ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

DETECTING MALICIOUS ACTIVITY INSIDE OF THE NETWORK

M.Sc. THESIS

Ayşenur KUMBASAR

Applied Informatics Department

Cybersecurity Engineering and Cryptography

DECEMBER 2023

ISTANBUL TECHNICAL UNIVERSITY ★ GRADUATE SCHOOL

DETECTING MALICIOUS ACTIVITY INSIDE OF THE NETWORK

M.Sc. THESIS

Ayşenur KUMBASAR
(707201002)

Applied Informatics Department

Cybersecurity Engineering and Cryptography

Thesis Advisor: Prof. Dr. Enver OZDEMIR

DECEMBER 2023

ISTANBUL TEKNİK ÜNİVERSİTESİ ★ LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

AĞ İÇERİSİNDEKİ KÖTÜ NİYETLİ AKTİVİTELERİN TESPİTİ

YÜKSEK LİSANS TEZİ

Ayşenur KUMBASAR
(707201002)

Bilişim Uygulamaları Anabilim Dalı

Bilgi Güvenliği Mühendlisliği ve Kriptografi

Tez Danışmanı: Prof. Dr. Enver ÖZDEMİR

ARALIK 2023

Ayşenur Kumbasar, a M.Sc. student of İTU Graduate School student ID 707201002, successfully defended the thesis/dissertation entitled "DETECTING MALICIOUS ACTIVITY INSIDE OF THE NETWORK", which she prepared after fulfilling the requirements specified in the associated legislations, before the jury whose signatures are below.

Thesis Advisor :	**Prof. Dr. Enver ÖZDEMİR** İstanbul Technical University	.....
Jury Members :	**Dr. Ogr. Üyesi Sefer BADAY** Istanbul Technical University	.....
	Dr. Ogr. Üyesi Elif Segah ÖZTAŞ Karamanoğlu Mehmetbey Üniversitesi University	.....

Date of Submission : 15 December 2023
Date of Defense : 20 December 2023

To my family,

FOREWORD

I would like to express my deep appreciation and thanks for my advisor.

I would like to thank my beloved family who always supported me throughout my education life.

December 2023

Ayşenur KUMBASAR
(MSc. Student)

TABLE OF CONTENTS

Page

ABBREVIATIONS

AUC **:** Area Under the Curve
CERT **:** Community Emergency Response Team
CNN **:** Convolutional Neural Network
DBN **:** Deep Belief Network
EUROSTAT **:** European Statistical Office
FP **:** False Positive
FN **:** False Negative
GCN **:** Graph Convolutional Networks
GridSearch CV: Grid Search Cross-Validation
IT **:** Information Technology
IDS **:** Intrusion Detection System
IPS **:** Intrusion Prevention System
KNN **:** K Nearest Neighbour
MAC: **:** Message Authentication Codes
ML **:** Machine Learning
PWC : PricewaterhouseCoopers
RBF **:** Radial Basis Function
ROC **:** Receiver Operating Characteristic
SEI **:** Software Engineering Institute
SSL **:** Secure Socket Layer
SVM **:** Support Vector Machines
TLS **:** Transport Layer Security
TN **:** True Posivite
TP **:** True Negative
US **:** United States

LIST OF TABLES

Page

LIST OF FIGURES

Page

DETECTING MALICIOUS ACTIVITY INSIDE OF THE NETWORK

SUMMARY

In today's world with the global development and digitalization, applications and services used in banking and finance sectors, as in all sectors, have started to adapt to the online world quickly. The increase in the rate of transition to the Internet environment shows that the issue of security is becoming more and more important and serious for banks and customers.

Companies serving in the financial and banking sectors are an attractive target for cyber attackers in terms of damage to the target system and data obtained by attackers. The protection of information systems containing important and sensitive business and customer information, such as databases, servers, computers, networks used, is of high importance. In the same way, providing a secure and robust online communication environment in the services provided to customers and ensuring that data is transmitted in reliable environments is one of the most important elements in the banking sector

Banks are also making major investments in security systems to ensure secure communication and the protection of personal and business information and documents as a precaution against this increasing number of cyber attacks. With these systems, they have the potential to prevent such attacks by detecting and responding to abnormal and unauthorized activities.

However, research shows that the majority of cyber attacks are carried out by insiders. Most security products in use focus on external threats. However, if the attacker is a person working within the organization, these systems may be insufficient to detect such activities. The inside attacker has legitimate access privileges to sensitive data, systems, networks that outsiders do not have. It is difficult to predict and prevent as the malicious user inside follows legitimate paths and methods. Since the systems have detailed information about the internal organization such as the corporate network, they can misuse sensitive and confidential data and cause irreversible damage to the organizations by creating great losses. Therefore, it can be said that the cost of damage caused by internal threat is much higher than external threat.

This study focuses on detecting insider threats by monitoring users with a behavioural focus. By examining normal user behaviour and malicious user behaviour with SVM, KNN and Random Forest algorithms, it is aimed to detect internal threats and help minimize the damage that can be done to the institution with preventive controls that will come with it.

AĞ İÇERİSİNDEKİ KÖTÜ NİYETLİ AKTİVİTELERİN TESPİTİ

ÖZET

Küresel gelişim ve dijitalleşme ile birlikte günümüz dünyasında tüm sektörlerde olduğu gibi bankacılık ve finans sektörlerinde de kullanılan uygulama ve hizmetler hızla online dünyaya adapte olmaya başlamıştır. İnternet ortamına geçiş hızındaki artış, güvenlik konusunun bankalar ve müşteriler için giderek daha önemli ve ciddi hale geldiğini göstermektedir.

Finans ve bankacılık sektörlerinde hizmet veren şirketler, hedef sisteme verdiği zarar ve saldırganların elde ettiği veriler açısından siber saldırganlar için cazip bir hedef konumundadır. Kullanılan veritabanları, sunucular, bilgisayarlar, ağlar gibi önemli ve hassas iş ve müşteri bilgilerini içeren bilgi sistemlerinin korunması büyük önem taşımaktadır. Aynı şekilde müşterilere sunulan hizmetlerde güvenli ve sağlam bir online iletişim ortamının sağlanması ve verilerin güvenilir ortamlarda iletilmesinin sağlanması bankacılık sektöründeki en önemli unsurlardan biridir.

Siber tehditlerden kaynaklanan finansal etki, günümüzde şirketler için büyük bir endişe kaynağıdır. Bir veri ihlali sebebi ile, bütün ticari sırlar, hassas veriler ifşa edilebilir. Yaşanan ihlal, kuruluşların itibarının zarar görmesine, müşteri kayıplarına ve yeni müşteri kazanmada zorluklara sebep olabilir.

Artan siber saldırılara önlem olarak bankalar, güvenli iletişim ve kişisel ve ticari bilgi ve belgelerin korunmasını sağlamak için güvenlik sistemlerine de büyük yatırımlar yapıyor. Bir kuruluşun veya işletmenin performansını büyük ölçüde etkileyebilecek siber saldırılara karşı korunmak için bir güvenlik planına sahip olmaları büyük önem taşımaktadır. Bu sistemler ile anormal ve yetkisiz faaliyetleri tespit edip müdahale ederek bu tür saldırıları önleme potansiyeline sahiptirler.

Kritik altyapıların güvenliğini sağlamak için yapılan yatırımlar, genel olarak izinsiz girişi önleme, siber saldırıyı tespit etme, saldırganın aktiviteleri sonucunda karşılaşılabilecek fiziksel etkileri hafifletme gibi önemli unsurları sağlamak amacındadır. Kuruluşlar, ihtiyaçlarına ve bu ihtiyaçların önem derecelerine göre ulusal

standartlar ve kurum politikalarına uygun olarak en uygun yatırımları gerçekleştirmektedir. Böylelikle düzenli yapılan kontroller sonucu bir güvenlik ihlali veya dışarıdan izinsiz giriş gibi durumların tespiti ve önlenmesi erkenden sağlanarak, ihlal olması durumunda bile yaşanılan kayıpların en aza indirilmesi sağlanmış olacaktır.

Ancak kullanılan güvenlik ürünlerinin çoğu dış tehditlere odaklanır ve içerden kaynaklanan bir tehditi tespit etmede yetenekli değillerdir. Araştırmalar, siber saldırıların büyük çoğunluğunun içeriden kişiler tarafından gerçekleştirildiğini gösteriyor. Ancak saldırgan kurum içinde çalışan bir kişi ise bu sistemler bu tür faaliyetleri tespit etmekte yetersiz kalabilmektedir.

İçerideki saldırgan, dışarıdakilerin sahip olmadığı hassas verilere, sistemlere ve ağlara meşru erişim ayrıcalıklarına sahiptir. İçerideki kötü niyetli kullanıcı zararlı aktivitesini gerçekleştirebilmek için meşru yolları ve yöntemleri izlediği için tahmin edilmesi ve engellenmesi zordur. Kurum içerisinde çalışan, gerçek kimlik bilgileriyle oturum açan yetkili herhangi bir kullanıcının çalıştığı kurumun çıkarları doğrultusunda hareket edip etmediğini anlamak mümkün olamamaktadır. Ayrıca içerden gelen tehditler, genellikle eylem gerçekleştikten sonra tespit edilebildiğinden, kritik verilerin kurtarılması da çok zor olabilmektedir.

İçeriden gelen saldırganlar, sistemler, kurum ağı gibi kurum içi organizasyon hakkında detaylı bilgilere sahip olduklarından hassas ve gizli verileri suistimal edebilmekte ve büyük kayıplar yaratarak kurumlara geri dönülmez zararlar verebilmektedir. Dolayısıyla iç tehdidin neden olduğu zararın maliyetinin dış tehditten çok daha yüksek olduğu söylenebilir.

