CYBERSECURITY CULTURE INFLUENCES ON PHISHING

A Dissertation Proposal Submitted to the Solent University Southampton

In Partial Fulfilment of the Requirements for the Degree MSC Cyber Security

MSc Cyber Security Engineering

Academic Year 2021-2022

Research Project (MAA112)

Student Number: 15804488

Name: Aditya P.S Parihar

Tutor: Dr Olufemi Isiaq

Supervisor: Kalin Penev

This report is submitted in fulfilment of the requirements of Southampton Solent University for the degree of MSc Cyber Security Engineering.

Table of Contents

List of figures

List of tables

CHAPTER 1 – INTRODUCTION

Overview of the study

Cybercrime is a form of digital misconduct in which an individual commits a crime by gaining illegal access to a processer, a mobile device, as well as a website to carry out other crimes including cash withdrawals and transfers. The same theoretical foundation will underlie this study's investigation of *"Cybersecurity in Banking Influences on Frauds."* This study will include an introduction, historical context, a review of related literature, and pertinent statistics. A particular literature review and study approach will also be effectively explained in addition to this.

Background of the study

Cybersecurity is the process of defending against malicious intrusions on networks, computers, servers, mobile devices, electronic systems, and data. It is stimulating to put effective cybersecurity protections in place in the contemporary world since there are more strategies than folks and hackers are flattering more creative. Cybercriminals frequently use phishing attacks to target employees in the banking industry in order to get beyond technical and security defenses. Phishing assaults on American companies and organizations increased by 40% in 2018. Because of the gathered data, sensitive information, and financial assets of banking companies, cybercriminals are drawn to banks.

The banking cybersecurity culture merits investigation given the significance of U.S. banks to national and international markets, globalization, and personal financing. Cybercriminals constantly strive to employ different phishing techniques to get access to the information systems substructure, sensitive and private data, intellectual property, and financial resources since banks are high-value targets. Through phishing scams, a cybercriminal can steal an employee's login information and get beyond network security measures. Through adherence to organizational security regulations, a rise in cybersecurity awareness, institutional learning, and leadership, cybersecurity culture strives to promote positive security behavior (Priya, and Saradha, 2021).

By practicing good security behavior, employers can prevent phishing instances by influencing employee behavior. Banks can effectively train staff to defend against phishing attempts by utilizing a cybersecurity culture. Successful phishing attempts led to the theft of identities, financial fraud, espionage, and hacktivism, which in turn caused unheard-of financial damages. The absence of scholarly study and theory contributions on cybersecurity cultural influences on decreasing phishing vulnerability is one problem affecting phishing susceptibility in the U.S. banking industry (Tiwari, et al., 2021).

Computer systems and networks are somewhat standardized in cybersecurity culture, but end users are not; this is because, according to studies, cultural values shape and form human behavior. The techniques of already-existing disciplines including data technology, computer science, data safety, accounting, marketing, plan, finance, and risk organization are the foundation of the cybersecurity sector. The development of security controls and mitigating procedures to improve protection and comprehend risk exposure was considered to be highly dependent on the cybersecurity culture.

The information security risk agenda and valuation for detailing and assessing the organization's cybersecurity philosophy are both integrated into the cybersecurity culture research methodology that is given to enterprises. the rising cost of data breaches, ransomware attacks, reputational harm, and human mistake, together with an increase in cybersecurity threats.

Aims and objectives

Aim

The main aim of this research is ***"to identify the significance of cyber security within banking sector which impact fraudulent".***

Research objectives

The main objectives of this research are discussed as underneath:

- To analyze the significance of cyber security and its role in the banking sector.
- To determine the role of cyber security within the banking sector.
- To analyze the factors of cyber security that influence fraud in banking culture.

Research questions

- What is the concept of cyber security and its role in the banking sector?
- What is the significance of cyber security within the banking sector?
- What are the factors of cyber security that influence fraud in banking culture?

Problem statement

Because finance and banking are directly related to people's assets and funds, fraud risks are rising across all businesses as a result of the expansion of technology in numerous sectors. Cybercrime can be defeated if the Cyber Security and Data Technology Act comprises suitable penalties for cyber phishing. Around the world, several cyber frauds will be discovered. To address the issue of cyber security and provide a workable solution, this effort will look further. On the other hand, security lapses may make it harder to have faith in financial institutions. Customers may decide to devote their money away as a result of data breaks brought on by shoddy cyber security measures.

Outline of the methodology

Research area	use concept
research philosophy	Pragmatism
Research approach	deductive primary data and inductive secondary data
Data collection	both primary and secondary
research design	Explanatory research
sampling method	probability sampling method
data analysis	descriptive analysis

Scope of the study

This study will support the creation of a fraud-fighting cyber security explanation for the lending sector. In other words, cybercrime is described as digital misconduct where a criminal gains illegal access to a processor or other electronic device, as well as the internet, in terms of committing a variety of crimes, like cash transfers and extractions. Cybercrime can harm a corporation's reputation, growth prospects, and earnings over time.

