bir terimdir. Bu terim, geliştiricilerin, kendi yazılımlarında farkettiği bugları yeni farkettiği ve bu hatayı gidermek için "Sıfır Günü" var demektir. Siber suçluların, yazılımın geliştiricilerinin dikkat etmediği bu açıklara düzeltme şansı vermeden yaptığı saldırıya ZeroDay saldırısı denir (Technopat, 2021).




**Şekil 1.7.** Zero Day Tanımları (Technopat, 2021)

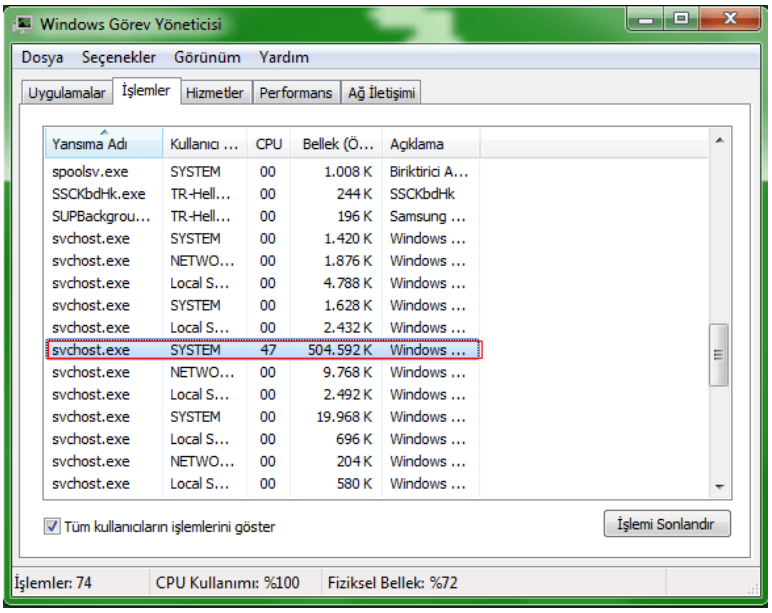
### 1.5.2. Malware

Malware, Malicious Software ifadelerinin kısaltmasıyla oluşturulan, kötü amaçla kullanılan yazılımların genel ifadesidir. Bu tür yazılımlar, kullanıcı adı, parola, kişisel bilgiler gibi verilerinizi çalma, bilişim sisteminizi çökertme, yavaşlatma ve yetkisiz erişime açma amaçlarıyla geliştirilen yazılımlardır. Scumware olarak da bilinmektedir. Bir bilişim sistemine

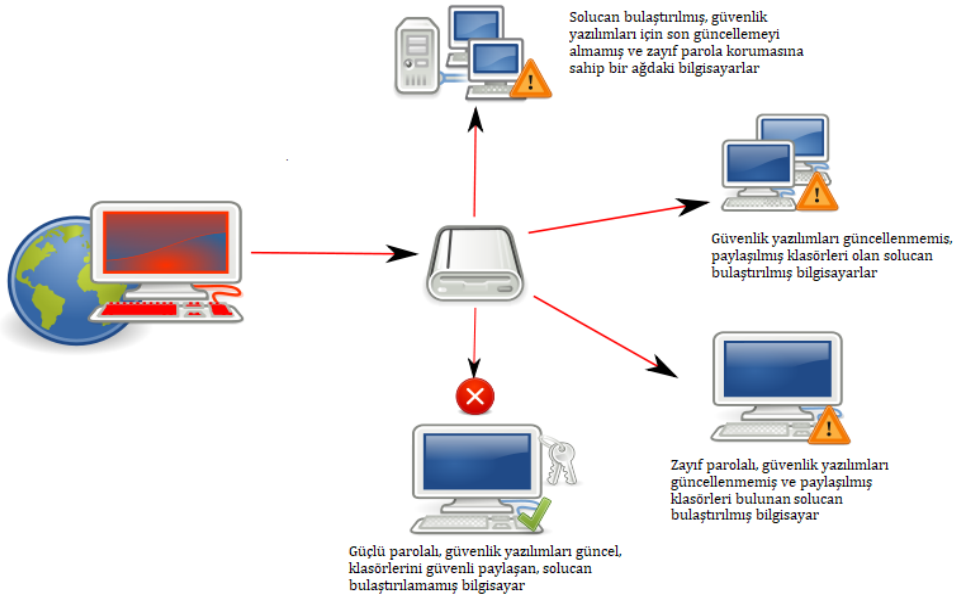
Malware genllikle, herhangi bir web sitesi ziyaretinden, virus içeren yazılım indirilip çalıştırılmasıyla bulaşabilmektedir (Uslu, 2021).

Bilişim sisteminin kullanıcısından habersiz bir biçimde casus ve fidye yazılımları gibi kötü amaca hizmet eden siber saldırılardır. Siber suçlu tarafından geliştirilen kod, normal bir yazılımın kodlarına eklenmektedir. Geliştirilen kodlar aracılığıyla bilişim sisteminin bağlı olduğu ağ, siber suçluya erişim için açılır veya sistemi servis dışı bırakarak verileri geliştiricisine iletmektedir. Malwarelar genellikle kişisel verilerden ziyade kazanç sağlayacağı finansal verileri hedef almaktadır (Biju ve diğerleri, 2019). Malware türlerini inceleyelim.

Virüs, içerisinde işletim sistemi barındıran tüm bilişim sistemlerinde, bir yazılım veya bir kod parçası olarak depolanan kötü amaçlar için çalışan programlara verilen genel bir isimdir. Bir virus, kullanıcıdan bağımsız çalışarak kendisini önbellek ve açık olan herhangi bir belge veya program gibi çalıştırılabilen sistemlere kopyalar (Bilgisayar Sistemleri). Virüsler, genellikle maillerle gelen eklerle, mesaj içerikleriyle veya bilişim sisteminde kurulu olan yazılımlardaki güvenlik açıklarıyla yayılırlar. Bu yüzden, gelen mail ve mesajların kimden geldiği bilinmiyorsa kesinlikle açılmamalı, ekleri indirilmemelidir (Bilgisayar Sistemleri). Virüsler, web siteleri aracılığıyla da yayılabilirler. Virüsler, lisanssız bir şekilde web sitelerinden indirilen yazılımlar veya dosyalar gibi yöntemlerle de bulaşabilirler (Bilgisayar Sistemleri). Virüslerden korunabilmek için yukarıda belirtilen durumlara dikkat edilmeli, bilgisayar ve güvenlik uygulamaları en güncel hali ile kullanılmalıdır (Bilgisayar Sistemleri).



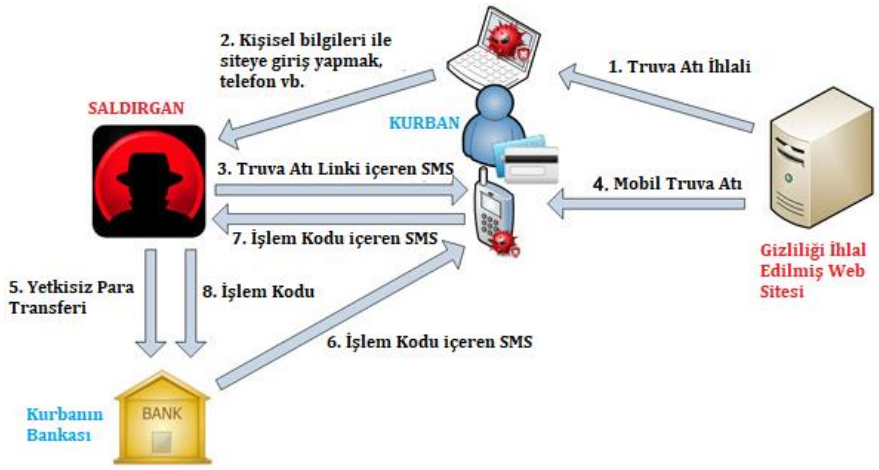

**Şekil 1.8.** Virüs Bulaşmış svchost.exe (Technopat, 2022)




**Şekil 1.9.** Worm için Alınabilecek Önlem (Thetartan, 2022)

Bilişim sistemlerinin ağ sistemlerindeki açıklardan yayılan solucanlar, ağdaki diğer bilişim cihazlarına on binlerce kopyasını göndermektedir. Solucanların birçoğunun amacı bilişim cihazlarının donanım kaynaklarını kullanıp performans azaltma olsa da artık çoğu solucanın amacı bulaştığı bilişim sistemlerindeki verileri silmek ve ele geçirmektir. E-posta eklerindeki programlar, korsan CD-DVD gibi farklı ortamlar, en sevdikleri bulaşma yöntemleridir (Vikipedi, 2022).

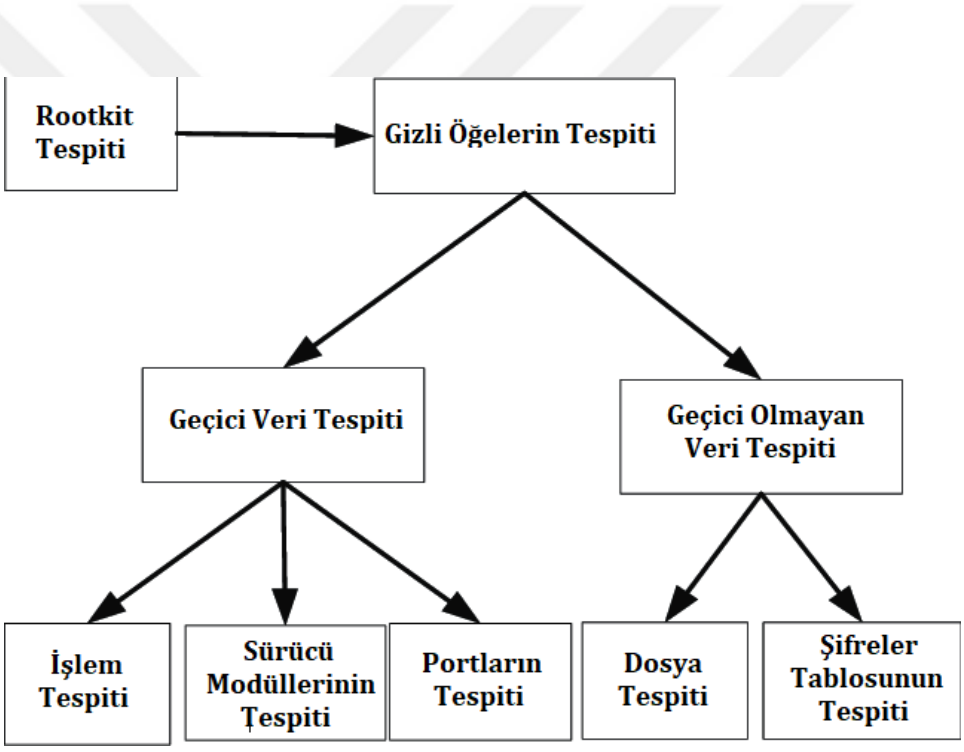
Truva Atı (Trojan), eski Yunan mitolojisindeki hikayesine benzer şekilde, yanlış yönlendirme gibi metotları kullanarak kendisini ve kötü amacını gizlemektedir. Hedef aldığı bilişim sistemine sızdıktan sonra, bilişim sisteminin kullanıcısı tarafından çalıştırılabilmesi için farklı yöntemlerle yerleşirler (Goodrich ve Tamassio, 2010).




**Şekil 1.10.** Truva Atı Örnek Banka İşlemi (Security Affairs, 2022)

En çok rastlanan kötü amaçlı yazılım türü Truva atlarıdır. Backdoor oluşturmak, bulaştığı bilişim cihazını kontrol etmek, bilişim sisteminde depolanan verileri ele geçirmek gibi birçok kötü amaca hizmet etmektedir (Goodrich ve Tamassio, 2010). Bu malware için Truva terimi, 1974 yılında ABD ordusu raporlarında kullanılmıştır (Goodrich ve Tamassio, 2010). Truva atlarının birçok türü vardır. Backdoorlar, en tehlikeli ve en çok kullanılan türüdür (City of Vancouver, 2016). Bulaşmış olduğu bilişim sistemini kimsenin ruhu duymadan siber suçlunun kontrolüne açmaktadır (City of Vancouver, 2016). Diğer bir tür olan PSW Trojan, bilişim sistemindeki şifreleri çalmak için geliştirilmiştir (City of Vancouver, 2016). Buna benzer amaçla geliştirilen daha birçok Truva türü vardır.

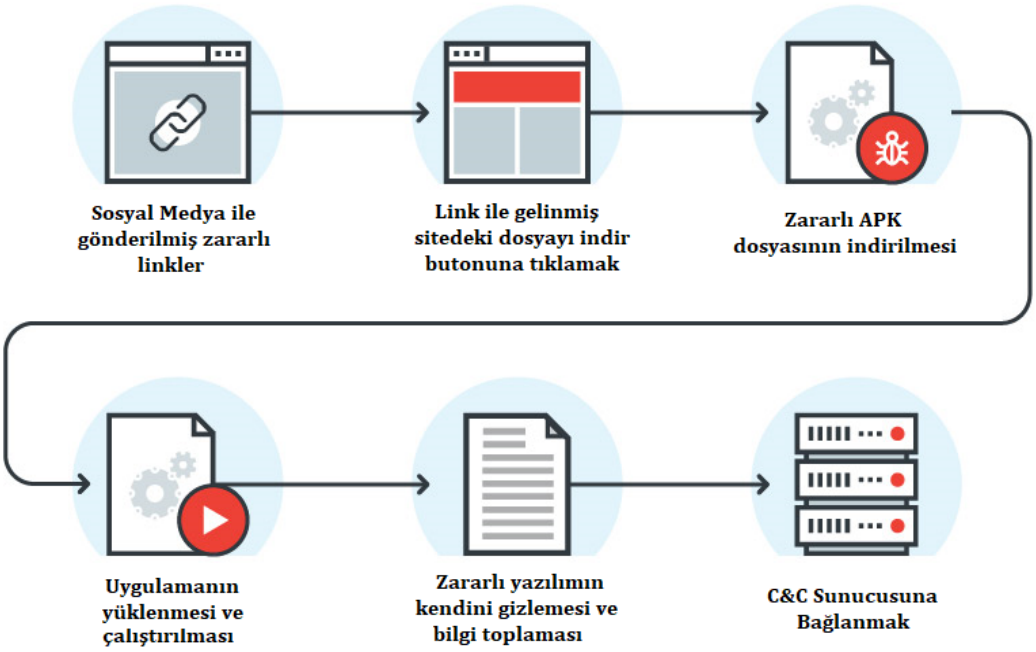
Diğer bir malware türü olan Rootkit, bilişim cihazlarının tam yetki ile yazılımı geliştiren siber suçluya uzaktan kontrol edebilmesini sağlayan, işletim sisteminin çekirdeğine sızan kötü amaçlı yazılımlardır (VM Aracı, 2022).




**Şekil 1.11.** RootKit (WM Aracı, 2022)

Rootkit kesinlikle normal bir virüs yazılımıyla karıştırılmamalıdır. Virüslerin amacı, bilişim sistemlerinize sızdıktan sonra kendini çoğaltmak ve geliştirilme amacı doğrultusunda görevini gerçekleştirmektir. Rootkit 'in ise amacı kendini çoğaltmak değil, bulaştığı andan itibaren geliştiricisine tam yetki ile uzaktan erişim açmaktır. Bahsedilen bu detayları inceleyecek olursak, Rootkitler diğer kötü amaçlı yazılımlara kıyasla daha tehlikelidir. İşletim sisteminin çekirdeklerine kendilerini gizledikleri için Antivirüs yazılımlarına yakalanmazlar [39].