Dış tehditlere odaklanan güvenlik duvarı, izinsiz giriş tespit ve önleme sistemleri gibi geleneksel güvenlik önlemlerinde, içerideki tehditleri tespit etmek veya önlemek çok başarılı sonuçlar getirememektedir. Örneğin bir saldırgan, sistemde oturum açmak için yetkili bir kullanıcının bilgilerini başarılı bir şekilde kullanırsa, bu anormal davranış güvenlik mekanizmalarında tespit edilemeyebilir. Makine öğrenmesi araçları kullanılarak, içeriden gelen potansiyel tehditlerin algılanması, kullanıcı davranışlarının analiz edilmesi, ve bir anomali tespit edildiğinde uyarılması sağlanabilir.

İçerideki saldırganlar, kötü niyetli aktivitelerini gerçekleştirebilmek için verileri kurcalama, manipule etme gibi eylemler gerçekleştirebilir. Kurum içerisinde iyi bildiği güvenlik ihlallerini kullanarak ve güvenlik önlemlerini kolaylıkla atlatarak erişim yetkilerini arttırabilir. Kalıcı erişim için bir arkakapı kurabilir. Erişim kazandığı veya halihazırda bulunan erişimini kullanarak kritik verilerin dışarıya aktarımını gerçekleştirmeye çalışabilir. İçeriden gelen tehditleri tespit edebilmek için farklı göstergeler izlenebilir. Bu göstergeler içerisinde alışılmadık zamanlarda sistem üzerinde etkinlik göstermesi olabilir. Sistem üzerinden veri çıkarmaya çalışması sonucu trafik üzerinde gözlemlenebilecek hacim artışı veya kullanıcının alışılmadık kaynaklara erişim sağlaması izlenerek kötü niyetli aktiviteler tespit edilebilir. Fakat bu gibi göstergeler daha çok saldırganların zararlı aktivitelerini gerçekleştirme adımlarında tespitine yardımcı olabilir.

Kötü niyetli eylemleri gerçekleştirebilmek için sistem üzerinde normal bir kullanıcının geçirdiği süreden daha fazla bir süre harcaması gerekmektedir. Bu yüzden makine öğrenme araçları ile kullanıcıların bir sistem üzerindeki oturum süreleri karşılaştırılarak elde edilen sonuç kötü niyetli kullanıcıların eylemlerini gerçekleştiremeden hazırlık aşamasında tespitinde önemli bir sonuç sağlayabilir.

Bu çalışma, kullanıcıları davranış odaklı izleyerek içeriden gelen tehditleri tespit etmeye odaklanmaktadır. SVM, Random Forest ve KNN makine öğrenme algoritmaları ile normal kullanıcı davranışı ve kötü amaçlı kullanıcı davranışı incelenerek, iç tehditlerin tespit edilmesi ve beraberinde gelecek önleyici kontrollerle kuruma verilebilecek zararın en aza indirilmesine yardımcı olunması amaçlanmaktadır.

Çalışmamızda makine öğrenmesi algoritmasında Marnegie Mellon Üniversitesi tarafından yayınlanan CERT Insider Threat veritabanının 6.2 sürümünü kullandık. CERT veri seti, içeriden gelen tehditleri tespit etmek için kullanılan çerçeveleri test etmek için yapay olarak geliştirilmiş sentetik bir veri setidir. Verisetinden elde ettiğimiz sistem üzerinde kullanıcıların oturum açma ve oturum kapama sürelerini farklı SVM, Random Forest ve KNN algoritmaları ile hesaplayarak doğruluk skorlarını karşılaştırdık. Makine öğrenimindeki en etkili veri sınıflandırma yöntemlerinden biri olan Random Forest algoritmasının, kullandığımız verisetindeki anormal aktiviteleri sınıflandırma konusunda daha iyi bir performansa sahip olduğunu tespit ettik.

1. INTRODUCTION

With the development of the Internet, the banking sector, like many other sectors, has started to transfer its services and applications mostly online. In proportion to the rapid development of technology, cyber attackers and the tactics and techniques they use are also developing. One of the most critical elements, especially for the banking and finance sector, is the confidentiality of sensitive data and communication. The fact that these data are compromised can reduce a bank to very bad levels, both financially and in terms of reputation. Although the first thing that comes to mind in terms of security and cyber attacks is external threats, the statistics obtained from different studies show that insider threats come first among the threats that can cause the greatest harm to an institution. And the number of insider threats is substantial. However, most existing security mechanisms focus on detecting external threats. In this study, the success of the SVM machine learning algorithm for insider threat detection was measured.

In the second part, basic information about online banking and its development over the years are given. Security, which is one of the most important concepts especially for the banking sector, is one of the main problems. With the development of the internet over the years, the banking and finance sectors have also moved to online environments and undoubtedly brought security problems with it.

In the third chapter, the concept of e-commerce, which is an important application and structure for the banking sector, is mentioned. The components of the e-commerce structure and how the process between the bank and the relevant institution work are explained.

In the fourth chapter, secure communication, which is an important and critical factor for e-commerce, is explained. Important information security concepts that must be provided for secure communication have been defined. In addition, the basic security mechanisms used to realize these information security elements are explained.

In the fifth chapter, the security mechanisms used to ensure the security of information and communication in institutions and the investments made in these mechanisms are mentioned.

In the sixth chapter, the concept of insider threat, which is one of our main topics, is explained. While mentioning the importance of the subject, the data breach notifications made by the employees within the institution that took place in the banks serving in Turkey were also mentioned.

In the seventh chapter, the basic concepts of machine learning are given. Machine learning algorithms are briefly mentioned and the details of the algorithms used in this study are given.

In the eighth chapter, the model and dataset used in this study for insider threat detection are explained in detail and the results are given.

1.1 Literature Review

[1] It is recommended to use the Deep Belief Network (DBN) model for internal threat detection. In the study conducted with the CERT database, more successful results were obtained in catching insider threats compared to the classical methods.

[2] For insider threat detection, a Graph Convolutional Networks (GCN) based model has been designed using Users' properties. With the model obtained, it was stated that better results were obtained in random forest, logistic regression, SVM, CNN machine learning algorithms.

[3] this study is on insider threat detection using unsupervised learning algorithms. Synthetic data including psychometric scores of users were produced to be used in algorithms. It has been observed that performance is much better than previous approaches.

[4] In this study, user behavior-based detection method was used against existing statistical analysis insider threat detection approaches. According to the comparative results obtained, a simple, flexible and efficient method was obtained with this approach.

[5] In this study, a stream mining method is proposed to classify dynamic data streams of unlimited length. Results with high accuracy have been obtained for insider threat streams compared to traditional supervised learning algorithms.

2. ONLINE BANKING

The Internet has played an important role in changing how we interact with other people and how we do business today. The Internet has played an important role in changing how we interact with other people and how we do business today.

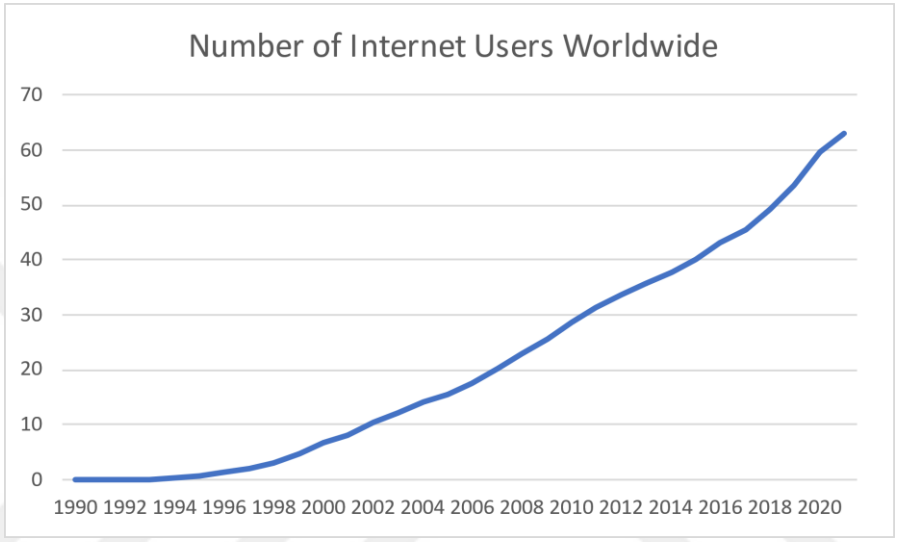


Figure 1: The Number of Global Internet Users

The number of global internet users is expected to increase continuously between 2023 and 2028, with a total of 860.7 million users (+16.14 percent) [6]. According to this estimate, in 2028, the number of users will reach 6.2 billion users for the fifth year in a row [6]. Figure 1 shows the increase in the number of internet users worldwide.

As a result of the Internet, electronic commerce has emerged, allowing businesses to interact more effectively with their customers and other companies inside and outside their industries.

One of the sectors that uses this new communication channel to reach its customers is the banking sector. The banking industries are one of the business lines that use these new communication media to provide value-added service and convenience to their customers.

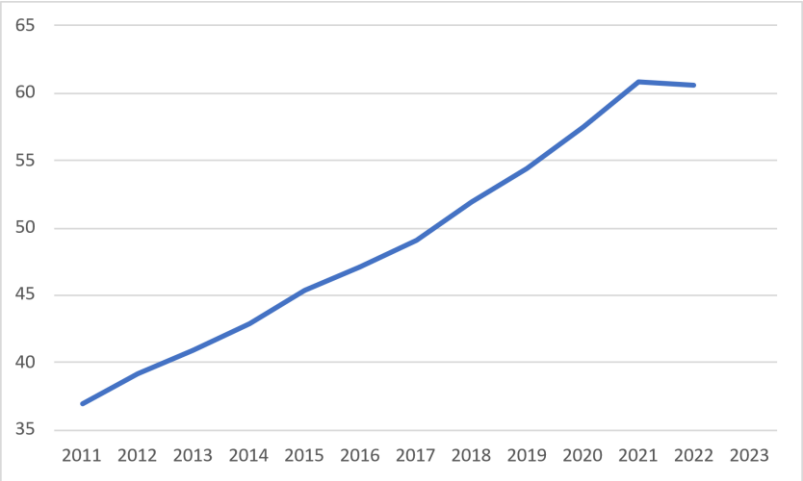


Figure 2: Individuals using the internet for internet banking

Figure 2 Individuals using the internet for internet banking

According to EUROSTAT, the rate of individuals using the internet for internet banking in December 2022 was 61.39% [7]. Figure 2 shows the increase in the number of internet banking users worldwide.