Significance of the study

The present study is essential in terms of providing and making people understand the significance of cyber security within the banking sector and making people aware of fraud. It has been analyzed that cyber security is important in terms of protecting all essential documents and data from any sort of damage and theft. It leads includes sensitive data, personally recognizable information, intellectual property, administration, and industry information systems (Gyamfi, and Abdulai, 2018). It is important to research this topic to spread knowledge about cyber security factors like data leakages that can make it complicated to trust monetary institutions. For banking sectors, a weak cyber security system is a serious problem as it can quantity to data breaches that can easily harm their customers and take their cash somewhere else.

The rationale of the study

The reason behind conducting this research is to develop knowledge about cyber security and its significance in the banking sector to make people understand fraud and also spread awareness about the same. Cyber threats can come from anywhere in the organization due to this work needs to include cyber security awareness training to make people understand and educate staff members about common cyber threats like phishing, ransomware attacks, social engineering scams, and so forth which are designed to steal intellectual property or any other personal data. Hence, banks need to be on their full security and guard instead of most businesses.

Structure of the study

Chapter 1: Introduction

This section of the study conducts a thorough analysis of the background, purpose, objectives, and overall research questions, which aids in providing the reader with important information that has been provided by the researcher. In addition, the purpose of this investigation, its significance, and the justification for selecting the research issue will be mentioned to give readers the best comprehension of the research background and to successfully finish the entire study. It is known as the most important part of the dissertation that defines the purpose of conducting the research. The main objective of this chapter is to define the research aim and objectives. It is the part of a study in which there is an explanation of accurate information about the topic, organization, and methodology that will be used in completing the research work.

Chapter 2: Literature Review

To conduct the research efficiently, the secondary sources that are being gathered by the research are discussed in this chapter of the study. In terms of the current study, the researcher must effectively employ a variety of secondary sources that include online articles, journals, books, and other material that is being gathered from libraries and Google Scholar to provide pertinent information on the selected issue. This chapter gives each response that was used in the investigation the proper credit. The literature review depends upon various previous studies that are done only research topic. This chapter will include different sources like journals, and the article's former company websites that will be used in the analysis of research data. The main objective of the literature review is to assist the investigator in attaining research aims and objectives along with research questions.

Chapter 3: Research methodology

This portion of the study contains a thorough justification of the methodology the researcher employed to carry out the entire study. This chapter employs an appropriate research methodology that results in the collection of sources using both secondary and primary data. In addition to this, surveys and literature studies using the approach of gathering data online are also used. This chapter of research provides help in aiding the collection and evaluation of data that will be associated with the research topic (Soni, 2019). This section consists of different forms of methodologies which will include approach, choice, philosophy, and strategy. It is useful in collecting and evaluating appropriate information and data about the research topic. Industry search from a quantitative method is used along with secondary information and thematic analysis is conducted.

Chapter 4: Data collection and analysis

This research chapter includes a useful description of the various data that the researcher acquired in order to finish the study and produce useful results. Here, the acquired data will be appropriately discussed, along with its interpretation using graphs and tables. This chapter discusses the conclusions drawn from the data that was gathered. According to this chapter, it is significant in assisting in giving evaluation and gathering of information that will be used in the research. In this research, thematic analysis is conducted which will facilitate the investigator for evaluating quantitative data which will be gathered through a literature review. An important tool for research is a thematic analysis which will be used by the researcher to make the research outcomes.

Chapter 5: Conclusion and recommendation

This section summarizes the overall discussion of the research's findings, which also includes the discussion, findings, and a review of the literature. This chapter demonstrates how to combine factors utilizing research objectives and consider factors for the outcome. In this chapter, there is information about the significance of results and discussion which will lead to the accomplishment of the aim and objective (Soomro, et.al, 2019). The recommendation is an important part of this study because it helps in assisting the evaluation of banking fraud.

It has been concluded from the discussion above that it is not improbable that a bank will suffer losses as a result of a significant cyberattack. Almost all financial institutions have been affected in some form by a cyberattack, and the frequency of attacks is rising. Government agencies and financial institutions are getting more concerned about the possibility of cyberattacks and the potential financial repercussions. This study will concentrate on the same subject and make the same attempt to address the cyber security issue.

Summary of the chapter

This chapter includes an introduction at background about the research topic. It is concluded in this chapter that the trade landscape and digital banking are evolving at a high pace because of technological advancement. According to sound cyber safety practice former, different financial institutions simulate cyber security breach situations of various types in the form of cyber security drills it providing help to organizations in getting up for combatting and facing challenges in cyber security space. Moreover, organizations must emphasize security breaches that have happened in past (Tariq, 2018). There must be the development of a database that will provide insight for acting in the right direction.