Spyware, bilişim sistemlerindeki verileri yetkisiz olarak geliştiricisi olan siber suçluya iletmektedir. Bu malware türü genellikle reklam amaçlıdır. Güvenilmeyen bir siteden indirilen yazılımın içerisine kendini gizleyerek, bilişim sistemine sızarlar [40].




**Şekil 1.12.** Spyware (WM Aracı, 2022)

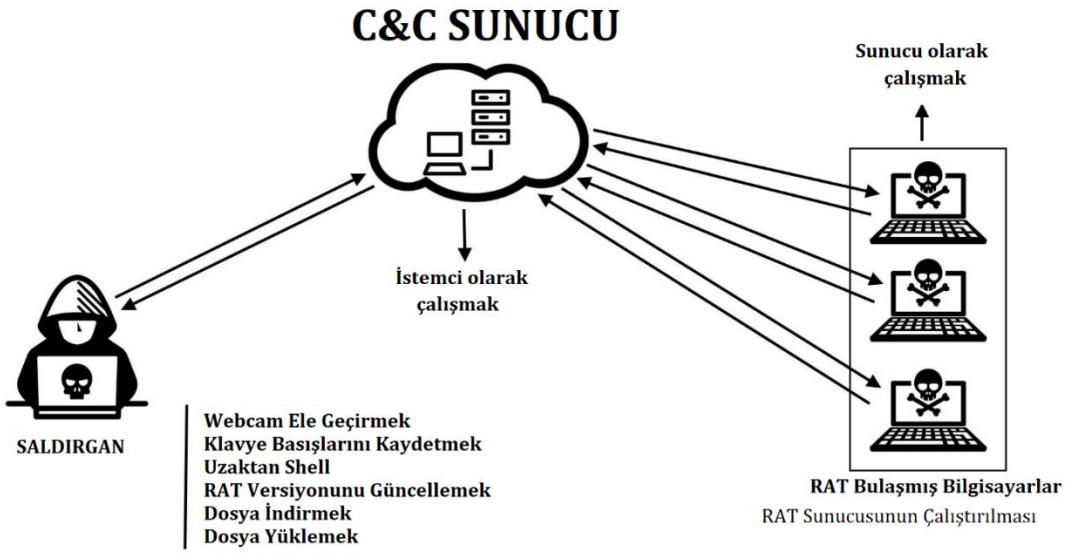
Diğer malware türlerinin aksine Spyware bir yazılım, casus yazılımlar kategorisindedir. Amacı sızdığı bilişim sistemine zarar vermek değildir. Asıl amacı bilişim sisteminden çalacağı verilerdir. Spyware, genellikle ücretsiz olarak sunulan yazılımların içerisine yazılır ve kurulduğu bilişim sistemine kolayca entegre olur. Kullanıcıdan habersiz verileri aktarmaya başlayabilir [40].

Malware türlerinden birisi olan Karma Tehdit (Blended Threat) adından da anlaşılacağı üzere Trojan, Virus ve Worm yazılımlarının karmasıyla oluşturulmuş saldırıdır [42].

Başka bir malware türü olan RAT (Remote Access Trojan), arka planda yönetici olarak çalışan, TCP/UDP portlarını kullanarak gelişiricisine sistemdeki verileri ileten, Antivirüs yazılımları ile tespit edilemeyen tehlikeli bir malware 'dir. Kısaca özetleyecek olursak siz bilgisayarınızı kullanırken neler yapabiliyorsanız, RAT'ı geliştiren siber suçluda aynı işlemleri yapabilmektedir (Sağıroğlu ve Alkan, 2018).

Bir RAT, bilgisayarınızı bir zombi bilgisayar olarak RAT geliştiricisinin kurmuş olduğu Zombi Bilgisayar topluluğundan oluşan BOTNET'e dahil edecektir. Yani bilgisayarınızı kendi saldırılarını yapmak için kukla olarak kullanacaktır. Siber suçlu, bir web sayfasına entegre etmiş olduğu script kodlarıyla arka planda gizlice gerekli dosyaları yükleyip açacaktır (Sağıroğlu ve Alkan, 2018).

Diğer bir malware türümüz olan Adware, genellikle başka bir program yüklerken kullanıcıdan habersiz veya küçük bir kutucukla işaretlenerek yüklenen yazılımlardır. Adwarelar, bilgisayar kullanıcısının internette gezindiği sayfaları, alışverişlerini ve aktivitelerini analiz ederek tıklanması daha yüksek ihtimalli reklamlar sunmaktadır. Bir de kötü amaçla geliştirilen adwarelar vardır. Bu tarz malwarelar, internet tarayıcılarına kendini yerleştirerek, ana sayfa değişikliği ve can sıkıcı pop-up reklamları açmak gibi işlemler yapmaktadır (Bayindir ve diğerleri, 2016).




**Şekil 1.13.** RAT (Remote Access Trojan, 2022)

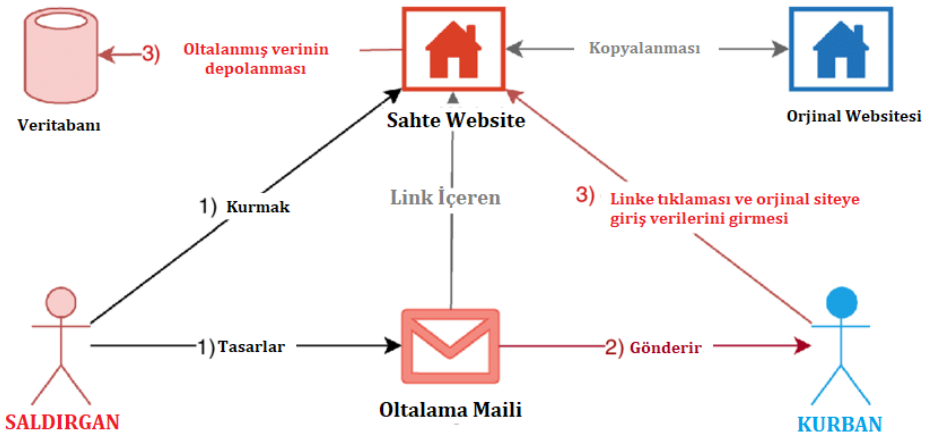
Tüm Adwarelar, veri topladığı verileri analiz edip reklam stratejisi kurmak için geliştirilmez. Bazı adwarelar, ekranda kullanıcıyı bunaltacak çok fazla reklam penceresi açar ve bu reklamları zorla tıklattırabilmektedir. Tüm Adware yazılımlar bilgi toplamak veya daha etkili reklam stratejileri üretmek için geliştirilmez. Bazı adwarelar ekranda çok fazla can sıkıcı reklam pencereleri açarak kullanıcıya zorunlu şekilde açtırabilmektedir (Bayindir ve diğerleri, 2016).


**Şekil 1.14.** Adware Sürekli Açılan Reklamlar (Abbott ve diğerleri, 2015)

Eğer kullandığınız tarayıcı, kendi kendine ana sayfasını değiştiriyorsa, bilgisayarınıza Adware yazılımı bulaşmıştır. Adwarelar kendilerini yalnızca tarayıcılarınız değil, tarayıcıdan bağımsız şekilde ekranınızda uyarı penceresi olarak sizi hedef alan ilginizi çekecek bir soru veya görselle açılabilir (Bayindir ve diğerleri, 2016).

### 1.5.3. Phishing (oltalama)

Hedef alınan kişinin kimlik bilgilerini çalmak için kullanılan en etkili siber saldırıdır. Siber suçlular, mail ekleri ve domaini orjinal adresine benzeyen sahte banka ve başka web siteleri gibi yöntemleri kullanarak saldırmaktadır [44]. Bu saldırı türü, siber suç madurları, balık tutulmasına benzetildikleri için oltalama adını almıştır. Lisanslı ve popüler web sitelerinin domain adresleri birkaç harf değiştirilerek orjinalinin kopyası oluşturulmaktadır. Bu şekilde milonlarca kullanıcı zarara uğramıştır. Bu tarz sahte domainleri çok dikkatli kullanıcılar farkedebilir. Sitenin URL adresine dikkatlice bakacak olan internet kullanıcısı, orjinal domaini olmadığını ve bazı harflerinde yapılan değişikliklerin olduğunu farkedecek ve siber saldırıdan kurtulacaktır (Das ve diğerleri, 2019)




**Şekil 1.15.** Oltalama (Jampen ve diğerleri, 2020)

### 1.5.4. Servis dışı bırakma saldırıları

Bir sunucunun vermiş olduğu hizmeti engellemek en çok hasar veren saldırılardan birisidir. Siber suçlu, saldırdığı sunucuyu geçici veya kalıcı olarak bağlantılarını keserek, bilişim sistemini veya bağlı olduğu ağı kullanılmayacak hale getirmektedir. Bu saldırı 2 farklı yöntemle kullanılabilmektedir. İlk yöntem, siber suçlu bilişim sistemine zararlı paketler göndererek, hedefindeki bilişim sistemi üzerinde çalışan protokolleri ve uygulamaların çalışmasını engellemek, ikinci yöntem ise hedef aldığı bilişim sisteminin donanım kaynaklarını ele geçirerek, sistem kullanıcısının kullanmasını engellemektedir (Technopat, 2021). Hükümetlerin, bankaların ve ticari kuruluşların sunucuları genel olarak bu saldırıya maruz kalan bilişim sistemleridir. Amacı verileri çalmak veya yok etmek olmasa da, temizlenmesine kadar sistemde ciddi mali kayıplara yol açmaktadır.

Bu saldırının en çok kullanılan yöntemlerinden birkaç tanesini açıklayacak olursak; Sistemin ağ adresi, siber suçlu tarafından yoğun şekilde trafik almasıyla oluşan bir saldırı türü olan Buffer Overflow Attack, bilişim sisteminin ağ veya uygulamalarında mevcut olan açıklar üzerinden sistemi istismar etmektedir. Diğer bir saldırı yöntemi olan ICMP Flood Attack, siber suçlu tarafından sisteme gönderilen sahte paketler, bilişim sistemine yoğun şekilde ping göndererek sunucunun ağ üzerinde servis dışı kalmasını sağlamaktadır. Smurf saldırısı veya ölüm pingi olarak da adlandırılmaktadır. Başka bir yöntem olan SYN Flood Attack ise saldırıya uğrayan sunucu, saldırının başlamasından itibaren tüm donanım kaynaklarının SYN paketlerinden dolayı kullanılarak, sunucu istemcilerden gelen veri alışveriş isteklerine cevap veremeyecektir (Sinha ve diğerleri, 2019). Başka bir tür olan DOS saldırısı, bilişim sistemine yine ağ üzerinden yoğun şekilde trafik gönderilmesidir. Bu saldırı türünün en büyük farkı, hedefe birden fazla kaynaktan saldırılmasıdır (Sinha ve diğerleri, 2019). DDOS 'da ise birden fazla hedef olabilir ve yine bu saldırı zombie bilgisayar topluluğu olan BOTNET üzerinden yapılabilmektedir (Technopat, 2021).

### 1.5.5. Spam

Spamlar, genellikle propaganda ve tanıtım amaçlı mesajların elektronik ortamda mail yolu ile gönderilmesidir (Beydogan ve Canbay, 2008). İstenmeyen elektronik postalar denilen bu saldırı türü, büyük bütçeli mail sunucularına sahip kurumların aldığı önlemlere ve yaptığı çalışmalara ragmen hala tam anlamıyla önlenebilmiş değildir.

## 1.6. Siber Suçlu

Elektronik cihaz, zararlı yazılımlar veya kullanıcılara ait sistemlerin yazılım açıklarından faydalanarak, internet ortamında kişisel veriler veya kredi kartı gibi bilgilerinin çalınması, kullanıcıların servislerini devre dışı bırakma veya sistemden herhangi bir veriyi kendi çıkarı

doğrultusunda değiştirme gibi amaçlarla yapılan eylemleri gerçekleştiren kişilere Siber Suçlu denir (Uslu, 2021).

Siber suçlulardan birisi de Hackerlardır. Teknik olarak donanımlı bu kişiler, bilgisayar ve ağlara yetkisi olmadığı halde girip, verilere erişen ve istediği gibi oynama yapabilirler (Coşar, 2019). Beyaz şapkalı, siyah şapkalı ve gri şapkalı olarak üçe ayrılırlar. Beyaz şapkalı hackerlar, her türlü yazılımı, web sitesini veya bilgisayarı, açıklardan yararlanarak kırıp, sistem sahibi ile iletişim kurarak bulduğu bu açıkların kapatılmasını ve diğer hackerlardan korunmasını sağlayan kişilerdir (Coşar, 2019). Yasal ve iyi niyetli amaçla saldırı yapan hackerlardır. Siyah şapkalı hackerlar, sistemlere yetkisiz erişimlerinden sonra, sistemleri devre dışı bırakan, kullanılmayacak hale getiren, verileri çalan, değiştiren kişilerdir (Coşar, 2019). Yasal olmayan ve kötü amaçla saldırı yapan hackerlardır. Gri şapkalı hacker ise siyah şapkalı hacker ve beyaz şapkalı hackerın ortasındadır. Sistemdeki açıkları siyah şapkalının yöntemi ile tespit eder ve sistem sahibine beyaz şapkalı gibi rapor eder. Bulduğu açığı düzeltmek için isterse para teklif edebilir. İyi niyetli ancak yasal olmayan hack yapan kişilerdir.

Diğer siber korsan veya siber suçlulardan birisi Hacktivistlerdir. Kendi düşüncelerine göre doğru veya yanlış olan politik veya toplumsal konuları, sorunları söylemek amacıyla belirli web sitelerine saldırarak mesajlarını siteye yerleştiren kişilerdir (Coşar, 2019). Başka bir korsanımız Script Kiddieler, Hackerlara özenen kişilerdir. Tam bir hacker olmasalarda genellikle kişisel e posta ve kişisel şifreleri çalan kişilerdir (Coşar, 2019). Başka bir korsan türümüz Lamerler ise ne yaptığını tam anlamıyla farkında olmadan ve yeterince bilgisi olmadan Hack yapan ve Script Kiddie ‘e benzeyen kişilerdir (Coşar, 2019). Diğer suçlumuz Yazılım Korsanı ise bilgisayar programlarının lisans anahtarlarını kırarak, bu programları izinsiz çoğaltan, dağıtan ve kazanç sağlayan kişilerdir (Coşar, 2019).

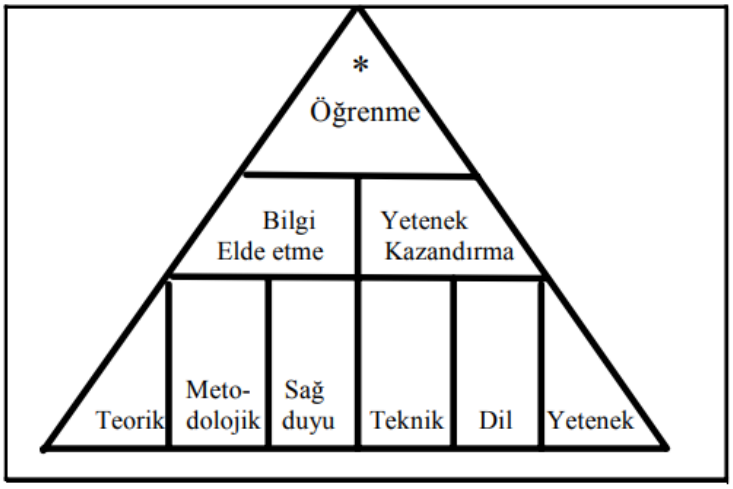
# 2. BÖLÜM

# YAPAY ZEKÂ

## 2.1. Yapay Zekâ Nedir?

Zeka kelimesi, düşünerek ve akıl yürüterek, bir gerçeği anlama ve yargılayarak sonuç çıkarma işlemi olarak sözlükte yer almaktadır (Bayık, 2019). Yapay zeka kavramının ortak bir karara varılmış bir tanımı yoktur. Bu sebepten dolayı yapay zeka kavramının tanımı, farklı yaklaşımlara karşı farklı anlamlarda olmaktadır. Yapay zeka, Yapay Zeka Ansiklopedisi 'nde "Bilgisayar biliminin, akıl yürüterek problem çözebilme gibi yeteneklere sahip olacak bilişim sistemleri tasarlayan bölümüdür" (Barr ve Feigenbaum, 1981, s. 3).