2.1 Security in Online Banking

The financial and banking sectors are among the primary targets of all kinds of crimes, including cybercrime.[8] The information stored in the bank and the main subject, money, are of interest to cybercriminals. To counter cyber threats, the bank must secure its communications and private data. In addition to the financial damage to be incurred by the bank, a possible breach will leave the image that the bank does not have sufficient infrastructure and competence to protect private data, so there will be much more material and moral damage than expected.

Financial institutions are responsible for 35% of all data breaches. 68% of companies in the industry report being hacked at some point. [8]

Listed below are examples of some of the key digital security incidents that have impacted financial firms in recent years: [9]

- In 2014, data of 20 million people was stolen from three Korean credit card companies (Kookmin Bank, Lotte Card and Nonghyup Bank). It was stated

that the stolen data included personal data such as identity numbers, addresses and credit card numbers.

- In 2014, JP Morgan Chase, the largest US retail bank, was the victim of an attack that compromised the data of 76 million people and some businesses in the US. The data included personal information such as names, addresses, phone numbers and e-mail addresses.
- In 2016, Tesco Bank was the victim of a cyberattack that affected 8,261 of 131,000 Tesco Bank personal current accounts. Some customers reported that they were unable to pay using their debit cards.
- In 2017, hackers stole the personal data of nearly 150 million people from the databases of Equifax, a consumer credit reporting agency.

3. E-COMMERCE

Electronic commerce is the purchase and sale of goods and services over the Internet. It started to be used in the early 1990s. Thanks to e-commerce, it allows all companies, large or small, to reach customers around the world and to spread their products and services over a wide area.

In general, three different methods of e-commerce are mentioned: [10]

M-commerce: Online transactions performed on mobile devices are known as m-commerce.

Corporate e-commerce: It is the purchase or sale of products from organizations to other organizations.

E-commerce in Social Media: Today, with the increase in the use of social media, it is now in our life in every field. As in many areas, in e-commerce, social media allows sellers to reach large audiences.

The e-commerce payment process steps are given in the Figure 3. [11] First, let's look at the main terms related to transactions:

Payment gateway: A payment gateway acts as a link between the e-commerce store and the Payment processor. It is an important component in the payment process. It transmits the information added to the e-commerce store to the Payment processor in an encrypted format. Authorization or refusal information is also transmitted back to the e-commerce site via the payment gateway.[12]

Payment processor: Payment processor is a financial services provider. It receives the information from the payment gateway and then verifies if it is correct. The trader checks if the customer has money in their account and then deposits this money into the merchant account.

Merchant account: It is an account opened at the bank by the e-commerce store. This is the account from which funds will be received after payment processing

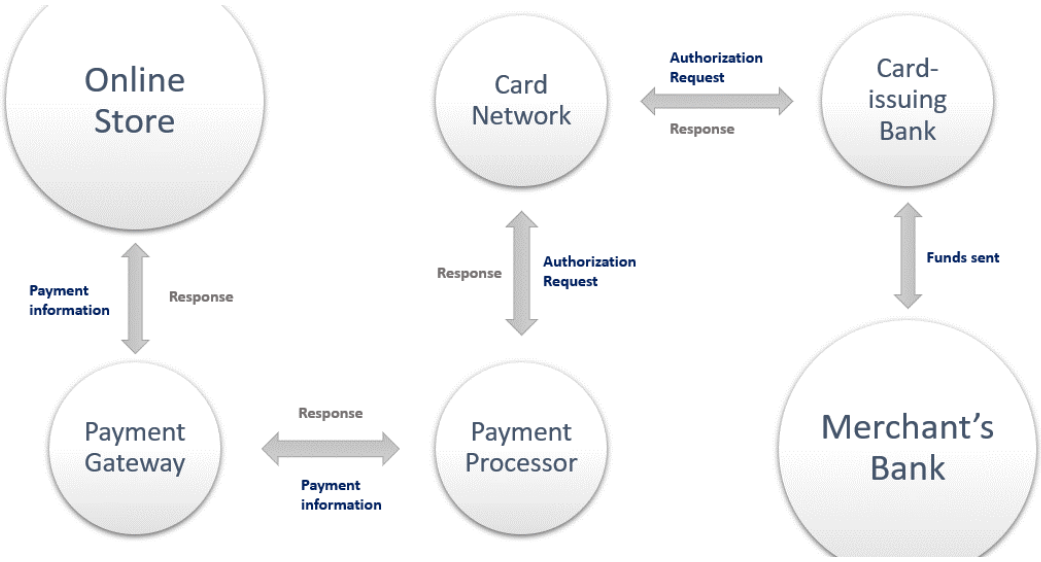


Figure 3: E-commerce Payment Processing

Payment process steps are listed below, respectively: [13]

1. The customer enters his credit card information on the payment page and when he initiates the payment transaction, the Payment gateway transmits the payment information to the Payment processor.
2. Payment processor sends card information to Card Network for verification.
3. After verification, Card Network requests authorization from the relevant bank to verify whether there are funds or limits available for payment and to ensure that the funds are received.
4. If the customer has sufficient balance, the bank sends an oany to the card network.
5. The card network forwards the bank confirmation to the Payment Processor and requests a money transfer from the customer bank to the Merchants bank.
6. Finally, the payment processor transfers the money to the bank account.

4. SECURE COMMUNICATION

Secure communication is when two parties communicate and a third party cannot eavesdrop on that communication. In today's digital world, secure data communication between two nodes is vital. Therefore, information security measures should be taken.

For a better understanding of the concept of Secure Communication, the basic security principles will be mentioned.

4.1 Basic Information Security Principles

In order to talk about the security of online communication and sensitive data, some basic elements must be provided. The following elements support each other. The fact that only a single element has occurred does not mean that security has been provided.

- **Confidentiality:** It is the inability to access information by unauthorized parties. In the communication between the two parties, the attacker can listen and read the messages between the sender and the receiver, which breaks the confidentiality of the communication. Encryption methods can be used in cryptography (symmetric/asymmetric) to achieve the goal of secrecy.
- **Integrity:** It is the protection of information against destruction or deletion by unauthorized persons. Changing the messages transmitted in inter-party communication by the attackers and sending them as if they were sent by the sender destroys the integrity of the transmitted message. Hash functions in cryptography are examples of methods used to ensure integrity.
- **Availability:** It means that the information is accessible and usable by authorized persons whenever it is needed. The attacker tries to prevent the continuity of the communication by making the communication channel between the two communicating parties unusable.

- **Authentication:** Authentication ensures that the message came from the person claimed in the message. An attacker can send a message to the receiver by impersonating the sender. In this case, if the recipient does not make a reliable authentication, they will communicate with the wrong people with the wrong messages. It can be achieved using a digital signature in cryptography.
- **Non-Repudiation:** It is the prevention of the data sender's denial of transferring the data or the receiver's receiving the data with the help of evidence. One way to achieve this in cryptography is through the use of digital signatures.
- **Accountability:** By following the actions of the users, in case of an incident, their transactions are examined and the person responsible is revealed. In this way, violation detection and prevention are supported. At the same time, undeniability is ensured.

Cryptographic algorithms that provide the specified basic security are used as components of security protocols to prevent malicious effects that communication may suffer.

4.2 Cryptography

The concept of cryptography can simply be defined as the science of cryptography. Cryptogari is a sub-concept of cryptology and it is called the methods used to ensure the confidentiality of information.

One of the areas where cryptography is most used is communication, but it is used in many areas such as banking transactions, shopping, and e-mails that we carry out over the internet today. Even Enigma, which was used for secret communication in World War II, is an important example.

From the earliest times to the present day, encryption methods are used to ensure a secure and confidential communication between the parties. Encryption methods transferred from encrypted messages written on ceramics to computer environments basically have the same logic.

The message to be transmitted to the other party is called plaintext. Making this plaintext unreadable except for legitimate parties communicating is called Encryption (Figure 4). The ciphertext obtained after the encryption process is decrypted with the Decryption method and the receiving party is provided to access the plaintext. Enctpytion and decryption operations are done with the help of a key.

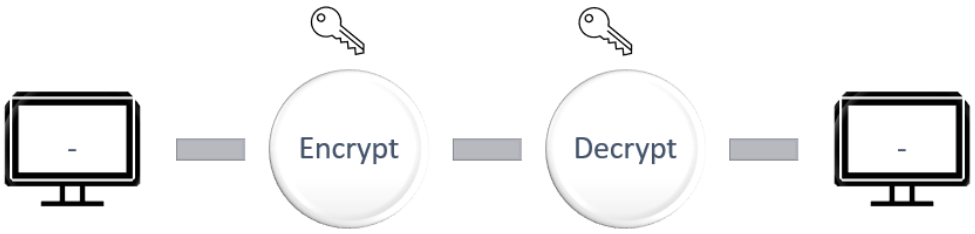


Figure 4: Encryption

According to the type of key used for encryption in cryptography, two different encryption algorithm categories can be mentioned.

4.2.1 Symmetric encryption algorithms

A common secret key is used for encryption and decryption in symmetric encryption algorithms as shown in the Figure 5. Before the sender and receiver start communication, they determine a common key and perform all encryption and decryption operations according to this key. The sending party encrypts the plaintext to be sent with the Secret Key known only between the parties and sends the resulting ciphertext to the other party. The receiving party, on the other hand, reaches Plaintext by decrypting the ciphertext with its public and secret key.

Figure 5: Symmetric Encryption

Encryption and decryption processes are fast. However, there are some difficulties encountered in symmetric encryption algorithms. First of all, key storage is difficult. Since n*(n-1)/2 keys must be stored in a system with “n” users, it may not be scalable. Also, reliable key distribution between parties is difficult. And it is difficult to achieve communication integrity and authentication securely.

4.2.2 Asymmetric encryption algorithms

In asymmetric encryption, different keys are used for encryption and decryption. Both communicating parties have a public key that they share with everyone, and a private key that they do not share with anyone. The encryption key is called the public key, and the decrypting key is called the private key.

Large prime numbers are used to obtain public and private keys. The private key consists of the prime numbers that make up the public key, while the public key can be accessed from the prime numbers, the numbers that make up the multipliers from the public key cannot be reached.

When the Sender wants to send a message (plaintext) to the other party, it encrypts the plaintext using the Receiver's Public key (Figure 6). Since this encrypted text can only be decrypted with the recipient's private key, the confidentiality of the transmitted message is ensured.

Figure 6: Asymmetric Encryption

With asymmetric encryption, key confidentiality and public key distribution, which are the most important problems of symmetric encryption, are solved. The disadvantage of asymmetric encryption is that it works slower than symmetric encryption.

Since the keys used in symmetric encryption remain confidential only between the parties, the identity of the message sender and receiver parties can be assured (it is thought that the key is not compromised). In asymmetric encryption, public keys should be known by everyone, so it should be ensured that the identity is correct. Signing method is used to verify that the message came from a certain person. It is sent by adding the signature part encrypted with the sender's private key to the sent message as it is explained in th e Figure 7. When the receiver decrypts this encrypted message with the sender's public key, he/she is sure of the identity of the sender, because the message encrypted with the private key can only be decrypted with the public key of that private key.

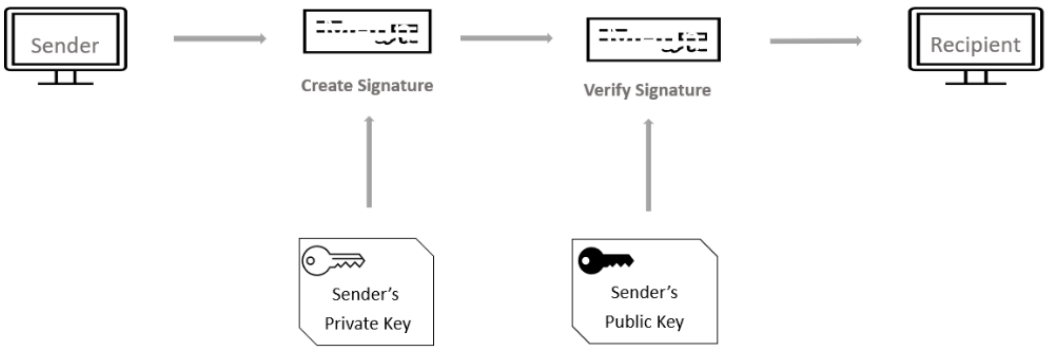


Figure 7: Public Key data encryption and decryption

4.2.3 Hash functions

The hash function is a mathematical algorithm that converts data into a fixed-length output. They are one-way functions. The original data cannot be obtained from the hash value. The same output is produced for every same input inserted into the function. However, even with the slightest change in the input, the value output from the function changes. Thanks to this feature of hash functions, the integrity of the data can be ensured.

For an ideal cryptographic hash function to be considered secure, it must have three properties: [17]

- **Collision Resistance:** Any two different inputs do not produce the same hash value as output.
- **Inverse image resistance:** It is the case that the original data cannot be found from the output value formed in the hash function.
- **Secondary inverse image tolerance:** It must be very difficult for two separate inputs to have the same hash value.

The storage of passwords can be given as an example of the areas where cryptographic hash functions are used the most. Passwords are stored in databases by calculating hash values. Thus, raw passwords cannot be obtained thanks to the "irreversible" feature of hash functions, and cost savings are achieved by reducing the size of the stored data.

4.2.3.1 Message authentication codes (MAC)

The difference of the MAC algorithm from other hash functions is the use of a key. It is used to verify the source and content of the transmitted message. A fixed length result (MAC) is obtained by processing the message to be sent and the symmetric key in the MAC algorithm by the sender. In the Figure 8, the process is visualized. Along with the output, the message is transmitted to the receiver. The receiver

calculates the received message with the same MAC algorithm and symmetric key. If the result calculated by the receiver and the result received from the sender are the same, it is confirmed that the message has been received unchanged.

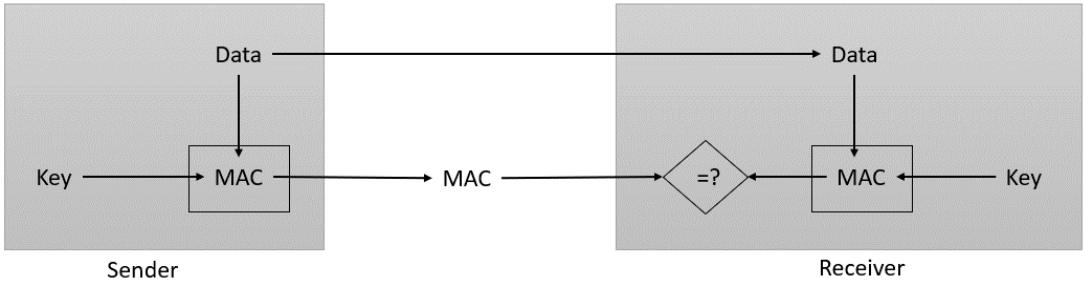


Figure 8: MAC Algorithm

4.3 Secure Socket Layer (SSL) / Transport Layer Security (TLS)

Different solution methods can be used to ensure communication security between the customer and the bank. One of these solutions, which can be considered as a standard, is SSL/TLS technology.

TLS is a technology used to secure communication by establishing an encrypted connection between client and server over a network. Today, SSL has left its place to TLS technology. But the term SSL is still widely used.

As it is explained in the Figure 9; first, the client establishes a connection with the server. The server sends the client the TLS certificate containing information about its identity. This certificate contains information such as the server's public key, domain name, and certificate authority. Thus, the client authenticates the server using this information. Then the client creates a unique session key that will be used to encrypt the communication with the server and sends this key to the server by encrypting it with the public key shared by the server. The server decrypts the password with its private key, thus providing encrypted communication between the server and the client.

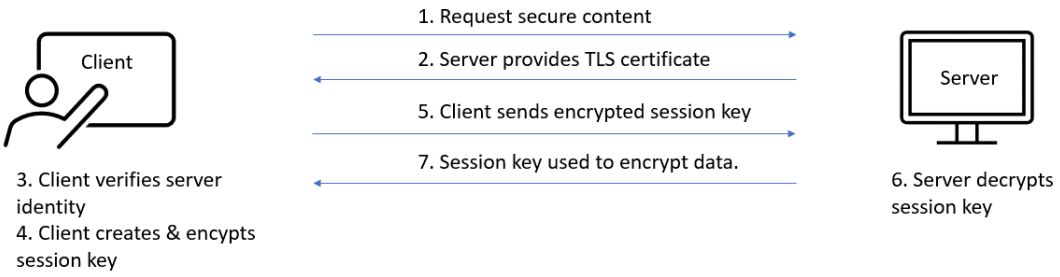


Figure 9: TLS Algorithm

5. INVESTMENTS TO ENSURE SECURITY

The security mechanism is the mechanism that detects, identifies, and prevents attacks against companies and systems and saves them from breaches.

Continuous investment in up-to-date security measures and technologies helps prevent heavy financial and information losses from cyber attacks.

Information security incidents are serious threats to business. Since the value of information is important, especially in the banking and finance sectors, there are many attacks and incidents targeting information systems, and they continue to increase. A sufficient amount of investment must be made in order for companies and organizations to achieve information security (Figure 10).

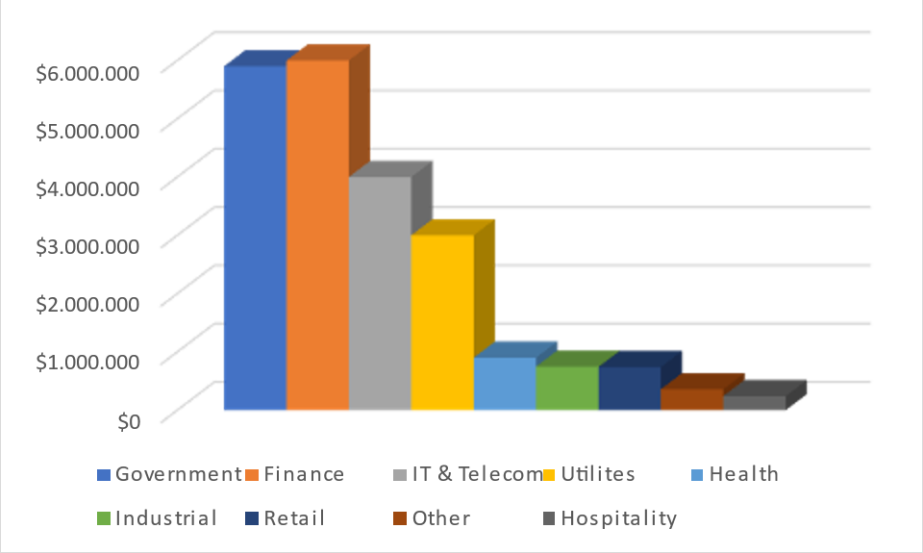


Figure 10: Security Budgets

It has been revealed that high investments are made in security, especially in the Public, IT and finance sectors. These investments should not be considered only as costs, they are a necessity for the continuity of the activities of the institutions. The expenditures made on these investments can actually provide a logical and both material and moral gain when considering the cost to be faced in the event of a security

breach. As can be seen in the study in the Figure 11, the investments made by thinking in this way and by being affected by the losses faced by other companies in data breaches are substantial.