Öğrenme sürecini bilgisayar sistemine dönüştürmek için araştırmacılar robot ve yapay zeka yapıları oluşturmak için çalışmalar yürütmüştür. Bu sayede günümüzde önemli bir yere sahip olan yapay zeka öğrenme yöntemlerine ait algoritmaların sayısı artmaya devam etmektedir. Bu öğrenme yeteneğinin en önemli özelliği, belirli işlere programlanıp sadece o işleri yapacak makineler yerine eğitilebilir makineler olmalarıdır. Bir insanın temel öğrenme öğelerinde örneğin bir gitar çalma yeteneğine sahip olmak için gitar akorları ile ilgili bilgilere ve teknik anlamda çaba harcaması gerekmektedir. Ardından kazandığı bu gitar çalma yeteneğini kazanmış oluyor (Keleş ve Ocak, 2007).




**Şekil 0.1.** Temel Öğrenme Öğeleri (Keleş ve Ocak, 2007)

Günümüzde bilgisayar kullanım sayısı arttıkça, bilgisayarların özellikleri ve alma amacını karşılayacak bir kullanıma el verişli olması da önem kazanmıştır. Geçmişte Yapay Zeka tabanlı oyun ve uygulamalar geliştirmeye çalışan az sayıda bilim adamı mevcutken, günümüzde,

birçok ülkenin üniversite araştırma merkezlerinde Yapay Zeka tabanlı teknoloji ve uygulamalar ile uğraşan çok sayıda bilim adamı vardır (Keleş ve Ocak, 2007).

## 2.2. Yapay Zekanın Tarihçesi

Geçmişin modern bilgisayarları kadar yakın bir geçmişe sahip olduğunu, yapay zeka kavramına bakarak söyleyebiliriz. Alan Mathison Turing, "Makineler düşünür mü?" sorusu ile tartışma sunmuştur. 2. Dünya Savaşında haberleşmelerin şifrelenmesi ve şifre çözülmesi için kullanılan kripto analiz kullanılarak üretilen elektromekanik cihazlar ile birlikte yapay zeka kavramları da ortaya çıkmıştır (Sucu ve Ataman, 2020).

Turing, elektromekanik cihazlardan birisi olan Enigma için, kriptolojide decription olarak tanımlanan şifre çözme çalışmalarına İngiltere 'nin Bletchey Parkı 'nda başlamıştır. Turing 'in çalışmalarından oluşan Bombe, Colossus ve Heath Robinson adlı bilgisayar prototipleri ile veri işleme algoritmalarını boole cebri ile geliştirmiştir (Sucu ve Ataman, 2020).

Yapay Zekanın tarihçesi, paragraflar halinde özetlenmiştir (Prim, 2006, s. 81-93).

Bilgisayar Mühendisleri, sisteme yalnızca veri yükleyerek düşünen bir sistem geliştirmek için akıllı bilgisayarlar oluşturmayı beklemişlerdir. Bu yüzden bu bir duraklanma yaşanmış ve 1965-1970 yılları arasında ki bu döneme Karanlık Dönem denmiştir.

Bilgisayar Mühendisleri, sürekli artan bir ivme ile yeni sistemler geliştirerek teknolojinin şuan ki geldiği konumun temelleri atılan bu dönem 1970-1975 arasındaki Rönesans Dönemidir.

Psikoloji ve dil gibi birçok bilim dallarından faydalanan mühendislerin 1975-1980 yılları arasındaki bu dönemi, Ortaklık Dönemi olarak bilinmektedir.

Günlük hayatta kullanılan ihtiyaçlara yönelik yapılan deney ve çalışmaların uygulamalara dönüşmeye başladığı ve halen günümüzde devam eden 1980 sonrası bu dönem Girişimcilik Dönemi olarak adlandırılmıştır.

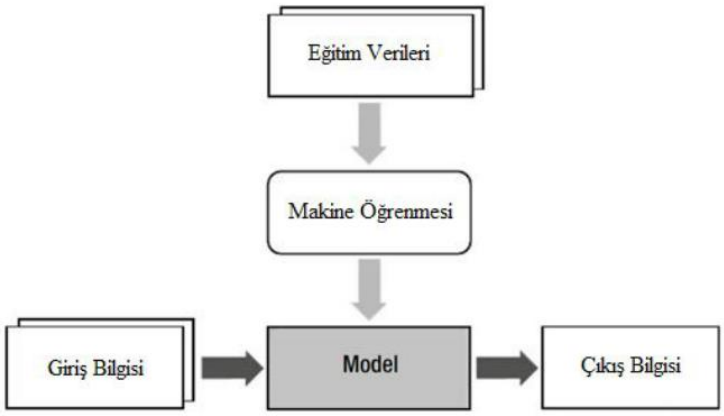
## 2.3. Yapay Zekâ Türleri

Yapay Zekâ; Makine Öğrenmesi, Derin Öğrenme, Doğal Dil İşleme, Bilgisayar Görmesi ve Açıklanabilir AI olarak beşe ayırabiliriz (Aslan). Biz bu çalışmamızda Makine Öğrenmesi yöntemini kullandık.

### 2.3.1. Makine öğrenmesi (machine learning)

Bilişim Sistemlerinin, mevcut verilerden yararlanarak öğrendiği yeteneğini uygulayabilmesi için belirli algoritmaları kullanarak sonuca varmasına makine öğrenmesi denir (Uzun, 2007). Makine öğrenmesi, insanın problem çözme yeteneklerini elindeki verilerle taklit ederek sonuç

çıkarma işlemini içermektedir . Verileri otomatik olarak analiz edebilmek için istatistiksel örüntü tanıma ve modelleme kullanan makine öğrenmesi, verilerden sürekli öğrenmeye, mevcut modelleri yenilemeye ve verilerden sonuçları çıkarmaya imkan sağlar (Chapman ve diğerleri, 2000, s. 16).




**Şekil 0.2.** Makine Öğrenmesi (Chapman ve diğerleri, 2000, s. 16)

Makine öğrenmesinde kullanılan algoritmalar eldeki verileri analiz ederek kendisini eğitir. Bilgisayarın kendisine kazandırdığı bu yetenek sayesinde makineye sorulan işlemlerin sonucu için bir tahminde bulunur. Literatürde 26 adet makine öğrenme algoritması mevcuttur. Bu algoritmalar, eldeki verilere göre başarı sonucu değişmektedir. Bu sebeple belirli bir yöntem ile tüm veriler ile işlem yapmak doğru değildir (Kocamaz, 2012).

Makine Öğrenmesi, dışarıdan herhangi bir insan müdahalesine maruz kalmadan yetkilendirildiği veriler ile bilgi toplayarak kendisini eğitmesi ve sonuca varması için tasarlanmış yapay zekâ dalıdır. Eğer bir insanın baş edemeyeceği büyüklükte bir veri seti mevcutsa makine öğrenmesi bu durumda bizi kurtaracaktır (Aslan). Makine Öğrenmesi, bahsedilen veriyi düzgün ve başarılı bir şekilde işleyerek sonuca ulaşacaktır.

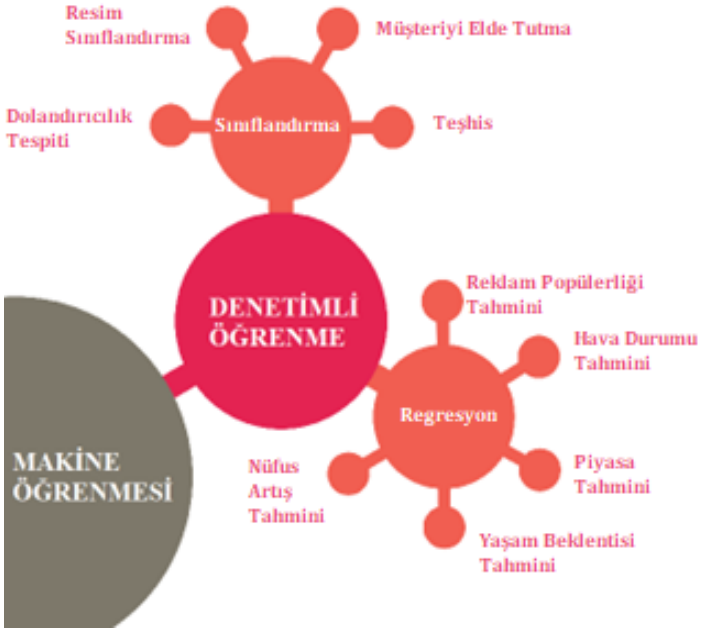
Makine Öğrenmesi, günümüzdeki firmaların en fazla çalışma sürdürdüğü yapay zeka türü olabilir. Gerçek Dünya Yapay Zekasına Yönelik Yönetici Klavuzu ‘nun bildirisine göre Makine Öğrenmesi, uzun süredir kullanılan bir türdür (Aslan).

Makine Öğrenmesi; Denetimli Öğrenme, Denetimsiz Öğrenme ve Takviyeli Öğrenme olarak 3 temel başlıkta incelenmektedir.


**Şekil 0.3.** Makine Öğrenmesi Türleri (Shewan, 2017)

2.3.1.1. Denetimli Öğrenme

Belirsizlik içerisinde tahminler yaparken kanıta dayanarak bir model oluşturan Makine Öğrenmesi yöntemine Denetimli Öğrenme denir. Denetimli öğrenme kategorisinde bulunan bir algoritmaya girdi olarak tanıtılan veri seti için sonuç olarak bilinen cevapları alır ve bu algoritmaya sonuç almak için gönderilecek veriler için uygun tahminler oluşturmak amacıyla bir model eğitir (Shewan, 2017).



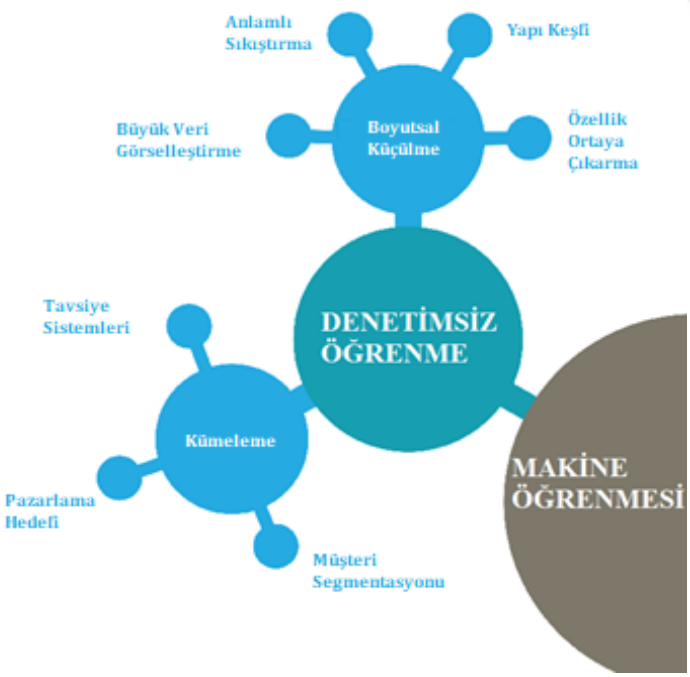

**Şekil 0.4.** Denetimli Öğrenme (Shewan, 2017)

Denetimli Öğrenme 'nin bir tekniği olan Sınıflandırma tekniği, algoritmaya girilen verileri kategorileştirir. Sınıflandırma için kullanılan en yaygın algoritmalar, Lojistik Regresyon, Yapay Sinir Ağları, Destek Vektör Makinesi, Karar Ağaçları, Diskriminant Analizi , K-En Yakın Komşu ve Naïve Bayes algoritmalarıdır (Shewan, 2017).

Diğer bir teknik olan Regresyon tekniği, bir değeri sürekli tahmin etmek için kullanılır. Bir evin büyüklüğü ve fiyatının tahmini en çok kullanılan örneklerindendir. En yaygın regresyon tekniğinin algoritmaları, Kademeli regresyon, yapay sinir ağları, doğrusal ve doğrusal olmayan model, regülasyon ve karar ağaçları algoritmalarıdır (Shewan, 2017).

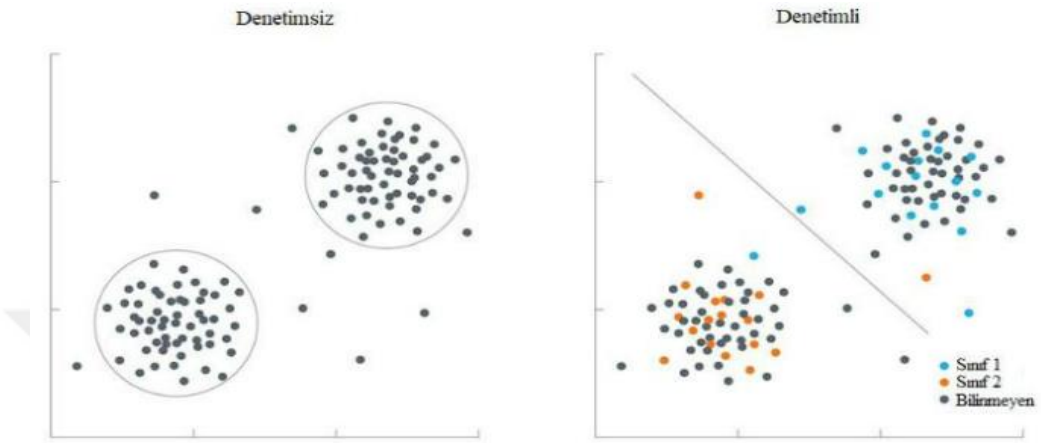
2.3.1.2. Denetimsiz Öğrenme

Denetimsiz Öğrenme modeli, denetimli öğrenme gibi denetlenmeye ihtiyaç duymayan bir tekniktir. Denetimsiz öğrenme algoritmasının oluşturduğu modelin, girdi olarak aldığı veri setindeki bilgileri anlaması için kendi kendine çalışmalıdır. Bu türde kullanılan algoritmalar, denetimli öğrenmeye göre daha karmaşık işlemler yapabilmemizi sağlar. Bu teknikte algoritmanın oluşturduğu modele sistem öğretilmemekte, girdi olarak verilen veri setinden kendi kendine öğrenmektedir. Denetimsiz Öğrenme tekniğinde kategorizasyon kullanılmaktadır (Shewan, 2017).




**Şekil 0.5.** Denetimsiz Öğrenme (Shewan, 2017)

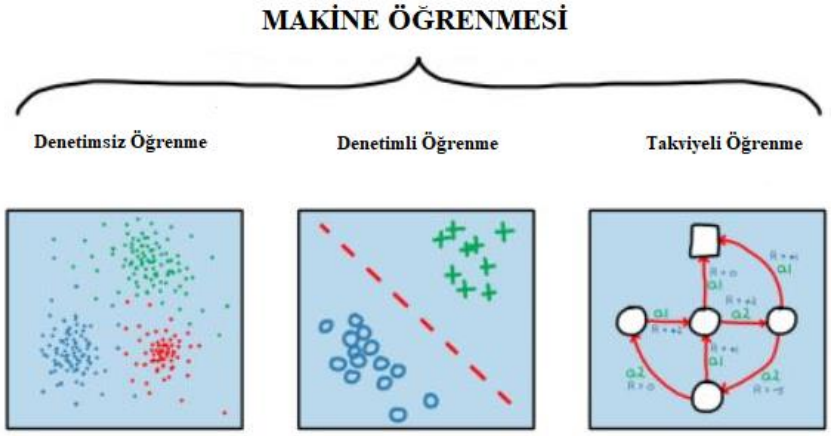
Denetimsiz Öğrenme tekniğinde en önemli kavram olan Kümeleme yöntemi, categorize edilmemiş bir veri seti ile bir model oluşturmaktadır. Kümeleme tekniğinde kullanılan algoritmalar, veri setlerini analiz eder, işler ve yapılabiliyorsa kümeleme veya başka bir deyişle gruplama işlemini oluşturur.