Figure 11: Reasons for Wanting to Increase IT Security Budgets

To strengthen protection against external threats, high-quality security systems must be used and invested. In order to provide security on complex networks, communication with the cryptographic methods mentioned in the previous sections should ensure the confidentiality of data as well as physical protection of assets. An example of physical security is the protection of access to network resources by security personnel in certain and highly secure rooms with restricted access.

In addition to cryptographic methods, some security devices can be used to ensure data privacy and intrusion prevention. Firewalls, intrusion detection and prevention systems can be cited as examples. Firewall is responsible for providing security between two or network. Controls are provided with certain policies and rules defined. IDS detects and reports malicious activity coming over the network. IPS prevents these activities from taking place.

The methods and tools to be used in order to ensure security should be based on a specific plan.

In order to meet the security requirements, a specific plan, policy, and procedure should be developed first. It should be checked that these documents are kept up-to-date, that they cover every asset in the institution, and that they contain appropriate

security control mechanisms by controlling them periodically with both internal and external people.

All these security needs can be increased even more. In order to meet these needs, institutions need to make significant investments. Although this may seem unnecessary or excessive for some, sanctions and material and moral losses that may be encountered as a result of a future violation or cyber incident can cause much more damage.

6. INSIDER THREAT

It was mentioned in the previous sections that the main goal of information security is to reach the basic elements of confidentiality, accessibility and integrity. Threats that can endanger these elements can be divided into two: Outsider threat and insider threat. Outsider threat is when external people pose an outside risk to an organization's security. These attackers are attackers with no outside connection, mostly aiming financial gain.

Insider threat is a security threat originating from an individual within the organization and is one of the most important cyber threats today. PwC's 2018 Global Economic Crime and Fraud Survey reveals that 52% of all scams are committed by individuals within the organization. [21]

A great deal of research has been done and many products have been developed to protect these assets. However, most current security products are focused on detecting and preventing external threats and detecting malware. Research shows that insider threats are just as dangerous as external threats, perhaps even more.

While external threats will never go away, many are being countered. As technology advances and the threat from within cannot be ignored, it becomes much more pervasive and dangerous.

According to the 2020 Cost of Insider Threats Global Report by the Ponemon Institute, the number of incidents increased by 47% between 2018 and 2020, with the average cost of an incident increasing by 31% to $11.45 million. [22]

In insider threats, three categories can be mentioned according to the intention of the person. One is a threat that deliberately tries to harm the organization for purposes such as financial gain or revenge. It acts by using its own credentials and powers while carrying out its malicious actions. Another insider is a user responsible for unintentional activities such as carelessness or ignorance. Finally, it is the attacker who steals legitimate employee information and impersonates him.

Insider threats are one of the most difficult risks to detect and countermeasure. The inside threat actor dominates the network, applications, services, hardware and infrastructure within the organization. It is easier for them to avoid being detected because they know better where, how and where they can be detected. They can also

do much more damage than external threats, as they have full access to systems and are subject to fewer security restrictions than external threats.

6.1 Cyber Incidents Involving Insider Threat

Examples of data breaches by insiders are listed below:

- In 2018, a report by McAfee reported that a former CocaCola employee suffered a data breach when he was leaving, when workers moved their data to a hard drive. 8,000 CocaCola employees were affected by the reported violation. [23]
- As a result of the evaluation of the images made by 2 employees working in 2 different bank branches, made for the Credit Registration Bureau screens through the Bank system, in the data breach notification made by Garanti Bank,
- It has been determined that the employees have shared the information about the customers they obtained from the inquiries with 3rd parties. [24]
- In the statement made by Turkiye Is Bankasi, it was stated that a branch employee misused the person's office and forwarded e-mails containing Risk Center inquiries to non-bank users in exchange for money. [25]
- DenizBank A.Ş. Within the scope of the controls carried out by the Board of Inspectors, it was determined by the Bank employee who works as the Customer Transaction and Sales Responsible that he misused his authority, contrary to information security policies, by making more inquiries from the query screens containing individual credit information. [26]
- It has been determined by ING Bank that although an employee cannot query a service in accordance with the authorizations defined on the Finsoft system, which is an ING Bank application, he can directly access the TBB Risk Center website by overriding the authorization system with a method that will disable the authorization system. It has been determined that the person who caused the data leak made inquiries with his Identification Number and tax identification number information many times in 2018,

and that the data generated as a result of these queries was sent out of the Bank through electronic communication means. [27]

- In the statement made by Yapı ve Kredi Bankası, it was stated that an employee made inquiries and transferred this information to third parties by using his authority to query the intelligence records of the Banks Association Risk Center outside of his duty, which was defined for his duty. [27]

7. MACHINE LEARNING

Machine learning is the application of artificial intelligence that is programmed to increase performance or reduce error according to a criterion through algorithms and training data instead of explicit instructions of computer systems.

With the increasing digitalization in recent years, one of the important issues in the world is machine learning. As in many areas, it provides many advantages in the banking sector. Below are examples of real-life use:

- **Health:** detecting cancerous tissues, determining disease based on the symptoms experienced. With the techniques currently used, highly accurate diagnoses are detected.
- **Finance:** Making important financial decisions such as stock market investment
- **Sales and Marketing:** shopping suggestions for users, suggestions for sellers what product can be sold and when
- **Production:** Increasing production speed, automating production
- **Airlines:** Using the autopilot feature
- **Cyber Security:** spam mail detection, ID/credit card fraud detection, Malware detection

7.1 Types of Machine Learning Algorithms

7.1.1 Supervised

The aim in supervised learning is to learn a context between the input values and the target variable and to make predictions about the new values based on this context. Learning models consist of "input" and "output" data, where the output is labeled with the desired value. With algorithms, the system starts to identify similarities, differences until it can compile all the training data and predict it on its own.

Classification and Regression can be done with supervised learning.

7.1.2 Unsupervised

In unsupervised learning, we only have input. The machine examines the unlabeled input data and starts identifying patterns using all associated data. They are algorithms developed to reveal the structure and relationships in data containing only inputs. The more examples you acquire, the greater your ability to describe and categorize.

Clustering and estimation can be done with unsupervised learning.

7.1.3 Semi-supervised

If there is a small amount of labeled data and a large amount of unlabeled datasets, the method of learning unlabeled data from the labeled data is the semi-supervised learning method.

Working with large volumes of data with target output variables can be very costly. In addition, it is often not possible to reach data with output value.

7.1.4 Reinforcement

In reinforcement learning, the learning machine, called the agent, reacts to the situations it encounters and receives a reward signal as a result. He works to maximize the reward points he receives. It is a machine learning technique that aims to reach the maximum reward through trial and error by using the feedback from the agent's own actions and experiences.

7.2 Machine Learning Algorithms

7.2.1 Linear regression

In the linear regression algorithm, based on the observations, the relationship between the two data is tried to be estimated using an optimal linear equation. In order to increase the correct prediction rate, the regions closest to the points are selected.

7.2.2 Decision tree

Decision trees is a supervised learning algorithm that works by splitting data according to a predetermined parameter. The first node is called the Root node. Root nodes are located below the nodes. The data is specified at the nodes in the tree structure, the leaves at the bottom of the tree are the decisions and give the result.

7.2.3 Naive bayes algorithm

It is an algorithm based on conditional probability. Estimates are made with the probability table created according to feauteres. Probability is calculated for each element in the data set used and classification is made according to the one with the highest probability value.

7.2.4 Neural betworks

A neural network tries to recognize relationships within a dataset by imitating the operation of the human brain. The learning process takes place with the inputs given to the algorithm and the expected outputs. It is expected to learn by training with previously labeled data.

7.2.5 K nearest neighbor (KNN) algorithm

The new data given to the Knn algorithm is classified by calculating the distance from the data in the learning data set. The distance between the predetermined “k” number of data consisting mostly of odd numbers and the new data is calculated. Minkowski distance calculation function is used for this calculation. Classification is done by finding the nearest neighbor.

To determine a new data class, the classes of its k closest neighbors are examined and the class of the new data is predicted in whichever class the neighbors are in the majority.

7.2.6 Random forest algorithm

Classification is performed using more than one decision tree. Thus, more accurate and stable results are tried to be obtained. With large datasets, output is produced with high accuracy. Different subtrees are created using random features. Each tree makes its own classification. Then, all predictions are brought together to make the final prediction. It has advantages such as strong generalization ability and performance, low risk of overfitting, and being resistant to missing data.

7.2.7 Support vector machine (SVM) algorithm

The basis of support vector machines is based on the supervised learning model and can be used in both classification and regression studies.

Margin is the separation gap between the closest data points between different classes. Hyperplane is a decision boundary that separates data in different classes.

SVM generally uses the plane with the maximum margin to separate the data of the two classes as shown in Figure 12. Hyperplanes are used to optimally separate the two classes from each other. In some classifications, data can enter the margin region, as seen in the example Figure 13. These situations are called Soft Margin. Hard Margin, on the other hand, is sensitive to outliers and works when the data can be linearly separated.

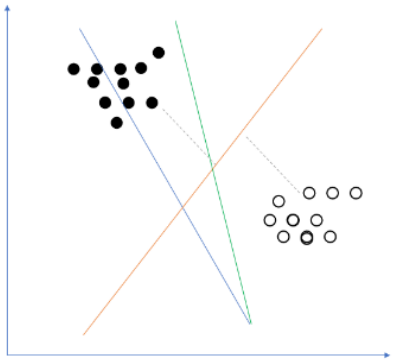


Figure 12: Hyperplane in SVM Model

In some classifications, data can enter the margin region, as seen in the example Figure 13. These situations are called Soft Margin. Hard Margin, on the other hand, is sensitive to outliers and works when the data can be linearly separated.