**Şekil 0.6.** Denetimli- Denetimsiz Karşılaştırma (Shewan, 2017)

2.3.1.3. Takviyeli Öğrenme

Bu yöntem, denetimli ve denetimsiz öğrenme yöntemlerinden farklıdır. Bir model, adım adım öğrenilerek, veri setinden üretilecek sonuçlar için doğruluğu control edecek bir eğitmen söz konusudur. Algoritma bir sonuç verir ve sonucun doğruluğuna göre makine ödüllendirilir ya da cezalandırılır. Makinenin asıl gayesi, çevresiyle etkileşimde kalması ve sonucunda ise kendi kendisine geri besleme yaparak en uygun modeli geliştirmesidir (Loon, 2014).




**Şekil 0.7.** Makine Öğrenmesi Özet (Akdağlı, 2021)

# 3. BÖLÜM

## MATERYAL VE METOT

### 3.1. Veri Seti Tasarımı

Veri seti, 2020-2021 yılları arasında toplam 365 günlük kredi kartlarıyla yapılan günlük alışveriş işlemlerini içermektedir. Toplam 10 bin müşteri ve 15 bin alışveriş noktasını içerecek şekilde belirlenen veriler, geliştirilen veri üretme algoritmasından geçirilerek elde edilmiştir. Veri seti oluşturulduktan sonra tüm kredi kartı ile yapılan işlemlerin ayrıntıları incelenmiştir.

Bir kredi kartı alışveriş işleminde ne tür verilere ihtiyacımızın olduğu belirlenmiş sonucunda kart kullanıcısının yaptığı işlem sonrasında 6 adet veri seti sütun başlığımız oluşmuştur.

- ISLEM_ID: Alışveriş işlemlerini birbirlerinden ayırdığımız işlem kimliklendirilmesi için kullanılan sütun,
- ISLEM_ZAMANI: Kredi kartının kullanım tarih ve zamanı tutmak için kullanılan sütün,
- MUSTERI_ID: Müşterileri birbirinden ayırdığımız müşteri kimliklendirilmesi için kullanılan sütun,
- ISLEM_YERI_ID: Alışveriş yapılan yerleri birbirinden ayırdığımız alışveriş noktalarının kimliklendirilmesi için kullanılan sütun,
- ISLEM_MIKTARI: Alışveriş işlemlerinin tutarı,
- DOLANDIRMA_MI: Kredi kartı dolandırıcılık tespiti. 0 ise normal işlem, 1 ise dolandırıcılık işlemi.

Sütün başlıkları belirlenen veri setinin örnek tablosu Tablo 3.1 ‘de gösterilmiştir.

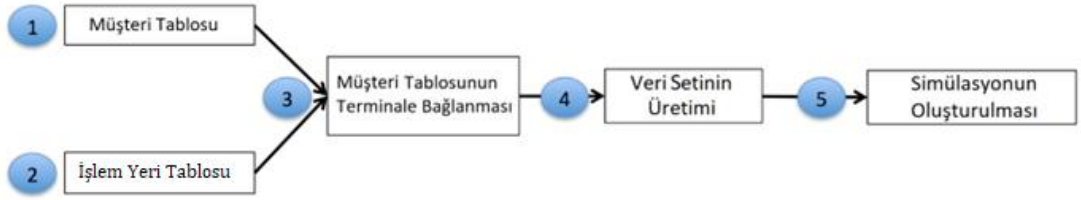
İlk olarak bir kredi kartı kullanım işleminde temel olarak ne tür verilere ihtiyacımız olduğunu belirlememiz gerekmektedir. Bir müşteri kredi kartını kullandıktan sonra, 6 adet sütun başlığımız olacaktır. Bunlar:

**Tablo 0.1.** Veri Seti Örnek Tablo

| ISLEM_ID | ISLEM_ZAMANI | MUSTERI_ID | ISLEM_YERI_ID | ISLEM_MIKTARI | DOLANDIRMA_MI |
|---|---|---|---|---|---|
| 0 | 2020-04-01 00:31:36 | 596 | 3156 | 57.16 | 0 |
| 1 | 2020-04-01 00:34:14 | 4961 | 3412 | 87.51 | 0 |
| 2 | 2020-04-01 00:36:41 | 2 | 1365 | 146.00 | 0 |
| … | … | … | … | … | … |

## 3.2. Simülasyon Planlaması

Simülasyon tasarımımız 5 adımdan oluşmaktadır:




**Şekil 0.1.** Simülasyon Planlama Adımları

- Müşteri Profilleri Oluşturma: Her müşterinin kredi kartı harcamaları farklıdır. Müşterinin özellikleri, coğrafi konumları, harcama sıklıkları ve harcama miktarları olarak Müşteri Profil Tablosu 'nda oluşturulacaktır.
- İşlem Yeri Profilleri Oluşturma: Alışveriş yapılan yerler olarak tanımlanan, sadece coğrafi konum içeren İşlem Yeri Tablosu 'nda oluşturulacaktır.
- Müşteri Profillerinin Terminale Bağlanması: Müşterilerin, yalnızca kendi konumuna belirli sınırlar içerisindeki alışveriş noktalarında işlem yaptığını varsayacağız. Çünkü bir müşterinin kendi coğrafi konumuna yakın yerlerden alışveriş yapması beklenir. Bu adımda her bir müşteri profiline, alışveriş yapabileceği işlem yeri setini içeren bir 'list_terminals' özelliği eklenecektir.
- Veri Seti Üretimi: Bu adımda algoritma, müşteri profilleri tablosundaki müşteri özelliklerine göre (harcama sıklıkları, harcama miktarları ve terminal listesi) veriler üretecek ve İşlem Tablosu olarak çıktı verecektir.
- Dolandırıcılık Simülasyonunun Oluşturulması: Simülasyonumuz, planlamanın son adımında, üretilmiş verilere göre dolandırıcılık işlemi olup olmadığını belirlemeye çalışacaktır.

## 3.3. Simülasyon Planlanmasının Adımları

### 3.3.1. Müşteri profil tablosu

Müşteriler; müşteri_id, x_musteri_id, y_musteri_id, ort_miktar, std_miktar, ort_gunluk_ alisveris özelliklerine sahip olacaktır.

MUSTERI_ID; Müşteri ID 'si, x_musteri_id; 100x100 Koordinat düzleminde gerçek koordinatın X konumu, y_musteri_id; 100x100 Koordinat düzleminde gerçek koordinatın Y konumu, ort_miktar; Alışveriş tutarlarının ortalaması (Normal dağılım izlediği varsayılacak), std_miktar; Alışveriş tutarlarının standart sapması (Normal dağılım izlediği varsayılacak) ve ort_gunluk_alisveris; Müşterinin günlük ortalama alışveriş sayısını (Poisson dağılım izlediği varsayılacak, bu sayı normal dağılımdan (0,4) çekilecektir.) temsil edecektir.

**Tablo 0.2.** Örnek Müşteri Tablosu

| | MUSTERI_ID | x_musteri_id | y_musteri_id | ort_miktar | std_miktar | ort_gunluk_alisveris |
|---|---|---|---|---|---|---|
| 0 | 0 | 54.881350 | 71.518937 | 62.262521 | 31.131260 | 2.179533 |
| 1 | 1 | 42.365480 | 64.589411 | 46.570785 | 23.285393 | 3.567092 |
| 2 | 2 | 96.366276 | 38.344152 | 80.213879 | 40.106939 | 2.115580 |
| 3 | 3 | 58.804456 | 92.559664 | 11.748426 | 5.874213 | 0.348517 |
| 4 | 4 | 2.021840 | 83.261985 | 78.924891 | 39.462446 | 3.480049 |

### 3.3.2. İşlem yeri tablosu

İşlem yeri olarak adlandırılan alış-veriş işlemlerinin gerçekleştiği noktalar, işlem id, x işlem yeri id ve y işlem yeri id özelliklerine sahip olacaklardır.

ISLEM_ID sütunu; alış-veriş işlemlerinin gerçekleştiği noktaların benzersiz numaralarını, x_islem_yeri_id sütunu; 100x100 lük bir koordinat düzleminde terminalin bulunduğu x konumunu, ve son olarak y_islem_yeri_id sütunu; 100x100 lük bir koordinat düzleminde terminalin bulunduğu y konumunu tutmaktadır. Yazılımda oluşturduğumuz generate_terminal_profiles_table fonksiyonu bahsettiğimiz bu özelliklerde ki işlem yeri tablosunu oluşturmuştur.

**Tablo 0.3.** Örnek İşlem Yeri Tablosu

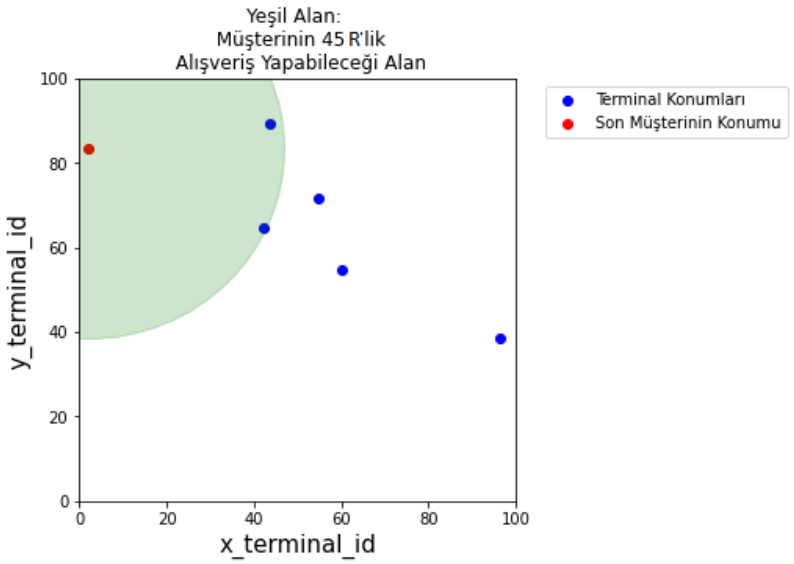
| | ISLEM_YERİ_ID | x_islem_yeri_id | y_islem_yeri_id |
|---|---|---|---|
| 0 | 0 | 54.881350 | 71.518937 |
| 1 | 1 | 60.276338 | 54.488318 |
| 2 | 2 | 42.365480 | 64.589411 |
| 3 | 3 | 42.758721 | 89.177300 |
| 4 | 4 | 96.366276 | 38.344152 |

### 3.3.3. Müşteri tablosunun işlem yeri tablosuna bağlanması

Müşteri tablosunda oluşturulmuş müşterilerin, bulunduğu mevcut konumu üzerinden belirli sınırlar içerisindeki alış-veriş noktalarından (işlem yeri) işlem yapabildiği kuralını kullanarak, Müşteri Tablosu ve Terminal Tablosu arasında bir ilişkilendirme yapıyoruz. Bahsedilen bu

sınırı bir çember olarak düşünüp belirli bir yarıçap oranı veriyoruz. Yazılımımıza get_list_terminals_within_radius adında bir fonksiyon ekliyor ve bu sınırları müşterilerle ilişkilendiriyoruz.

Bu sınırların çapını 100x100 koordinat düzleminde 45 birim olarak belirledik ve oluşturduğumuz bu fonksiyonu çağırdık. Bahsedilen bu sınırlandırma ilişkisini görselleştirmek için matplotlip.pyplot adlı kütüphaneyi yazılımımıza ekliyoruz. Yazılan kod scriptinden sonra oluşacak grafik Şekil 3.2 'de gösterilmiştir. Mavi renkli noktalar işlem yerlerini (alış-veriş noktası), Kırmızı renkli noktalar müşteri dediğimiz kredi kartı sahibinin bulunduğu konum ve müşteri etrafında oluşan yeşil renkli sınır ise 45 birim çapa sahip çemberimizi belirtmektedir.




**Şekil 0.2.** Müşteri- Terminal İlişkilendirme

Yazılımımıza eklediğimiz kod scriptinde, panda kütüphanesini kullanarak, her bir müşterinin (kredi kartı sahibinin) alışveriş yapabileceği terminalleri, 3.3.1. başlıkta oluşturduğumuz Müsteri Profilleri Tablosunda, mevcut_islem_yerleri adında bir sütun ekleyerek tutuyoruz.

**Tablo 0.4.** Terminallerle ilişkilendirilmesi Biten Müşteri Profil Tablosu

| | MUSTERI_ID | x_musteri_id | y_musteri_id | ort_miktar | std_miktar | ort_gunluk_alisveris | mevcut_islem_yerleri |
|---|---|---|---|---|---|---|---|
| 0 | 0 | 54.881350 | 71.518937 | 62.262521 | 31.131260 | 2.179533 | [0, 1, 2, 3] |
| 1 | 1 | 42.365480 | 64.589411 | 46.570785 | 23.285393 | 3.567092 | [0, 1, 2, 3] |
| 2 | 2 | 96.366276 | 38.344152 | 80.213879 | 40.106939 | 2.115580 | [1, 4] |
| 3 | 3 | 58.804456 | 92.559664 | 11.748426 | 5.874213 | 0.348517 | [0, 1, 2, 3] |
| 4 | 4 | 2.021840 | 83.261985 | 78.924891 | 39.462446 | 3.480049 | [2, 3] |

### 3.3.4. Veri setinin üretimi

Müşteri olarak adlandırdığımız kredi kartı sahiplerini tuttuğumuz Müşteri Profilleri Tablosu artık veri oluşturulması için hazır hale geldi. Müşteri profillerini, işlemlerin başlangıç tarihi ve ne kadarlık bir işlem süresi olacağını girdi olarak generate_transactions_table adında bir fonksiyon oluşturup yazılımımıza ekliyoruz. Örnek olarak bir müşteri için 10 günlük işlemler oluşturalım. Tablo 3.5 'de de bu müşteri için nasıl bir tablo oluşacağını görüyoruz.