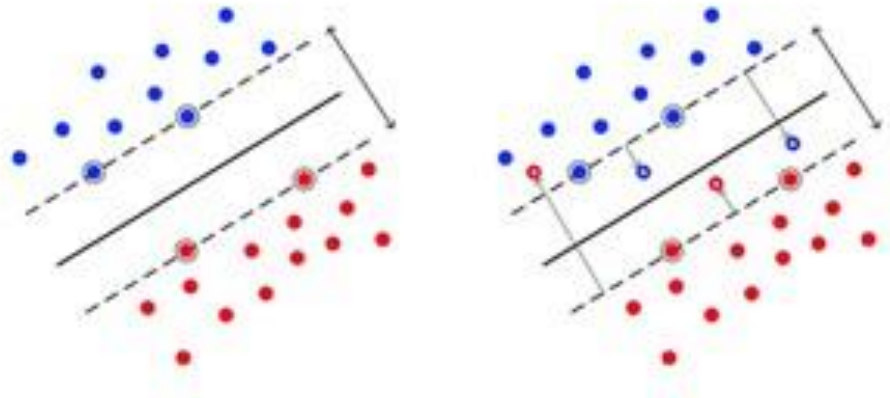


Figure 13: Hard-Margin vs Soft-Margin

The state between Hard magin and Soft Margin can be controlled with parameter C. As the C value increases, the margin decreases.

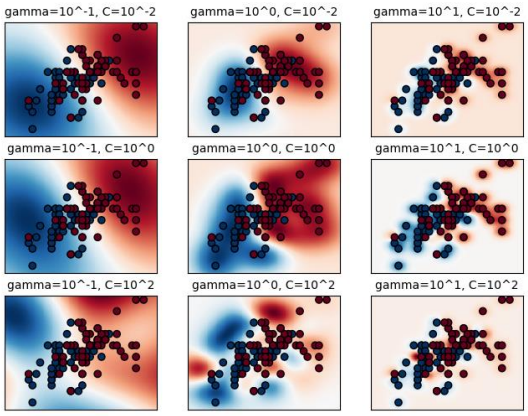


Figure 14: RBF Kernel

Using the rbf kernel function in the Figure 14, it can be seen how the change occurs with different c and gamma values. As you move to the right in the graphs, it is seen that the blue and red areas in the background become smaller and more overlap with the data in the graph. In other words, when an estimate is made, it is understood that more accurate results will come. As you go down, the C value increases and it is seen that there are deviations in the graph.

Kernel Functions: The use of one-dimensional functions will not be a realistic approach to classify real-life data. Kernel functions are used to adapt the SVM algorithm to real life data that cannot be separated linearly in small dimensions. By using the kernel mapping function, it moves the two-dimensional space space to the 3-dimensional space. Thus, data points are linearly separable like visualized in Figure 15.

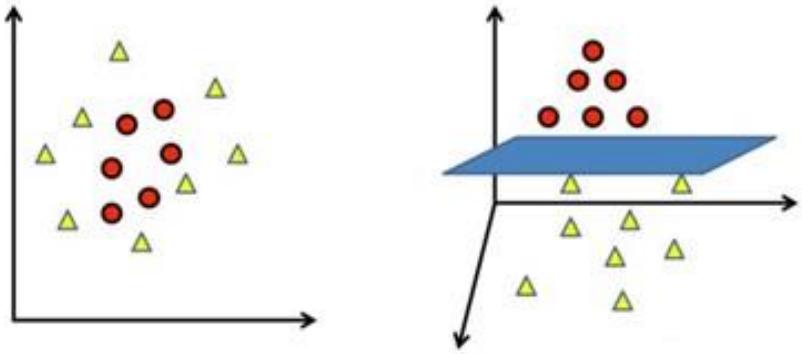


Figure 15: Kernel Functions

There are four commonly used functions:

Linear: Classes are separated by drawing a linear line in three-dimensional space.

Polynomial: Data are separated by plotting polynomials. The polynomial kernel controls the inputs as well as their combinations.

Sigmoid: The sigmoid function separates two data from each other. It is used in artificial neural networks.

Radial Basis: Classification is done by determining how similar each point is to the determined points.

7.3 Datasets in Machine Learning

A dataset is simply a collection of data. It can also be defined as a table consisting of rows and columns, each of which represents a variable, which may or may not be related to each other. The last columns in tables are usually labels. Tags are outputs after a model has been trained.

Features are the input data observed in the machine learning model. They are the characteristics used to train the model and are represented by columns in the table like Figure 16. Each column represents a different feature.

id	date	user	pc	activity
{F3X8-Y2GT43DR-4906OHBL}	1.02.2010 02:19	DNS1758	PC-0414	Logon
{B4Q0-D0GM24KN-3704MAII}	1.02.2010 02:31	DNS1758	PC-0414	Logoff
{T7J1-D4HK34KV-5476TCIJ}	1.02.2010 02:34	DNS1758	PC-5313	Logon
{S4Y6-D8MQ05SA-0759HLIS}	1.02.2010 02:53	DNS1758	PC-5313	Logoff
{F3P0-E7FH78CV-4874FRGZ}	1.02.2010 04:07	DNS1758	PC-0012	Logon
{M6C2-B1LK96JR-3409QEXP}	1.02.2010 04:10	DNS1758	PC-0012	Logoff
{J3S1-G0CN29NY-6126GYJV}	1.02.2010 06:16	ANC1950	PC-4921	Logon
{M1V4-S1FK13BQ-2952XBLU}	1.02.2010 06:25	SAB1954	PC-5091	Logon

Figure 16: DataSet Example

The dataset dataset used for machine learning and self-training is called Training Set. The test set is the data set used to evaluate the machine learning model. After the machine is trained using the Training data, the Test data is used to check whether it works as expected. Training data and Testing data must be different from each other. The dataset used to evaluate the performance of the model during the training phase can also be defined as the Validation dataset as shown in Figure 17.

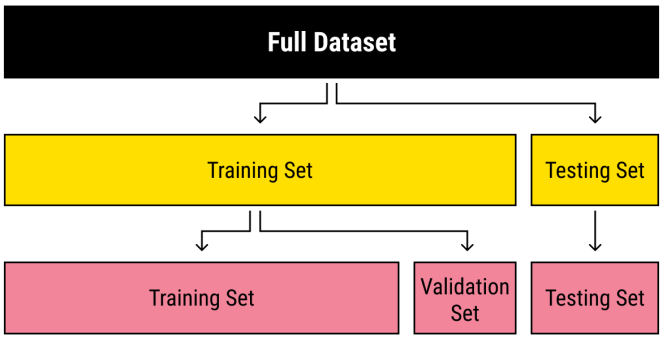


Figure 17: Datasets

7.4 Machine Learning Model

Machine learning models are mathematical expressions used to predict unknown data based on available data. The created model is required to perform successfully on new data that it has not encountered before. If the model does not show the expected success, two situations can be mentioned:

Overfitting: In the case of over-learning, the model performs very well on the training dataset but less well on the test set. The model has memorized the situations in the

training set and will fail to predict them when it encounters a different situation in the test data set.

The overfitting problem may occur in complex models or when the number of features of the model is much higher than the number of observations. To reduce overfitting, the training data can be increased and the complexity of the model reduced.

Underfitting: In the case of underfitting, the model performs unsuccessfully on both training and test sets. The model did not learn enough from the training data. Therefore, it cannot capture the connection between inputs and outputs. Model complexity can be increased to reduce underfitting. The number of features can be increased.

8. RESULTS

8.1 CERT Dataset

Researchers studying insider threats face a significant challenge due to the limited availability of real data, which is often denied use by organizations for security reasons. To overcome this problem, it is preferred to use synthetic data. The use of CERT database in systems used for insider threat detection has increased in the last decade.

The CERT® Insider Threat Center at Carnegie Mellon University's Software Engineering Institute (SEI) has collected and analyzed information on more than seven hundred insider cybercrimes, from national security espionage to trade secret theft. With the published guide, they share suggestions that can be applied by individuals, institutions and organizations, both by describing their research and by starting from these researches.

In this study, we used version 6.2 of the CERT Insider Threat database published by Marnegie Mellon University. The CERT dataset is an artificially developed synthetic dataset to test frameworks used to detect insider threats. In this data set, there are 5 different event sources as seen in Table 1.

Table 1: CERT Dataset

Datasets	**Fields**
logon.csv	id, date, user, pc, activity (Logon/Logoff)
device.csv	id, date, user, pc, file_tree, activity (connect/disconnect)
http.csv	id, date, user, pc, url, content
email.csv	id, date, user, pc, to, cc, bcc, from, activity, size, attachments, content
file.csv	id, date, user, pc, filename, content

As examined in previous chapters real cyber incidents carried out by insiders, time plays a very important role in identifying malicious users in the context of cyber security and detection of insider threats. Early detection is of great importance in preventing or reducing possible security breaches. Organizations can strengthen security measures by detecting anomalies and suspicious activities within a specified time frame and minimizing the impact of insider threats. In addition, the longer time

malicious actors spend on the targeted platform or system allows them to meticulously analyze the environment, make the necessary preparations and carry out violation attempts. It is this prolonged interaction that provides a critical window of opportunity for security professionals to detect irregularities. Recognizing abnormal time patterns or prolonged access is effective in identifying potential internal threats. Such deviations from established norms can serve as a red flag and lead to closer investigation of the user's behavior and intentions.

Therefore, in this study, we used the “time” users spend in a certain banking application/service as the main feature. We took these data from the logon logs, which contain logon/logoff activities, as indicated in Table 2.