**Tablo 0.5.** Bir Müşteri 'nin 10 Günlük İşlemi

| | ISLEM_ZAMANI | MUSTERI_ID | ISLEM_YERI_ID | ISLEM_MIKTARI | ISLEM_SURESI_saat | ISLEM_SURESI_gun |
|---|---|---|---|---|---|---|
| **0** | 1.01.2020 07:19 | 0 | 3 | 123.59 | 26345 | 0 |
| **1** | 1.01.2020 19:02 | 0 | 3 | 46.51 | 68522 | 0 |
| **2** | 1.01.2020 18:00 | 0 | 0 | 77.34 | 64816 | 0 |
| **3** | 2.01.2020 15:13 | 0 | 2 | 32.35 | 141182 | 1 |
| **4** | 2.01.2020 14:05 | 0 | 3 | 63.30 | 137138 | 1 |
| **5** | 2.01.2020 15:46 | 0 | 3 | 13.59 | 143211 | 1 |
| **6** | 2.01.2020 08:51 | 0 | 2 | 54.72 | 118266 | 1 |
| **7** | 2.01.2020 20:24 | 0 | 3 | 51.89 | 159887 | 1 |
| **8** | 3.01.2020 12:15 | 0 | 2 | 117.91 | 216947 | 2 |
| **9** | 3.01.2020 08:50 | 0 | 1 | 67.72 | 204609 | 2 |
| **10** | 3.01.2020 09:25 | 0 | 1 | 28.46 | 206749 | 2 |
| **11** | 3.01.2020 15:33 | 0 | 2 | 50.25 | 228794 | 2 |
| **12** | 3.01.2020 07:41 | 0 | 1 | 93.26 | 200484 | 2 |
| **13** | 4.01.2020 01:15 | 0 | 0 | 46.40 | 263735 | 3 |
| **14** | 4.01.2020 09:33 | 0 | 2 | 23.26 | 293638 | 3 |
| **15** | 5.01.2020 16:19 | 0 | 1 | 71.96 | 404349 | 4 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **16** | 5.01.2020 07:41 | 0 | 2 | 52.69 | 373279 | 4 |
| **17** | 6.01.2020 09:57 | 0 | 0 | 61.39 | 467869 | 5 |
| **18** | 6.01.2020 14:22 | 0 | 0 | 64.33 | 483766 | 5 |
| **19** | 7.01.2020 13:40 | 0 | 2 | 42.52 | 567649 | 6 |
| **20** | 8.01.2020 09:59 | 0 | 3 | 36.95 | 640745 | 7 |
| **21** | 10.01.2020 02:24 | 0 | 0 | 34.02 | 786274 | 9 |
| **22** | 10.01.2020 12:17 | 0 | 2 | 84.96 | 821838 | 9 |

Tablo 3.5 'de ISLEM_ZAMANI sütunu; işlemin yapıldığı tarih ve saati tutmakta, MUSTERI_ID sütunu; işlemde bahsi geçen müşteri (kredi kartı sahibi) kimliğini tutmakta, ISLEM_YERI_ID sütunu; işlemin gerçekleştiği alışveriş noktasının kimliğini tutmakta, ISLEM_MIKTARI sütunu; alışveriş işleminde ne kadar birimlik bir harcama yapıldığının bilgisini tutmakta, ISLEM_SURESI_saniye sütunu; alış veriş işleminde bahsi geçen müşterinin kredi kartının toplam kaç saniye kullanıldığı bilgisini tutmakta ve son olarak ISLEM_SURESI_gun sütunu ise yine bahsi geçen müşterinin kredi kartının toplam kaç gün kullanıldığının bilgisini tutmaktadır. Oluşturulan bu işlemlerin, müşteri profil özellikleri ile uyuşup uyuşmadığını Tablo 3.6 yardımı ile kontrol edebiliriz.

**Tablo 0.6.** Üretilen Müşteri Profili 'nin Kontrolü

| MUSTERI_ID | 0 |
|---|---|
| x_musteri_id | 54.88135 |
| y_musteri_id | 71.518937 |
| ort_miktar | 62.262521 |
| std_miktar | 31.13126 |
| ort_gunluk_alisveris | 2.179533 |
| mevcut_islem_yerleri | [0, 1, 2, 3] |
| Name: 0, dtype: object | |

ISLEM_YERI_ID kodlarda belirlediğimiz alışveriş noktalarından birisini almıştır (0, 1, 2 ve 3). İşlem tutarları müşterinin harcama miktarıyla uyuşuyor gibi göründüğünü ort_miktar = 62.26 ve std_miktar = 31.13 değerlerinden anlayabiliriz. Günlük işlem sayısı da ort_gunluk_alisveris = 2.18 değerine bakarsak, müşterinin alışveriş sıklığına göre değişmektedir. Veri setimizin oluşup oluşmayacağını görebilmek için 10 müşteri, 10 terminal ve 10 gün değerlerini kullanarak bir veri seti üretelim. Sonuçları ise Tablo 3.7 de görelim.

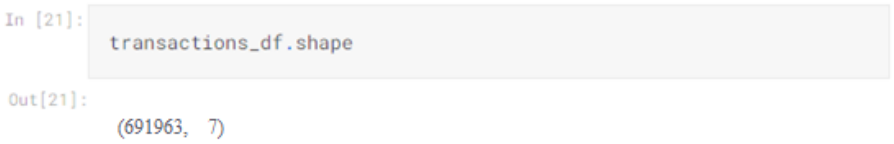
**Tablo 0.7.** Küçük Bir Veri Seti Oluşturma

| | ISLEM_ZAMANI | MUSTERI_ID | ISLEM_YERI_ID | ISLEM_MIKTARI | ISLEM_SURESI_saat | ISLEM_SURESI_gun |
|---|---|---|---|---|---|---|
| 0 | 2020-01-01 07:19:05 | 0 | 3 | 123.59 | 26345 | 0 |
| 1 | 2020-01-01 19:02:02 | 0 | 3 | 46.51 | 68522 | 0 |
| 2 | 2020-01-01 18:00:16 | 0 | 0 | 77.34 | 64816 | 0 |
| 3 | 2020-01-02 15:13:02 | 0 | 2 | 32.35 | 141182 | 1 |
| 4 | 2020-01-02 14:05:38 | 0 | 3 | 63.30 | 137138 | 1 |
| ... | ... | ... | ... | ... | ... | ... |
| 112 | 2020-01-09 10:07:56 | 4 | 3 | 82.13 | 727676 | 8 |
| 113 | 2020-01-09 10:39:49 | 4 | 2 | 159.12 | 729589 | 8 |
| 114 | 2020-01-09 14:21:56 | 4 | 3 | 12.51 | 742916 | 8 |
| 115 | 2020-01-10 09:35:22 | 4 | 2 | 89.09 | 812122 | 9 |
| 116 | 2020-01-10 07:38:44 | 4 | 3 | 51.23 | 805124 | 9 |

Girdi olarak aldığımız 10 müşteri, 10 işlem yeri ve 10 gün girişleri bize 117 adet veri oluşturmuştur.

### 3.3.5. Büyük veri seti üretimi

Büyük bir veri seti oluşturmak için tüm işlemleri ve gerekli alt yapıları oluşturduk. Bahsettiğimiz 3. Başlık altındaki tüm fonksiyonlarımızı tek bir fonksiyon üzerinden yürütebilmek için generate_dataset fonksiyonunu oluşturalım. Girdi olarak 10.000 müşteri, 15.000 işlem yeri ve başlangıç tarihi olarak da 2020/01/01 olarak alıp toplam 365 günlük işlem üretimine başlıyoruz. Üretilen veri büyük olduğu için bu aşama biraz uzun sürmüştür. Ürettiğimiz işlem sayısı 7.080.983 adet olup doğruluğunu test etmek için yazdığımız "transactions_df.shape" kod parçası için aldığımız sonuç Şekil 3.3 'de gösterilmiştir.




**Şekil 0.3.** Üretilen Toplam Veri Sayısı

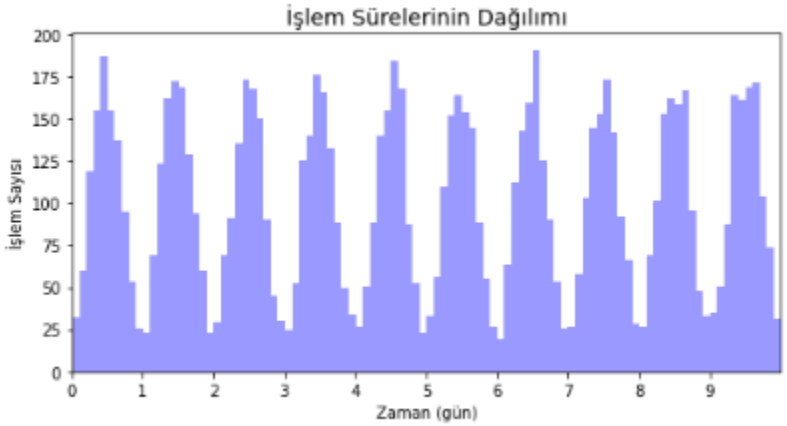
Veri setimizin ilk 5 ve son 5 verisini görmek için yazdığımız “transactions_df” kod parçası için aldığımız sonuç Tablo 3.8 ‘de gösterilmiştir.

**Tablo 0.8.** Üretilen Veri Setinin İlk ve Son 5 Verisi

| | ISLEM_ID | ISLEM_ZAMANI | MUSTERI_ID | ISLEM_YERI_ID | ISLEM_MIKTARI | ISLEM_SURESI_saat | ISLEM_SURESI_gun | DOLANDIRMA_MI | SENARYO |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 2020-01-01 00:00:31 | 596 | 298 | 57.16 | 31 | 0 | 0 | 0 |
| 1 | 1 | 2020-01-01 00:07:56 | 2 | 51 | 146.00 | 476 | 0 | 0 | 0 |
| 2 | 2 | 2020-01-01 00:10:34 | 927 | 1159 | 50.99 | 634 | 0 | 0 | 0 |
| 3 | 3 | 2020-01-01 00:10:45 | 568 | 1171 | 44.71 | 645 | 0 | 0 | 0 |
| 4 | 4 | 2020-01-01 00:13:44 | 541 | 964 | 59.07 | 824 | 0 | | |
| ... | ... | | ... | ... | ... | ... | ... | | |
| 691958 | 691958 | 2020-12-30 23:43:48 | 618 | 1494 | 7.19 | 31535028 | 364 | | |
| 691959 | 691959 | 2020-12-30 23:44:10 | 972 | 1398 | 17.19 | 31535050 | 364 | 0 | 0 |
| 691960 | 691960 | 2020-12-30 23:46:10 | 893 | 1336 | 28.86 | 31535170 | 364 | 0 | 0 |
| 691961 | 691961 | 2020-12-30 23:50:13 | 617 | 101 | 12.89 | 31535413 | 364 | 0 | 0 |
| 691962 | 691962 | 2020-12-30 23:51:09 | 515 | 1389 | 53.78 | 31535469 | 364 | 0 | 0 |

İşlem miktarları ve süreleri için dağılım grafiklerini oluşturacak kod scriptimiz oluşturulmuş ve çalıştırılmıştır. Şekil 3.4 ve Şekil 3.5 ‘de dağılım grafikleri gösterilmiştir.


**Şekil 0.4.** İşlem Miktarlarının Dağılım Grafiği




**Şekil 0.5.** İşlem Sürelerinin Dağılım Grafiği

İşlem miktarlarının dağılım grafiğine bakarsak, işlem miktarının büyük çoğunluğu küçük miktarlardan oluşmaktadır. İşlem sürelerinin dağılım grafiğine baktığımızda ise Gauss dağılımını takip ettiğini görmekteyiz. 2 grafiğinde dağılımı 3. Başlık altındaki diğer simülasyon adımlarındaki parametrelerle uyumludur.

### 3.3.6. Simülasyonun oluşturulması

Veri setinin oluşturulduğu ve veri üretiminin son adımı olan bu başlıkta aşağıdaki kuralları kullanarak yapılan alışveriş işlemlerinin dolandırıcılık olup olmadığını belirlemektedir.

Kural 1: Alışveriş miktarı eğer 300 'den büyük ise bu işlem dolandırıcılık işlemidir.

Kural 2: Her gün rastgele iki işlem yerinden oluşan liste çizilir. Önümüzdeki 28 gün içinde bu işlem yerlerindeki tüm işlemler sahte işlem olarak işaretlenecektir. Bu senaryo, örneğin kimlik avı yoluyla bir işlem yerindeki hileli işlemlerin sayısını takip eden özellikler eklenerek mümkün olacaktır. İşlem yeri yalnızca 28 gün boyunca tehlikeye girdiğinden, bu senaryoyla verimli bir şekilde başa çıkmak için konsept kaymasını içeren ek stratejilerin tasarlanması gerekecektir.

Kural 3: Her gün rastgele 3 müşteriden oluşan bir liste çizilir. Sonraki 14 gün içinde işlemlerinin 1/3 'ünün tutarları 5 ile çarpılır ve hileli olarak işaretlenir. Bu senaryo, bir müşterinin kimlik bilgilerinin sızdırıldığı, kartın mevcut olmadığı bir sahtekarlığı simüle eder.

Bu 3 kurala uyarak oluşturulan dolandırıcılık işlemleri DOLANDIRMA_MI sütununda tutulacaktır. Simülasyona eklenen kod script ve fonksiyonları ile dolandırıcılık işlemleri eklenmiştir. Her bir kural için ayrı ayrı eklenen işlemler Şekil 3.6 'da fonksiyonun çağrılması

ile 1 dakika 18 saniye içinde eklenmiş olup, dolandırıcılık işlemlerin yüzdesi Şekil 3.7 ‘de, eklenen dolandırıcılık işlemlerinin sayısı ise 3.8 ‘de gösterilmiştir.

```
[28]: %time transactions_df = add_frauds(customer_profiles_table, terminal_profiles_table, transactions_df)

Number of frauds from scenario 1: 354
Number of frauds from scenario 2: 24246
Number of frauds from scenario 3: 9256
CPU times: user 1min 18s, sys: 78.8 ms, total: 1min 18s
Wall time: 1min 18s
```

**Şekil 0.6.** Dolandırıcılık İşlemleri Ekleme

```
[29]: transactions_df.TX_FRAUD.mean()

[29]: 0.04892747155556005
```

**Şekil 0.7.** Dolandırıcılık İşlemlerinin Yüzdesi

```
[30]: transactions_df.TX_FRAUD.sum()

[30]: 33856
```

**Şekil 0.8.** Dolandırıcılık İşlemlerinin Sayısı

Simülasyonumuza 3 kurala göre toplam 39100 dolandırıcılık işlemi eklenmiş oldu. Bu işlemler, tüm veri setinin %0.5 ine tekabül etmektedir. Üretimi tamamlanan veri setinin ilk 5 verisi Tablo 3.9 ’da gösterilmiştir.

**Tablo 0.9.** Veri Setinin İlk 5 Verisi

| | ISLEM_ID | ISLEM_ZAMANI | MUSTERI_ID | ISLEM_YERI_ID | ISLEM_MIKTARI | ISLEM_SURESI_saat | ISLEM_SURESI_gun | DOLANDIRMA_MI | SENARYO |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 2020-01-01 00:00:31 | 596 | 298 | 57.16 | 31 | 0 | 0 | 0 |
| 1 | 1 | 2020-01-01 00:07:56 | 2 | 51 | 146.00 | 476 | 0 | 0 | 0 |
| 2 | 2 | 2020-01-01 00:10:34 | 927 | 1159 | 50.99 | 634 | 0 | 0 | 0 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| 3 | 3 | 2020-01-01 00:10:45 | 568 | 1171 | 44.71 | 645 | 0 | 0 | 0 |
| 4 | 4 | 2020-01-01 00:13:44 | 541 | 964 | 59.07 | 824 | 0 | 0 | 0 |

Veri seti üretme simülasyonunun tamamlanmasının ardından üretilen veriler fraudData isminde ve csv uzantılı olarak kayıt edilmiştir.

Veri seti, üzerinde işlemlerin yapılabilmesi için pandas kütüphanesinin bir özelliği olan Data Frame olarak yazılımla çağrılmaktadır. Veri setinin kontrolü için Data Frame 'e atanan bu verinin ilk 5 verisi Tablo 3.10 'da gösterilmiştir.

**Tablo 0.10.**Data Frame 'in İlk 5 Verisi

| | ISLEM_ID | ISLEM_ZAMANI | MUSTERI_ID | ISLEM_YERI_ID | ISLEM_MKTARI | ISLEM_SURESI_saniye | ISLEM_SURESI_gun | DOLANDIRMA_MI |
|---|---|---|---|---|---|---|---|---|
| 0 | 0 | 2020-01-01 00:00:31 | 596 | 298 | 57.16 | 31 | 0 | 0 |
| 1 | 1 | 2020-01-01 00:07:56 | 2 | 51 | 146.00 | 476 | 0 | 0 |
| 2 | 2 | 2020-01-01 00:10:34 | 927 | 1159 | 50.99 | 634 | 0 | 0 |
| 3 | 3 | 2020-01-01 00:10:45 | 568 | 1171 | 44.71 | 645 | 0 | 0 |
| 4 | 4 | 2020-01-01 00:13:44 | 541 | 964 | 59.07 | 824 | 0 | 0 |

### 3.4. Makine Öğrenmesi Modeli

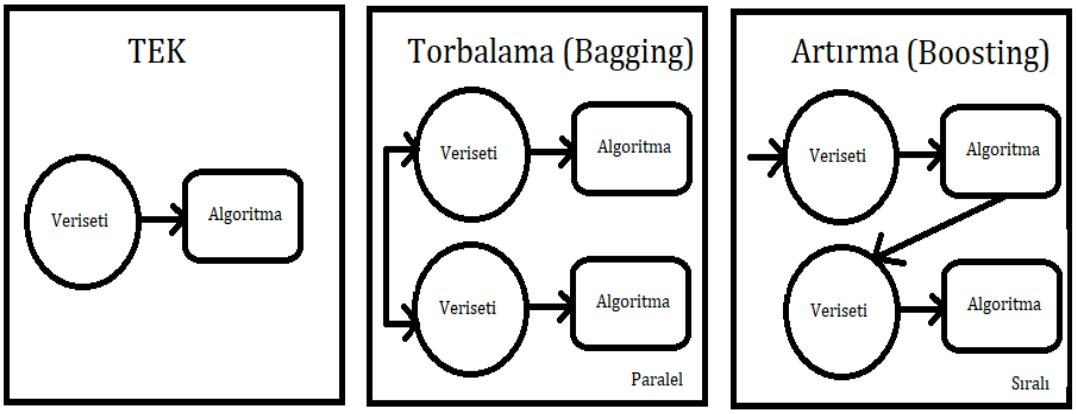
Makine Öğrenmesi için oluşturulan fonksiyonlara veri seti entegre edildikten sonra, bu veri setinin bir kısmı eğitim verisi diğer bir kısmı da geliştirilen modelin doğruluğunun testi için kullanılmıştır. Değer olarak veri setinin %70 'i eğitim için kalan %30 'u ise test verisi için kullanılmıştır. Modelimiz dolandırıcılık işlemi olup olmadığını öğrenebilmesi için oluşturulmuştur. Diğer sütunların ('ISLEM_ID',' MUSTERI_ID', 'ISLEM_YERI_ID', 'ISLEM_SURESI_saniye', 'ISLEM_SURESI_gun', 'ISLE_MMIKTARI', 'SENARYO') aldığı değerlere göre 'DOLANDIRMA_MI' sütununu karşılaştırmış ve gelişimini tamamlamıştır.

## 3.5. Makine Öğrenmesi Algoritmaları

Makine Öğrenmesine ait birçok algoritma vardır. Bu çalışmada kullanılan algoritmalar bu başlık altında anlatılmıştır. Herhangi bir makine öğrenme algoritmasının başka bir makine öğrenme algoritmasına üstünlüğü vardır diyemeyiz. Makine öğrenme algoritmasının ürettiği sonuçların doğruluk oranları, performansları, eğitilme süreleri gibi çoğu özelliği veri setinin büyüklüğüne ve yapısına göre değişiklik göstermektedir. Bu çalışmada, simülasyon ile üretilen veri seti için kullanılabilecek sınıflandırma algoritmalarının bazıları şunlardır; Karar Ağacı (Decision Tree), Rastgele Orman (Random Forest), Gradyan Artırılmış (Gradient Boosting) ve Yapay Sinir Ağları (Artificial Neural Network) sınıflandırma algoritmalarıdır.

### 3.5.1. Rastgele orman (random forest)

Topluluk (Ensemble) yöntemlerine dahil olan Rastgele Orman (Random Forest) algoritması, Torbalama yöntemi ve Karar Ağacı algoritmasının beraber kullanılması metoduna dayanmaktadır (Breiman, 2001). Birden fazla model eğitmek için bir tane öğrenme algoritması kullanan makine öğrenmesi kavramına Topluluk yöntemleri denir. Temel bir öğrenme algoritması, Artırma (Boosting) veya Torbalama (Bagging) yöntemlerini kullanmak için seçilmiş olmalıdır. Torbalama yönteminin rastgele özellik kullanımının doğruluk oranının artması için, özelliklerin farklı biçimlerde birleştirilmesiyle yeni ağaçlar oluşturulur ve en çok tanımlanan sınıf oluşturulan bu ağaçlardan seçilir (Breiman, 2001). Herhangi bir değişkenin yeni bir veri kümesinde yeniden tekrar etme olasılığı Torbalama yönteminde aynıdır ama bazı değişkenler yeni veri setlerinde daha fazla tekrar ettiğinden dolayı, artırma yönteminde değişkenler ağırlıklandırılır. Torbalama yöntemi için makinenin eğitim adımlarıyla pareleldir. Bu yüzden Topluluk yönteminde her bir model bağımsız olarak eğitilirken; Artırma, tahminlerin sırası ile yapıldığı bir yöntemdir. Yani bir sonraki tahmin adımları, önceden oluşan hatalardan öğrenirler (Breiman, 2001). Torbalama ve Artırma yöntemleri arasındaki fark Şekil 3.18 'de gösterilmiştir.




**Şekil 0.9.** Artırma ve Torbalama Yöntemleri (Breiman, 2001)

Esnek yapılı olan bir makine öürenmesi yöntemi olan Rastgele Orman, sınıflandırma ve regresyon algoritmaları için kullanılırlar. Rastgele orman, doğruluk oranı daha yüksek bir tahmin çıkarımı yapabilmek için karar ağaçlarının birleştirilmesiyle oluşmaktadır.

### 3.5.2. Gradyan artırılmış ağaçlar (gradient boosting trees)

Denetimli makine öğrenmesi yöntemlerinden birisi olan Gradyan Artırılmış Ağaçlar (Gradient Boosting Trees), sınıflandırma ve regresyon algoritmaları için kullanılırlar. Oluşacak her yeni bir ağaç, daha önce eğitilmiş ağaçlardan bilgiler ile kendini eğitir. X değişkenini gradyan algoritmasına girdi olacak özellikleri, Y değişkenini ise algoritmanın ürettiği sonucu temsil ettiği düşünüldüğünde, aşağıdaki denklemler gibi genelleştirilebilir.

Gradyan Artırma Algoritması 'nın ürettiği tahmini temsil eden $f_1$ fonksiyonu Şekil 3.19 'da gösterilmiştir.

$$f_1(x) = y,$$

**Şekil 0.10.** Tahmin Üreten Gradyan Artırma Fonksiyonu

Tahminlerin ve hedef değerin arasındaki farkın hesabını gösteren $h_1$ fonksiyonu Şekil 3.20 'de gösterilmiştir.

$$h_1(x) = y - f_1(x),$$

**Şekil 0.11.** Tahminler ve Hedef Arasındaki Fark ($h_1$)

Oluşan yeni ağacın denklemi Şekil 3.21 'de gösterilmiştir.

$$f_2(x) = f_1(x) + h_1(x),$$

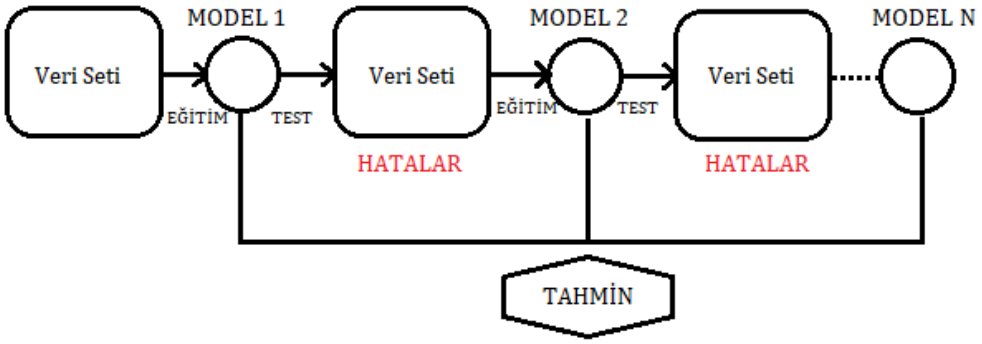
**Şekil 0.12.** Oluşan Yeni Ağaç

Tahminlerin ve hedef değerin arasındaki farkın yeniden hesaplanmasıyla oluşan $h_2$ fonksiyonu Şekil 3.22 'de gösterilmiştir.

$$h_2(x) = y - f_2(x),$$

**Şekil 0.13.** Farkın Yeniden Hesaplanması

Oluşturulan bu döngü, sürekli olarak tahminler ve hedef arasındaki farkı sıfıra en yakın şekle getirmeye çalışarak f fonksiyonunun başarısını artırmayı hedeflemektedir.



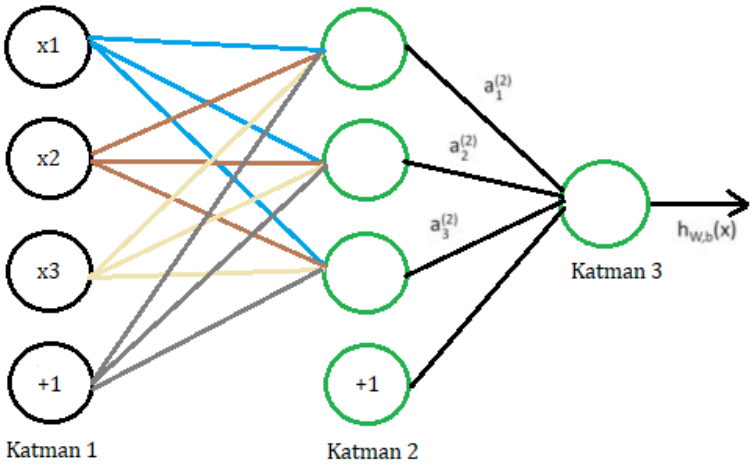

**Şekil 0.14.** Gradyan Artırma Yöntemi

Gradyan Artırma Yöntemi 'nin çalışma modeli Şekil 3.23 'de gösterilmiştir. Bu yönteminin temel amacı, hatayı en düşük orana çekmek ve bir sonraki oluşturulacak model için hedef çıktılar üretmektir. Yani önceki tahmin hatalarından kendine bir sonuç çıkartarak sonraki modellerin tahminlerinin ilerlemesine dayanmaktadır (Breiman, 2001).

### 3.5.3. Yapay sinir ağları (artificial neural network)

İnsan beyninin taklit ederek analiz işlemleri yapan ve bu verilerden yeni sonuçlar üreten yönteme Yapay Sinir Ağları (Artificial Neural Network) yöntemi denmektedir. Hataya karşı tolerans göstermesi, lineer olmayan (non-lineer) yapısı, adaptasyonu, girdi ve çıktıların eşleştirilmesiyle modellenebilmesi Yapay Sinir Ağları 'nı kullanmanın avantajlarındandır (Breiman, 2001).

Yapay Zekâ kavramının içerdiği ve araştırmacılar tarafından yoğun bir ilginin söz konusu olduğu Yapay Sinir Ağları, makinelerin öğrenmesine dayalı çalışmaları kapsayan bir alandır (Öztemel, 2003). Biyolojik sinir ağını taklit eden bir algoritma olduğu adına bakarak da çıkartılabilmektedir. Matematiksel bir model oluşturmak için birçok nöron birbirine bağlanırlar. Bu şekilde birbirine bağlanan nöronların oluşturduğu gruba sinir ağır denir.




**Şekil 0.15.** Basit Yapay Sinir Ağ Modeli

Basit bir yapay sinir ağ modeli Şekil 3.24 'de gösterilmiştir. Çoklu katman içerren modellerde, bir katmanın girdisi kendinden bir önceki katmanın girdisi olmaktadır. Optimum ağırlık değerleri bulunan sinir ağı modeline girdi olarak gösterilen veriler, Yapay Sinir Ağı için öğrenme model mantığıdır. Veri sayısının düşük miktarda olması, Yapay Sinir Ağı 'nın tahmin başarısını olumsuz etkileyebilmektedir.

### 3.5.4. Karar ağacı (decision tree)

Uygun bir özelliği belirlenen Karar Ağacı 'ndaki tüm düğümlerde genellikle bilgi kazanma yaklaşımı kullanılmaktadır. Bu yüzden maksimum entropi azalması mevcut düğümün test niteliği için tercih edilebilmektedir. Bu yöntemle elde edilen model, gereken bilgilerin sayısını minimumda tutarak alt kümelesini sınıflandırmaktadır. $C_i$i (1,2, 3,…, m) örnek sayısını temsil eden denkleme karşılık gelen m adet farklı değerler alabilen S tip özelliği, verilerin örnek sayılarını içeren sete denir.

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

**Şekil 0.16.** Gerekli Bilgi Miktarı

Eğitim modeline girilen veri setini sınıflandırmak için gerekli bilginin miktarının hesaplanmasını gösteren denklem Şekil 3.25 'de gösterilmiştir (Jin ve diğerleri, 2009).

Karar ağaçları, her ne kadar karmaşık veri setleri için kullanılsa da kullanımı basit ve eğitim modeli için etkili tahminler üreten sınıflandırma algoritmasıdır. Eğitim modelinin kesin bir tahmine

ulaşabilmesi için veri setinin çoğu özelliği belirli bir sıra ile kontrol edilerek karar verilmektedir. Karar ağaçlarında, sıranın ve özelliklerin belirlenmesinde birtakım sıkıntıları bulunsada, sınıflandırma kurallarının insan tarafından okunabilir olması bu algoritma için bir avantaj olarak kabul edilmektedir (Ben-Haim ve Tom-Tov, 2010).

Eğitilecek veri setleri için Karar Ağacı Modeli geliştirmek için, her bir özniteliği farklı değere sahip bir girdi almalıdır. Modelin çıktısı bir karar ağacını oluşturmaktadır. Modelde ki bir düğüm N oluşturulur (Jin ve diğerleri, 2009).

# 4. BÖLÜM

# ARAŞTIRMA BULGULARI

Bu bölüm, geliştirilen simülasyon ile üretilen veri setini eğitmek için kullanılan algoritmaların doğruluk oranlarını içermektedir. Araştırma da 24 adet makine öğrenmesi algoritması için test işlemi gerçekleştirilmiştir. 10 bin adet veri ile eğitilen makine öğrenmesi modelinin, test işlemi sonrasında Doğruluk (accuracy), ROC AUC ve F1 Skorları hesaplanmıştır.