Table 2: logon.csv

	date	**user**	**pc**	**date**	**time_spend**	**attack**
0	2.01.2010 12:29	AAC0610	PC-1834	02.01.2010	4:30	0
1	2.01.2010 07:54	AAC0610	PC-1834	02.01.2010	9:06	0
2	3.01.2010 07:54	AAC0610	PC-1834	03.01.2010	9:13	0
3	4.01.2010 12:26	AAC0610	PC-1834	04.01.2010	4:46	0
4	4.01.2010 07:46	AAC0610	PC-1834	04.01.2010	9:12	0

The data set in Table 2 consists of 6 attributes. User is a unique value that specifies users, pc is the computer users use, date is the date the activity took place, and time_spend is the time spent by the user between the time the user logged in and the time he was logged off. The Attack column is the attribute that will predict whether the user is an insider or not.

In the data set we use, there are 1437 users and a total of 40843 rows of data belonging to these users. Among these data, there are 4 users who perform insider threat activity. In the Attack column, the behavior of these users is labeled with 1, the value of other users is 0. When we checked the distrubution of the "attack" column in the data set, it was seen that there were 118 values of 1 and 40725 values of 0. When we view this distribution as a percentage, we see that there are 99.71% 0 labels and 0.22% 1 labels. It is understood from here that there is a class imbalanced problem.

While preparing the dataset for the algorithm, we divided it into two as test and train. We determined 0.1 as test (4085) and the remaining data as train (36758) data.

Since the variables in the dataset were not in the same unit, feature scaling was applied. Feature scaling serves to standardize the numerical values of these different features into a common range, making them directly comparable and ensuring that no feature affects the model's learning process. By doing this, feature scaling helps improve the performance of various machine learning algorithms and enables fair and effective comparisons between the importance of different features. This transformation facilitates accurate assessment of relationships between variables, thus increasing the model's ability to make meaningful predictions or classifications.

One method we use to evaluate the performance of the model for the SVM algorithm is GridSearch CV. C, gamma, and kernel hyperparameters were adjusted via the GridSearch technique. GridSearch CV helps define parameters that will improve the performance of the model. A model is created with separate combinations for the hyperparameters and their values to be tested in the model, and the hyperparameters are determined according to the results. Gamma is the parameter of the Gaussian Kernel and C is the parameter representing misclassification of the training data. The kernel enables separation of classes by translating data into higher dimensions. Parameters, C for the range { 1, 10, 100, 1000}, gamma for the range { 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9} and kernel {RBF, 'poly' , 'gamma'} was used.

Various evaluation metrics can be used to evaluate the accuracy of the model results we use. One of the methods that can be used to examine model performance is Confusion Matrix. It gives information about the performance of the confusion matrix model and the error types produced. It is used to show the predicted and actual class labels with the model. Correct and incorrect predictions are grouped under 4 different classes:

- **True Positive (TP):** observation that is predicted to belong to a class does indeed belong to that class.
- **True Negative (TN):** The number of values that are classified as negative and are actually in the negative class.
- **False Positive (FP):** The number of values that are classified as positive but are actually in the negative class.

- **False Negative (FN):** It is the number of values classified as negative but belonging to the positive class.

Another method that can be used to evaluate model performance is the Classification Report. Classification Report calculates precision, recall, f1 and support scores for the model. These scores are calculated using values from the confusion matrix.

- **Precision:** Expresses the proportion of correctly classified data (Equation 1).

$$Precision = \frac{TP}{\mathrm{TP} + \mathrm{FP}} \quad (1)$$

- **Recall:** Expresses the ratio of correctly classified positive values (Equation 2).

$$Recall = \frac{TP}{\mathrm{TP} + \mathrm{FN}} \quad (2)$$

- **F1 (F-score):** It expresses the harmonic average of Precision and Recall values (Equation 3).

$$\mathrm{F1} = 2 * \frac{Precision * Recall}{\mathrm{Precision} + \mathrm{Recall}} \quad (3)$$

- **Support:** is the number of occurrences in our dataset.
- **Accuracy:** It is the collection part of the correct classifications (Equation 4).

$$Accuracy = \frac{TP + TN}{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}} \quad (4)$$

- **Error Rate:** It is the rate of misclassifications (Equation 5).

$$Error\ Rate = \frac{FP + FN}{\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN}} = 1 - Accuracy \quad (5)$$

We used three ML classification models to classify the users into normal and abnormal: SVM, random forest and KNN. Table 3 shows the results of machine learning algorithms.

Table 3: Comparison with different classification methods

Model	Precision	Recall	F1 - Score	Accuracy	True Positive Rate	False Positive Rate	Specificity
SVM	0.9989	0.9978	0.9983	0.9967	0.9989	0.7500	0.2500
Random Forest	0.9975	0.9970	0.9972	0.9945	0.9975	1.0000	0.0000
KNN	0.9980	0.9973	0.9976	0.9953	0.9980	0.7857	0.2143

There are three different classification models we used. The SVM model demonstrated the most effective performance by successfully detecting most of threat users. The initial step of the SVM was to determine the optimal combination of C and gamma parameters. A high value of C attempts to minimize misclassification of the training data, while a low value contributes to smoothing the model. A lower C causes the model to accept the training data more generally and tolerate some errors. A high value of C allows the model to fit the training data more tightly. In this case, the model learns from the training data in more detail, but becomes prone to overfitting. Overfitting can cause the model to become too specific to the training data and lose its ability to generalize. On the contrary, a high gamma value may cause overfitting problems. The SVM model exhibited the highest overall performance compared to other algorithms. It also had a higher recall rate and fewer false alarms.

The Receiver Operating Characteristic (ROC) curve is a graphical tool used to evaluate the performance of a classification model. This curve visualizes the trade-off between sensitivity and specificity of a model. Sensitivity represents the true positive (TP) rate and specificity represents the true negative (TN) rate. The ROC curve allows analyzing the performance of the model at different thresholds. The curve is drawn from the upper left corner to the lower right corner, and a curve that represents the ideal case will be closer to the upper left corner of the curve.

Area Under the ROC Curve (ROC AUC) is a metric that expresses the area under the ROC curve. ROC AUC evaluates the classification performance of a model to an overall extent. The value range is 0 to 1, the closer to 1 the better. A ROC AUC of 0.5 means that the model is making random predictions, while approaching 1 represents excellent classification ability.

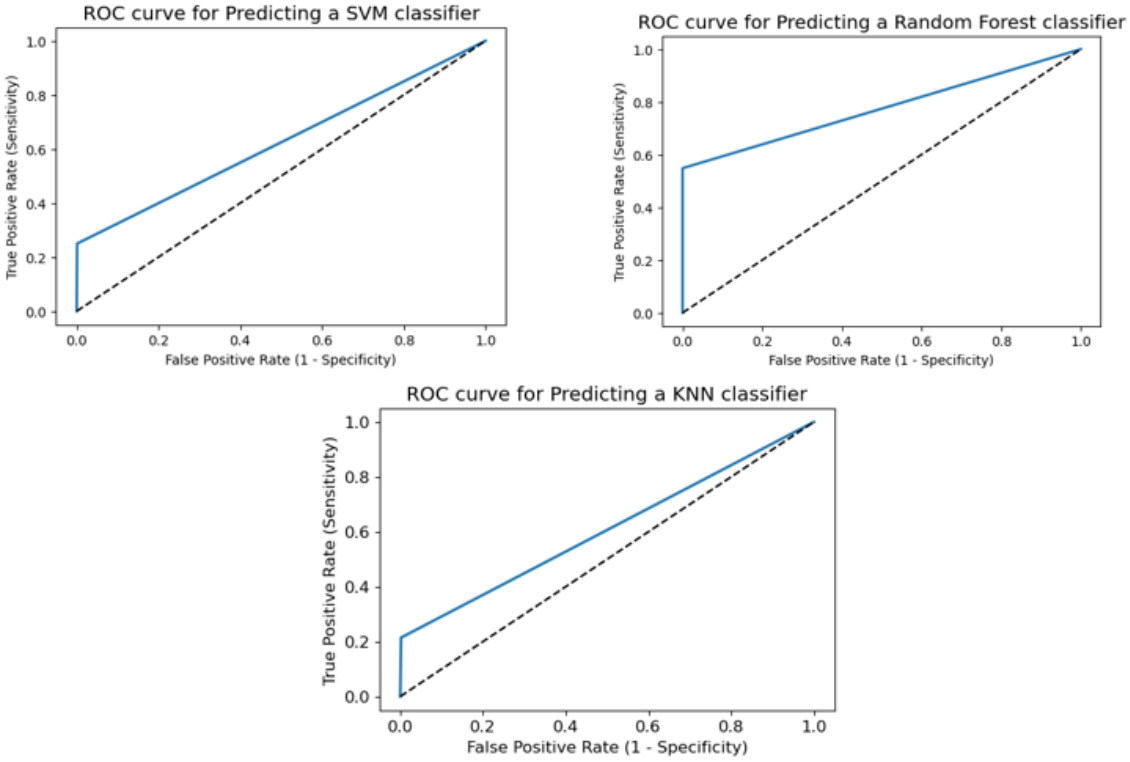


Figure 18: ROC Curve

The performance of SVM, KNN and Random Forest algorithms was also evaluated through ROC curves. The ROC AUC value of SVM was determined as 0.6244, the value of KNN was 0.6062 and the value of Random Forest was 0.7742. ROC curves visually demonstrated the ability of each algorithm to detect attack situations. In particular, Figure 18 ROC Curve reveals that the Random Forest model has a higher performance than other models.

9. CONCLUSION

This study examined in detail the performance of SVM, Random Forest and KNN algorithms applied for Insider Threat detection on an imbalanced data set. First, the SVM model showed moderate performance with a ROC AUC value of 0.6244. While this model showed some success in detecting "attack" situations, it had difficulty in correctly classifying non-attack situations. In particular, the low True Negatives (TN) and specificity value indicate that the model tends to incorrectly label non-attack situations as "attacks".

The Random Forest model demonstrated high performance with a ROC AUC value of 0.7742. A high True Positives (TP) value indicates that the model is successful in correctly detecting "attack" situations, while low specificity and zero TN value indicate that the model tends to mislabel non-attack situations. It is seen that in the unbalanced data set, this model focuses on "attack" situations and therefore has difficulty in correctly classifying non-attack situations.