**Tablo 0.1.** Algoritmaların Test Sonuçları

| Model | Doğruluk (Accuracy) | ROC AUC | F1 Score |
|---|---|---|---|
| **Hafif Gradyan Artırma Makinesi (LGBM)** | 0.89 | 0.89 | 0.89 |
| **Gradyan Artırma Makinesi (GBM)** | 0.88 | 0.88 | 0.88 |
| **Rastgele Orman Sınıflandırıcı** | 0.88 | 0.88 | 0.88 |
| **Sınıflandırma ve Regresyon Ağacı (CART)** | 0.87 | 0.87 | 0.87 |
| **CatBoost Modeli** | 0.87 | 0.87 | 0.87 |
| **Lojistik Regresyon** | 0.87 | 0.87 | 0.87 |
| **AdaBoost Sınıflandırıcı** | 0.87 | 0.87 | 0.87 |
| **Lineer SVC** | 0.87 | 0.87 | 0.87 |
| **Destek Vektör Sınıflandırıcı (SVC)** | 0.87 | 0.87 | 0.86 |
| **Ölçülü Sınıflandırıcı (Calibrated Classifier)** | 0.86 | 0.86 | 0.86 |
| **Olasılıksal Dereceli Azalma (SGD) Sınıflandırıcı** | 0.86 | 0.86 | 0.85 |
| **Karesel Ayırma Analizi** | 0.85 | 0.85 | 0.85 |
| **NuSVC** | 0.84 | 0.84 | 0.84 |
| **Gaussian Naive Bayes** | 0.84 | 0.84 | 0.83 |
| **Doğrusal Ayırma Analizi** | 0.83 | 0.83 | 0.83 |
| **Ridge Sınıflandırıcı** | 0.83 | 0.83 | 0.83 |
| **En Yakın Sendroid (Nearest Centroid)** | 0.82 | 0.82 | 0.81 |
| **Karar Ağacı Sınıflandırıcı** | 0.81 | 0.81 | 0.81 |
| **Bernoulli Naive Bayes** | 0.79 | 0.79 | 0.79 |
| **K-En Yakın Komşu Sınıflandırıcı** | 0.79 | 0.79 | 0.78 |
| **Etiket Yayılımı** | 0.76 | 0.76 | 0.76 |
| **Algılayıcı (Perceptron)** | 0.73 | 0.73 | 0.73 |
| **Pasif Agrasif Sınıflandırıcı** | 0.71 | 0.71 | 0.70 |
| **Dummy Sınıflandırıcı** | 0.50 | 0.50 | 0.34 |

Tablo 4.1 ‘de gösterilen Doğruluk ROC AUC ve F1 Skor ‘un anlamları şu şekildedir:

- Doğruluk: Doğru tahmin sayısının oranını temsil etmektedir. Doğru tahmin edilen dolandırıcılık işlemlerinin toplam dolandırıcılık işlemine bölünmesidir.

$$\text{DOĞRULUK (Accuracy)} = \frac{\sum TP + \sum TN}{\text{Bütün Gözlemlerin Sayısı}}$$

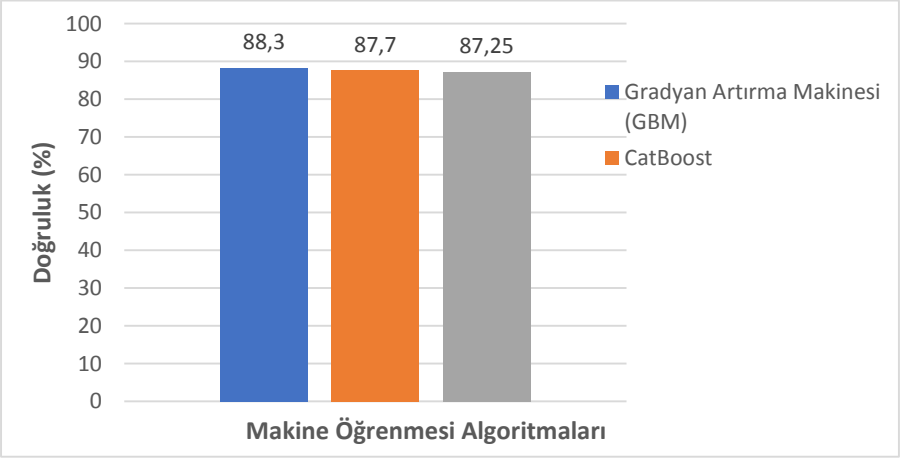
**Şekil 0.1.** Doğruluk Oranı Hesaplaması

- ROC AUC: ROC eğrisi, sınıflandırma algoritmaları için performans ölçümüdür. Olasılık eğrisini temsil etmektedir ve bu eğrinin altında kalan alana da AUC değerini vermektedir.
- F1 Score: Kesinlik (Precision) ve Duyarlılık (Recall) değerlerinin harmonik ortalamasını göstermektedir.

$$F_1 = 2 * \frac{\text{Kesinlik} * \text{Duyarlılık}}{\text{Kesinlik} + \text{Duyarlılık}}$$

**Şekil 0.2.** F1 Score Hesaplaması

Bu algoritmaların içerisinden seçilen doğruluk oranları yüksek olan algoritmalardan rastgele 3 adet algoritma seçilmiştir. Şekil 4.3 ‘deki sütun grafiğinde gösterilmiştir.




**Şekil 0.3.** Seçilen Algoritmaların Doğruluk Oranları

Şekil 4.2 ’deki grafik incelendiğinde en başarılı algoritmanın Gradyan Artırma Makinesi (GBM) olduğu görülmektedir. Test işlemi yapılan 10 bin dolandırıcılık işleminden 8830 tanesini doğru

tahmin etmiştir. Doğruluk oranı ise %88,3 'dür. Hemen ardından %87,7 doğruluk oranı ile CatBoost algoritması takip etmektedir. Algoritmaların karışıklık matrisleri Tablo 4.2 'den Tablo 4.5 'e kadar gösterilmiştir.

Tablolarda gösterilen TP, FP, TN, FN terimlerinin anlamları şu şekildedir;

- TP (Doğru Karar): Gerçek durumda, dolandırıcılık işlemlerinin doğru tahmin edildiği anlamına gelir.
- FP (Tip I Hatası): Gerçek durumda, dolandırıcılık işlemi olmayan kredi kartları, dolandırıcılık olarak bulunmuştur.
- TN (Doğru Karar): Gerçek durumda, dolandırıcılık işlemi olmayan kredi kartları, doğru tahmin edildiği anlamına gelmektedir.
- FN (Tip II Hatası): Gerçek durumda dolandırıcılık işlemi olan kredi kartları, dolandırıcılık değil olarak tahmin edilmiştir.
- Kesinlik (Precision): Pozitif (TP ve FP) olarak tahmin edilen değerlerin kaç tanesinin doğru tahmin olduğunu göstermektedir.

$$\text{Kesinlik} = \frac{TP}{TP + FP}$$

**Şekil 0.4.** Kesinlik Hesaplaması

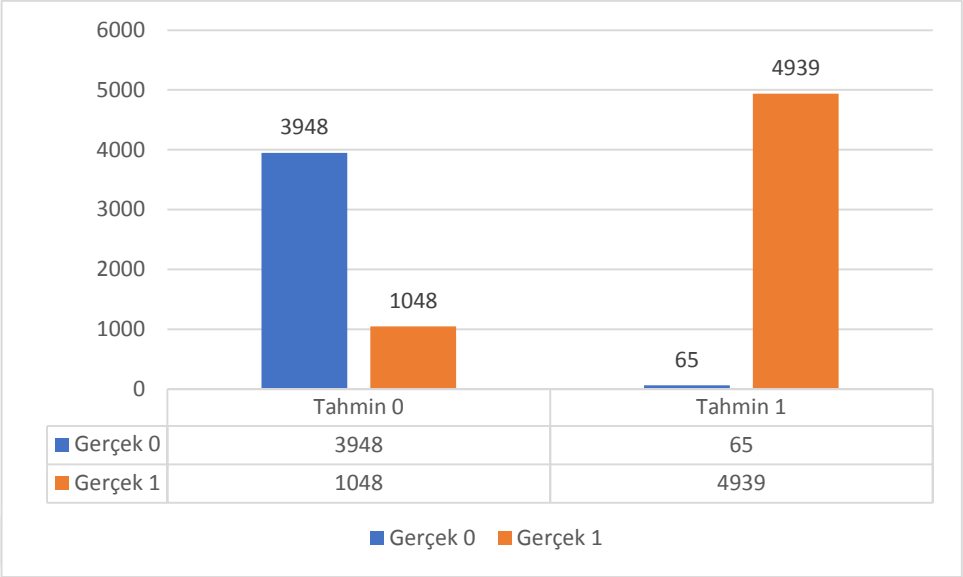
- Duyarlılık (Recall): Pozitif olarak tahmin edilmesi gereken işlemlerin kaç tanesini doğru olarak tahmin edildiğini göstermektedir.

$$\text{Duyarlılık} = \frac{TP}{TP + FN}$$

**Şekil 0.5.** Duyarlılık Hesaplaması

**Tablo 0.2.** Gradyan Artırma (GBM) Algoritması Karışıklık Matrisi

| **Doğruluk: %88,8** | **Gerçek 1** | **Gerçek 0** | **Toplam** | **Precision (Kesinlik)** |
|---|---|---|---|---|
| **Tahmin 1** | 4939 (TP)<br>Doğru Karar | 65 (FP)<br>Tip 1 Hatası | 5004 | %98,7 |
| **Tahmin 0** | 1048 (FN)<br>Tip 2 Hatası | 3948 (TN)<br>Doğru Karar | 4996 | %79 |
| **Toplam** | 5987 | 4013 | 10000 | |
| **Recall (Duyarlılık)** | %82,4 | %98,3 | | |



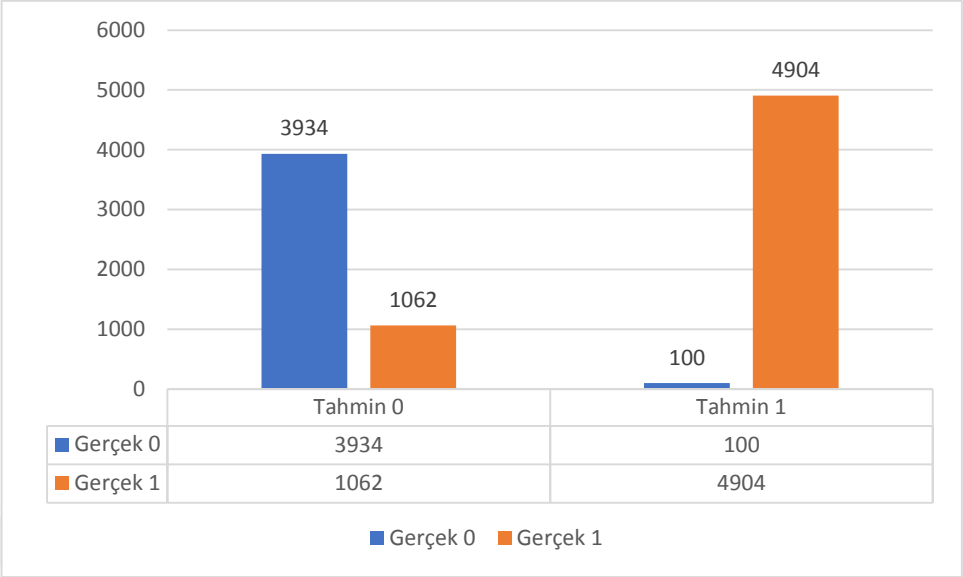

**Şekil 0.6.** Gradyan Artırma Algoritması Tahmin Sayıları

Tablo 4.2'ye göre Gradyan Artırma (GBM) Algoritması 'nın doğru tahmin ettiği örnek sayısının tüm örnek sayısına oranı %88,8 'dir. Şekil 4.6 'da dolandırıcılığa maruz kalmış kredi kartları turuncu renk ile, dolandırıcılık olmayan kartlar ise mavi renk ile gösterilmiştir. Bu bilgiye göre 4013 dolandırıcılık olmayan vakanın 3948 'inin dolandırıcılık olmadığı doğru tahmin edilmiştir. Dolandırıcılığa maruz kalan 5987 adet kartın ise 4939 tanesinin dolandırıcılık olduğu doğru tahmin edilmiştir.

**Tablo 0.3.** CatBoost Algoritması Karışıklık Matrisi

| **Doğruluk: %88,3** | **Gerçek 1** | **Gerçek 0** | **Toplam** | **Precision (Kesinlik)** |
|---|---|---|---|---|
| **Tahmin 1** | 4904 (TP)<br>Doğru Karar | 100 (FP)<br>Tip 1 Hatası | 5004 | %98 |
| **Tahmin 0** | 1062 (FN)<br>Tip 2 Hatası | 3934 (TN)<br>Doğru Karar | 4996 | %78,7 |
| **Toplam** | 5966 | 4034 | 10000 | |
| **Recall (Duyarlılık)** | %82,1 | %97,5 | | |

Tablo 4.3'e göre CatBoost Algoritması 'nın doğru tahmin ettiği örnek sayısının tüm örnek sayısına oranı %88,3 'dür. Şekil 4.7 'de dolandırıcılığa maruz kalmış kredi kartları turuncu renk ile, dolandırıcılık olmayan kartlar ise mavi renk ile gösterilmiştir.



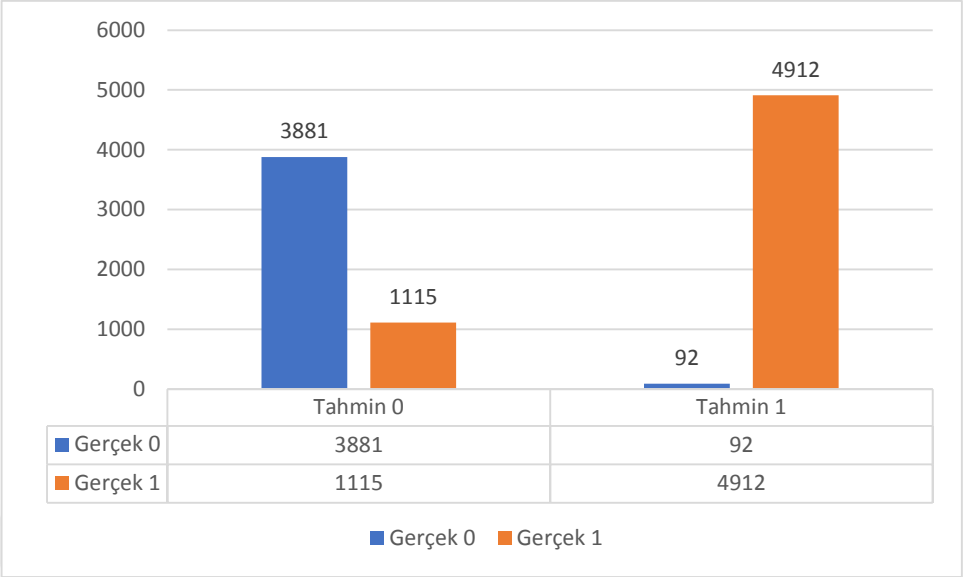

**Şekil 0.7.** CatBoost Algoritması Tahmin Sayıları

Bu bilgiye göre 4034 dolandırıcılık olmayan vakanın 3934 'ünün dolandırıcılık olmadığı doğru tahmin edilmiştir. Dolandırıcılığa maruz kalan 5966 adet kartın ise 4904 tanesinin dolandırıcılık olduğu doğru tahmin edilmiştir.

**Tablo 0.4.** Sınıflandırma ve Regresyon Ağacı (CART) Algoritması Karışıklık Matrisi

| **Doğruluk: %87,9** | **Gerçek 1** | **Gerçek 0** | **Toplam** | **Precision (Kesinlik)** |
|---|---|---|---|---|
| **Tahmin 1** | 4912 (TP) Doğru Karar | 92 (FP) Tip 1 Hatası | 5004 | %98,1 |
| **Tahmin 0** | 1115 (FN) Tip 2 Hatası | 3881 (TN) Doğru Karar | 4996 | %77,6 |
| **Toplam** | 6027 | 3973 | 10000 | |
| **Recall (Duyarlılık)** | %81,4 | %97,6 | | |

Tablo 4.4'e göre CatBoost Algoritması 'nın doğru tahmin ettiği örnek sayısının tüm örnek sayısına oranı %87,9 'dur. Şekil 4.8 'de dolandırıcılığa maruz kalmış kredi kartları turuncu renk ile, dolandırıcılık olmayan kartlar ise mavi renk ile gösterilmiştir. Bu bilgiye göre 3973 dolandırıcılık olmayan vakanın 3881 'inin dolandırıcılık olmadığı doğru tahmin edilmiştir. Dolandırıcılığa maruz kalan 6027 adet kartın ise 4912 tanesinin dolandırıcılık olduğu doğru tahmin edilmiştir.




**Şekil 0.8.** Sınıflandırma ve Regresyon Ağacı (CART) Algoritması Tahmin Sayıları

# SONUÇ VE ÖNERİLER

Özel ve kamu kuruluşlarının giderek daha da fazla karşılaştığı dolandırıcılık girişimlerinden dolayı, açtığı milyarlarca dolar mali zararlarda düşünüldüğünde insandan daha hızlı dolandırıcılık tespit sistemlerinin önemi artmaktadır. İnternet üzerinde artık olmazsa olmaz E-Ticaret kullanımının büyümesiyle, kredi kartı kullanılan çevrimiçi alışveriş ödemeleri de büyük ölçüde artmıştır. Bunun sonucunda ise internet üzerinde kullanılan kredi kartlarının dolandırıcılığında da büyüme meydana gelmiştir. Önemi artan otomatik tespit sistemleri, bankaların uğrayacağı zararları minimum seviyesine indirmek için artık zorunlu bir hal almıştır. Sürekli artış gösteren veri göz önüne alındığı zaman insan yeteneğini aşmakta ve tespit için yetersiz kalmaktadır. Bu yüzden dolandırıcılık işleminin hızlı ve doğru tespit edilmesi için otomatik tespit sistemlerine ihtiyaç vardır. Otomatik tespit sistemleri geliştirmek için Makine Öğrenimi (ML) tekniklerinden faydalanılmıştır. Bu çalışmada bankalara ait olan kredi kartı işlemlerini içeren veri setleri gizlilik nedeni ile paylaşılmadığı için dolandırıcılık işlemleri içeren veri seti bir simülasyon geliştirilerek üretilmiştir. İlk olarak similasyonun çalışmasını deneme amacıyla 117 adet veri üretilmiştir. Ardından 1754155 adet veri üretilmiştir. Üretilen veri setinde 14681 adet dolandırıcılık işlemi olan veri mevcuttur. Python programlama dili ile çeşitli sınıflandırma algoritmaları kullanılarak veri seti eğitilip test edilmiştir. Algoritmaların alışveriş işleminin dolandırıcılık olup olmadığını tespit eden kalıpları daha iyi bir şekilde anlayabilmesi için dolandırıcılık ve normal işlemlere ait kayıt sayısının birbirlerine yakın sayıda veri seti olması faydalı olabilir. Bu yüzden 10.000 veri içeren bir set oluşturulmuş ve bu sette 4996 adet dolandırıcılık işlemi ve 5004 adet normal alışveriş işlemi mevcuttur. Veri sayısını 10000 birime düşürdüğümüzde algoritmaların hız faktörleri artmıştır. Algoritmalar, yapılan işlemlerin dolandırıcılık mı normal bir işlem mi olduğunu tespit etmeye çalışırken 1 veya %99,9 gibi hatalı sonuçlar üretmiştir. Bu yüzden dolandırıcılık işlemi %49,96, normal bir işlem ise %50,04 yani birbirlerine yakın olarak alınmıştır. Üretilen veri seti toplam 24 farklı sınıflandırma algoritması ile eğitilip test edilmiştir. Eğitim için verilerin %80 'i kullanılmış, test işlemi için ise geriye kalan %20 kısmı kullanılmıştır. Eğitilen modellerin içerisinden en iyi sonuçları veren algoritmaların rastgele 3 tanesi açıklanmış, rastgele seçilen 3 tanesinin ise karışıklık matrisi gösterilmiştir. Üretilen veri setinde dolandırıcılık mı değil mi, dolandırıcılık işlemi var mı yok mu, dolandırıcılıksa lojik 1 değilse 0, 0 mı 1 mi gibi sorular sorabildiğimiz için regresyon değil sınıflandırma algoritmaları kullanılmıştır. Bu çalışmada cross-validation denilen çapraz doğrulama yöntemi de denenmiştir. Veriler gruplara ayrılmış ve test edilmiştir. Veri sayımızı 10000 'e düşürmemize rağmen hız faktörü negatif yönde değişim gösterirken, doğruluk oranlarında pozitif yönde dikkate alınacak bir değişim olmamıştır. Bu yüzden normal olarak verinin %80 'i eğitim, %20 'si test işlemi olarak alınmıştır.

Geliştirilen sistem, kredi kartı ile yapılan alışveriş işlemlerinin dolandırıcılık mı normal alışveriş işlemi mi olduğunu tespit ederek banka ve kredi kartı kullanıcılarına daha güvenli bir

kullanım sağlayacak bir sistem olabilmektedir. Böylece kredi kartının daha fazla maddi zararlara yol açmadan blokesi sağlanabilecektir.

Adli bilimler alanında da makine öğrenmesi yöntemlerinin kullanılabilmesi, bu alanda yapılacak diğer çalışmaların da olabileceğini göstermiştir. Bu çalışma siber güvenlik ve adli bilişim alanında makine öğrenmesi ile geliştirilen tespit sistemlerinin olabileceğini göstermiştir.

Veri üretmek için geliştirilen simülasyon, gizlilik ve veri güvenliği nedeniyle ulaşılamayan veriler için gerçeğe yakın sonuçlar üretebilecek bir simülasyon geliştirilebileceğini makine öğrenmesi ile çalışma yürütebilecek kişilere göstermiştir. Bu çalışma, insanın yeteneklerini aşan ve çok zaman kaybetmesine rağmen doğru sonuca veri fazlalığından dolayı varamayacağını göz önüne alarak adli bilimler alanında kullanılacak otomatik tespit sistemlerinin önemini göstermiştir. Bu çalışmada en yüksek doğruluk (accuarcy) oranını Hafif Gradyan Artırma Makinesi (LGBM) %89 'luk bir doğruluk oranı ile en başarılı algoritma olmuştur.

# KAYNAKÇA


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

Adalı, E. (2016). *Bilgisayar ve Bilgi Güvenliği*, İstanbul: İTÜ Yayınları.

Akdağlı, E. (2021). *Pekiştirmeli Öğrenme (Reinforcement Learning) Nedir?,* Erişim tarihi: 12.06.2022, https://www.muhendisbeyinler.net/pekistirmeli-ogrenme-reinforcement-learning-nedir/

Alkan, M. (2012). *Siber Güvenlik ve Siber Savaşlar*. Ankara: Siber Güvenlik Siber Savaşlar TBMM İnternet Komisyonu.

Andress, J. ve Winterfeld, S. (2013). Cyber warfare: techniques, tactics and tools for security practitioners, Amsterdam: Elsevier.

Aslan, E.D. *Yapay Zeka ve Türleri*, Erişim Tarihi: 18 Ağustos 2022, https://bilginc.com/tr/blog/yapay-zek-ve-turleri-5422/

Ayvaz, T., *Firewall Nedir?*, Erişim tarihi: 07 Aralık 2022, https://www.mediaclick.com.tr/tr/blog/firewall-nedir.

Barr, A. ve Feigenbaum, E. (1981). The Handbook of Artificial Intelligence, *3. Los Altos: William Kaufmann INC,* Cilt 1.

Bayık, F. (2019). Aristoteles ve Descartes Bağlamında Akıl ve Zekâ Kavramlarının Farkları. Kaygı. Bursa Uludağ Üniversitesi Fen-Edebiyat Fakültesi Felsefe Dergisi, 18 (1), 172-187

Bayindir, R., Colak, I., Fulli, G., ve Demirtas, K. (2016). Smart Grid Technologies And Applications. Renewable and Sustainable Energy Reviews, C. 66, 499-516.

Bengio, Y., Simard, P. ve Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks*, 5 (2): 157- 166.

Ben-Haim, Y., ve Tom-Tov, E. ,2010. A Streaming Parallel Decision Tree Algorithm Elad Tom-Tov. *Journal of Machine Learning Research*,11, :849– 872.

Beydogan, T. A. ve Canbay, C. (2008). Providing Cyber Security and Protection of Critical Iinformation and Infrastructures. *International Telecommunications Society17th Biennial ConferenceMontreal (ITS)*.

Bıçakçı, S. ve Ergun, D. (2015). Türkiye'de Siber Güvenlik, *Edam Siber Güvenlik Kağıtları Serisi*, (1).

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

Bilgisayar Sistemleri. *Virüs Nedir?*, Erişim Tarihi: 14 Ocak 2022, https://bilgisayarsistemleri.net

Breiman, L. (2001), Random forests. *Machine learning,* 45(1), 5-32.

Chapman, P., Clinton, J., Kerber, R., Khabaza, T., Reinartz, T., Shearer, C. ve Wirth, R. (2000), *CRISP-DM 1.0: Step-by-step data mining guide*. Germany: *SPSS inc*.

Check Point, (2020), *Check Point Güvenlik Raporu 2020*, Erişim tarihi: 08 Ekim 2022, research.checkpoint.com.

City of Vancouver, (2016). *Internal audit summary report*, Erişim tarihi: 01 Şubat 2022, https://vancouver.ca/your-government/internal-audit-reports.aspx.

Computer Repair Slogan, (2016), *Different Types of Malware*, Erişim tarihi: 9 Mart 2022, https://www.computerrepairslogan.com.au/different-types-of-malware/.

Coşar, M (2019). *Siber Güvenlik*, (Ders Konu Slaytı), Çorum: Hitit Üniversitesi Yüksek Lisans Enstitüsü.

Cyware (2019), *Cardinal Rat the Remote Access Trojan That Targets Fintech Companies*, Erişim tarihi: 9 Mart 2022, https://cyware.com/news/cardinal-rat-the-remote-access-trojan-that-targets-fintech-companies-4c546fe7.

Das, S., Kim, A., Tingle, Z. ve Nippert-Eng, C. (2019). All About Phishing: Exploring User Research

Garrido, A.P. (2016), *What is the difference between Bagging and Boosting?,* Erişim tarihi: 15.06.2022, https://quantdare.com/what-is-the-difference-between-bagging-and-boosting/

Goodrich, M. ve Tamassio, R. (2010). *Introduction To Computer Security*. Londra: Pearson.

Haykin, S. (1999). *Neural Networks and Learning Machines.* New Jersey: Prentice Hall.

Herman, M. (2022), *Cyber Security vs Information Security*, Erişim tarihi: 16 Aralık 2022, https://www.uscybersecurity.com/blogs/cybersecurity-information-security-difference.

Jampen, D., Gür, G., Sutter, T. ve Tellenbach, B. (2020), *Example of An Email Based Phishing Attack*, Human Centric Computing and Information Sciences, (10)33, 1-41.

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

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

Kaspersky, *Siber Güvenlik Nedir?*, Erişim tarihi: 24 Ocak 2022. https://www.kaspersky.com.tr/resource-center/definitions/what-is-cyber-security

Kaur, C. J., Bhandari, A., ve Behal, S. (2019). *Distributed Denial of Service Attacks: A Threat or Challenge. New Review of Information Networking,* 24(1), 31-103.

Keleş A. ve Ocak, R. (2007), *Öğrenme-Öğretme Sürecinde Yapay Zekâ Ve Web Tabanli Zeki Öğretim Sistemi Tasarimi ve Matematik Öğretiminde Bir Uygulama,* (Doktora Tezi), Erzurum: Atatürk Üniversitesi Fen Bilimler Enstitüsü.

Keycdn, (2018), *DDoS Attack*, Erişim tarihi: 24 Ocak 2022, https://www.keycdn.com/support/ddos-attack.

Kocamaz, A.F. (2012). *Makine Öğrenmesi Tabanlı Bir Uzman Tasarımı,* (Doktora Tezi), Edirne: Trakya Üniversitesi Fen Bilimler Enstitüsü.

Loon, R.V. (2014), Machine Learning Explained: Understanding Supervised, Unsupervised, and Reinforcement Learning.

Netrusion, *EventTracker Enterprise and the Cyber Kill Chain*, Erişim tarihi: 15 Şubat 2022, https://www.netsurion.com/articles/eventtracker-enterprise-and-the-cyber-kill-chain.

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

Öztemel, E. (2003). *Yapay Sinir Ağları.* İstanbul: Papatya Yayıncılık.

Pingree, L., MacDonald, N., and Firstbrook, P. (2015). Best Practices for Mitigating Advanced Persistent Threats. Gartner.

Pirim, H. (2006). Yapay Zeka. *Journal of Yasar University*. 1 (1), 81-93.

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

Sağıroğlu, Ş., Alkan, M. (2018). *Siber Güvenlik ve Savunma Farkındalık ve Caydırıcılık*, Ankara: BGD Yayınları.

Samonas, S., ve Coss, D. (2014). The Cia Strikes Back: Redefining Confidentiality, Integrity and Availability In Security. *Journal of Information System Security*, 10(3).

Security Affairs, *Example of Banking Trojan Attack*, Erişim tarihi: 9 Mart 2022, https://securityaffairs.com/?s=Example+of+Banking+Trojan+Attack

Shewan, D. (2017). *10 Companies Using Machine Learning in Cool Ways*, Erişim tarihi: 04.06.2022, https://www.wordstream.com/blog/ws/2017/07/28/machine-learning-applications

Singh, A., Vaish, A., and Keserwani P. K. (2014). *Information Security: Components and Techniques*. Allahabad, India: Indian Institute of Information Department Technology,

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

Sucu, İ. ve Ataman, E., (2020). Dijital Evrenin Yeni Dünyası Olarak Yapay Zeka ve Her Filmi Üzerine bir Çalışma. *Yeni Medya Elektronik Dergi*.4 (1), ss. 40-52.

Technopat, (2021), *svchost.exe Çok Fazla CPU ve Ram Tüketimi*, Erişim tarihi: 16 Şubat 2022, https://www.technopat.net/sosyal/konu/svchost-exe-yueksek-ram-kullanimi.1506238/

The Tartan, (2010), *How Things Work: Computer Worms*, Erişim tarihi: 16 Şubat 2022, http://thetartan.org/2010/2/15/scitech/computerworms.

Through a Systematic Literature Review, *ArXiv*, abs/1908.05897.

Topçuoğlu, A. (2001). Yapay Zeka, Bilim ve Teknik Dergisi, 39.

Uslu, A. (2021). *Siber Saldırı Nedir?*, Erişim Tarihi: 14 Ocak 2022, https://www.niobehosting.com/blog/siber-saldiri-nedir/

Uzun, E. (2007). *İnternet Tabanlı Bilgi Erişimi Destekli Bir Otomatik Öğrenme Sistemi*, (Doktora Tezi), Edirne: Trakya Üniversitesi Fen Bilimleri Enstitüsü.

Vmaraci, *Rootkit Nedir?*, Erişim tarihi: 9 Mart 2022, https://vmaraci.com.

Vmaraci, *Spyware Nedir?,* Erişim tarihi: 9 Mart 2022, https://vmaraci.com.

Wikipedia, Solucan (Virüs), Erişim tarihi: 16 Şubat 2022, https://tr.wikipedia.org/wiki/Solucan_(vir%C3%BCs)

Xu, E. ve Guo, G. (2018), *Windows, Android Users Targeted by Maikspy Spyware*", Erişim tarihi: 9 Mart 2022, https://www.trendmicro.com/en_us/research/18/e/maikspy-spyware-poses-as-adult-game-targets-windows-and-android-users.html

Yan, T., Deng, Q., Qu, B., ve Zhang, Z. (2021), *Palo Alto Networks Discloses New Attack Surface Targeting Microsoft IIS and SQL Server at Black Hat Asia 2021*, Erişim tarihi: 14 Şubat 2022, https://unit42.paloaltonetworks.com/iis-and-sql-server/.