The KNN model showed a lower performance with a ROC AUC value of 0.6062. While high TP and low False Negatives (FN) values indicate that the model is successful in correctly detecting "attack" situations, low specificity value indicates that the model tends to incorrectly label non-attack situations as "attack".

Overall, the Random Forest model seems to stand out with its high ROC AUC value and TP value. However, it should be noted that all models have difficulty correctly classifying non-attack situations and may need improvement in terms of specificity.

This study evaluated in detail the performance of Insider Threat detection algorithms under imbalanced data set conditions. The results can help us understand the strengths and weaknesses of each model and provide guidance for future studies. However, further research and model improvements could increase efficiency in this area and provide security professionals with more effective tools.

REFERENCES

[1] **Lin, L., Zhong, S., Jia, C., & Chen, K.** (2017, August). Insider threat detection based on deep belief network feature representation. In 2017 International Conference on Green Informatics (ICGI) (pp. 54-59). IEEE.

[2] **Jiang, J., Chen, J., Gu, T., Choo, K. K. R., Liu, C., Yu, M., ... & Mohapatra, P.** (2019, November). Anomaly detection with graph convolutional networks for insider threat and fraud detection. In MILCOM 2019-2019 IEEE Military Communications Conference (MILCOM) (pp. 109-114). IEEE.

[3] **Aldairi, M., Karimi, L., & Joshi, J.** (2019, July). A trust aware unsupervised learning approach for insider threat detection. In 2019 IEEE 20th International Conference on Information Reuse and Integration for Data Science (IRI) (pp. 89-98). IEEE.

[4] **Singh, M., Mehtre, B. M., & Sangeetha, S. (2021, May).** User behaviour based insider threat detection in critical infrastructures. In 2021 2nd International Conference on Secure Cyber Computing and Communications (ICSCCC) (pp. 489-494). IEEE.

[5] **Parveen, P., Weger, Z. R., Thuraisingham, B., Hamlen, K., & Khan, L**. (2011, November). Supervised learning for insider threat detection using stream mining. In 2011 IEEE 23rd international conference on tools with artificial intelligence (pp. 1032-1039). IEEE.

[6] **Oberlo.** (2023). How Many People Use The Internet? https://www.oberlo.com/statistics/how-many-people-use-internet

[7] **Trading Economics.** (2023, May). Individuals using the internet for internet banking. https://tradingeconomics.com/euro-area/individuals-using-the-internet-for-internet-banking-eurostat-data.html

[8] **X-Force Threat Intelligence Index 2023 Report.(**2023). IBM Security. https://www.ibm.com/reports/threat-intelligence

[9] **UpGuard.** (2023, May, 04). 10 Biggest Data Breaches in Finance. https://www.upguard.com/blog/biggest-data-breaches-financial-services

[10] **Maamar, Zakaria. (2003).** Commerce, e-commerce, and m-commerce: what comes next? Commun. ACM 46, 12 (December 2003), 251–257.

[11] **Tsang, B.** (2022). e-Commerce Payment Processing: An Essential Guide For Online Sellers. https://choco-up.com/blogs/ecommerce-payment-processing

[12] **Shopify.** (2022.May 9), Ecommerce Payment Processing: An Ultimate Guide. https://www.shopify.com/blog/ecommerce-payment-processing

[13] **Fatonah, S., Yulandari, A., & Wibowo, F. W.** (2018, December). A review of e-payment system in e-commerce. In Journal of Physics: Conference Series (Vol. 1140, No. 1, p. 012033). IOP Publishing.

[14] **DigiCert.** WHAT IS SSL CRYPTOGRAPHY? https://www.digicert.com/faq/cryptography/what-is-ssl-cryptography

[15] **Cisco**, What Is Encryption?. https://www.cisco.com/c/en/us/products/security/encryption-explained.html#~q-a

[16] **Bruce J Mack's Software Development Reference Materials**. (2013). http://guides.brucejmack.net/SOA-Patterns/WSSP/13.1PublicKeyEncryptDigSigDoc.htm

[17] **YAŞAR, S.N., DİKİCİ, F. C., TANYILDIZI, E., KARAKÖSE, E.,** (2021). A Generator Design Based on the Middle Square and SHA3 Algorithm for the Requirements of Randomness in Science and Engineering Studies. Fırat Üniversitesi Fen Bil. Dergisi 33(1), 81-91, 2021

[18] **Mitsophonsiri, K., Punthawanunt, Suphanchai, Mitatha, S., Yupapin, Preecha.** (2011). Data Security Transmission via a Noisy Channel. Procedia Engineering. 8. 487-492. 10.1016/j.proeng.2011.03.088.

[19] **Villiers, R. D.** (2021, Feb). Demystifying SSL / TLS certificates. https://www.lawtrust.co.za/knowledge-hub/blog/lawtrust-blog/2021/02/24/demystifying-ssl-tls-certificates

[20] **Kaspersky.** (2022). IT Security Econmics Report. https://go.kaspersky.com/rs/802-IJN-240/images/IT%20Security%20Econmics%20Report%209.18.17.pdf?aliId=488652022

[21] **Global Economic Crime and Fraud Survey.** (2018). Pulling fraud out of the shadows. https://www.pwc.com/gx/en/news-room/docs/pwc-global-economic-crime-survey-report.pdf

[22] **Ponemon Institute.** (2020). 2020 Cost of Insider Threats: Global Report.

[23] **McAfee.** (2018). Insider Threat at Coca-Cola Compromises 8,000 Employees' Information. https://www.mcafee.com/blogs/consumer/consumer-threat-reports/insider-threat-at-coca-cola-compromises-information/

[24] **Kamuoyu Duyurusu (**Veri İhlali Bildirimi) – T. Garanti Bankası AŞ. (2021). https://www.kvkk.gov.tr/Icerik/7004/Kamuoyu-Duyurusu-Veri-Ihlali-Bildirimi-T-Garanti-Bankasi-AS

[25] **Kamuoyu Duyurusu** (Veri İhlali Bildirimi) – Türkiye İş Bankası A.Ş. (2019). https://www.kvkk.gov.tr/Icerik/5526/Kamuoyu-Duyurusu-Veri-Ihlali-Bildirimi-Turkiye-Is-Bankasi-A-S-

[26] **Kamuoyu Duyurusu** (Veri İhlali Bildirimi) – DenizBank A.Ş. (2019). https://www.kvkk.gov.tr/Icerik/5516/Kamuoyu-Duyurusu-Veri-Ihlali-Bildirimi-DenizBank-A-S-

[27] **Kamuoyu Duyurusu** (Veri İhlali Bildirimi) - ING Bank A.Ş. (2019). https://www.kvkk.gov.tr/Icerik/5375/Kamuoyu-Duyurusu-Veri-Ihlali-Bildirimi-ING-Bank-A-S-

[28] **Kamuoyu Duyurusu** (Veri İhlali Bildirimi) – Yapı ve Kredi Bankası AŞ. (2021). https://www.kvkk.gov.tr/Icerik/6950/Kamuoyu-Duyurusu-Veri-Ihlali-Bildirimi-Yapi-ve-Kredi-Bankasi-AS

[29] **M. Somvanshi, P. Chavan, S. Tambade and S. V. Shinde,** "A review of machine learning techniques using decision tree and support vector machine," 2016 International Conference on Computing Communication Control and automation (ICCUBEA), Pune, India, 2016, pp. 1-7, doi: 10.1109/ICCUBEA.2016.7860040.

[30] **Abdullah, D. M., & Abdulazeez, A. M**. (2021). Machine learning applications based on SVM classification a review. Qubahan Academic Journal, 1(2), 81-90.

[31] **Huang, S., Cai, N., Pacheco, P. P., Narrandes, S., Wang, Y., & Xu, W.** (2018). Applications of support vector machine (SVM) learning in cancer genomics. Cancer genomics & proteomics, 15(1), 41-51.

[32] **Osisanwo, F. Y., Akinsola, J. E. T., Awodele, O., Hinmikaiye, J. O., Olakanmi, O., & Akinjobi, J.** (2017). Supervised machine learning algorithms: classification and comparison. International Journal of Computer Trends and Technology (IJCTT), 48(3), 128-138.

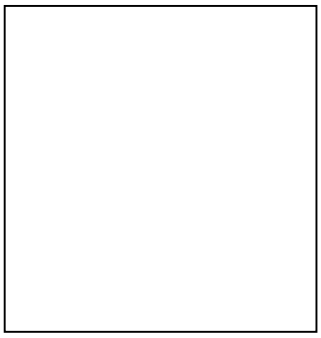
[33] **Dima, C.** (2016). Basics of support vector machines. https://www.cristiandima.com/basics-of-support-vector-machines

[34] **MLMath,** (2019). Math behind SVM(Support Vector Machine). https://ankitnitjsr13.medium.com/math-behind-svm-support-vector-machine-864e58977fdb

[35] **Scikit Learn.** RBF SVM parameters. https://scikit-learn.org/stable/auto_examples/svm/plot_rbf_parameters.html

[36] **Al-Behadili, Husam., Grumpe, Arne., Dopp, Christian., Wohler, Christian**. (2015). Non-linear distance based large scale data classifications. 613-617. 10.1109/PIC.2015.7489921.

CURRICULUM VITAE



Name Surname **:** Ayşenur KUMBASAR

EDUCATION **:**

- **B.Sc.** **:** 2020, Istanbul University, Department of Engineering, Computer Engineering

List of Publications and Patents:

PUBLICATIONS/PRESENTATIONS ON THE THESIS

- Hejazi N., Kumbasar A., Ozdemir E., 2023: Detecting Malicious Acitivity Inside of The Network. 2nd International Graduate Research Symposium (IGRS'23), May 16-18, 2023 Istanbul, Turkey