Moreover, organizations must reiterate the significance of cyber security culture. Due to the monitoring of cyber security tools, it is easy to take action whenever required. By monitoring and implementing cyber security tools, it is easy to effectively drive away cyber security risks.

In the banking sector, fraud has increased. In this research, there will be a discussion about potential solutions that can be conducted for cyber security in banking influence on frauds (Zhang, 2018). There is also discussion about online banking offenses or cyber security measures.

Chapter-2 Literature review

As stated by Golden and Kohlbeck, (2020), it is revealed that cyber security is known as the activity of the protecting of different computing resources and sensitive and confidential information from any type of malicious attack. This is known as part of electronic security as well as information technology (IT) study. This is known as the defense of computer systems from different types of cyber-attacks. In the context of banking and finance, cyber security plays a significant role. This sector is considered a confidential one in which security and safety are the key elements. There are a lot of blockchain technology used and a diversified way that considers various practical approaches to minimize cyber-attacks. There are different types of risks which are related to cyber security. Out of these risks, online manipulation is known as one of the biggest risks. There are different types of fraud that are related to the banking system. As stated by Hasham, Joshi, and Mikkelsen, (2019), it is shown that verbal manipulation or malicious attacks can come through online fraud. In recent years, Internet usage has increased and because of this, online fraud has also increased. It is known as the biggest threat to financial organizations. The aspect of cyber phishing also referred to as online fraudulent this has become a very serious issue across the world. Different theories are developed on the association between cyber security and fraud in the banking sector. According to the views of Wechsler and Siwakoti (2022), The fraud triangle theory clarifies the cause by which different individuals or organizations are linked with fraud. There is also a description of the internal and external causes for individuals through which they committed fraud. Different theories of fraud include fraud diamond theory, scale theory, and self-control theory.

Cyber risk has become one of the major risks across the World. According to the reports of DTCC, Cyber risk has become one of the top risks in the banking and financial sector. The aspect of cyber-attack explains the risk which is related to information and technology assets. This is having negative consequences on affecting confidentiality. It also includes liability and property risk in cyber fraud.

According to the views of Afriyie (2022), The aspect of cyber security in banks includes a combination of methods, technologies, and strategies that are used for safeguarding devices, networks, and programs from different data security risks like hacking, data theft, virus, and unauthorized access. The main objective of cyber security in banking includes protecting the assets of users. They are influenced by the structure, aim policies, leadership, and practice of the banking sector (Priya, and Saradha, 2021). Developing an effective cyber security culture is very important for influencing employees and developing a strong as well as effective human firewall.

There is the inclusion of attitudes, value, and beliefs in the context of security in the banking sector for developing a cyber-security culture.

Cyber phishing is an attack that includes stealing money and it can be used for stealing the identity of a client like debit or credit card number, password, bank details, and many more (Jansen and Van Schaik, 2018). By using these details, cyber phishing is conducted and frauds take place. It is defined as a semantic attack that is used to exploit human vulnerabilities. The aspect of susceptibility is regarded as an attack by which different techniques of cyber phishing are evaluated.

As stated by Koibichuk, (2021), it is revealed that the banking sector is under attack For several years. Nowadays, cyber fraud has become very common. Several companies and individuals are performing most of the transactions that are made online, and the chances of a data breach have increased. The significance of cyber security in banking sector transactions is basically for protecting sets of customers. Nowadays people are most activities are performed online. In all these situations, personally identifiable information is redirected to different locations and used for malicious activities. This impact customers a lot. This is involved in harming the bank when they're attempting towards recovering data. When this is taken hostage, banks need to pay thousands of dollars for releasing the information. In return for this, the trust and loyalty of customers are last. There are three types of risk data related to banking on the web. These are mentioned below:

More risks from mobile apps – There are a lot of individuals who Access Bank accounts on mobile apps. Most of these individuals stand for having minimal or no security. It increases the risk of potential attack. That is why banking software solutions are required at different endpoints for preventing malicious activities.

Breaches at third-party organizations – Due to the upgradation of banking in cyber security, Hackers have turned to shared banking systems and third-party networks for gaining access (Nasution, et.al, 2018). When these are not protected by the bank then attackers can have access to the data very easily.

Increased risk of cryptocurrency hacks – In the context of standard funds, there is an increase in hacks in the context of cryptocurrency. This sector is unsure about how to implement cyber security software for banking which is why the ability of attackers to grab large amounts of currency is very high (Thennakoon, et al., 2019).

In the context of security on the Internet, there is a consideration for enhancement and replacement of current protection applications. various factors are considered in the context of banking software development. Some of these are explained below:

Security audit

There is a need for a thorough audit which is imperative before any type of cyber security software implementation. According to this review, it is easy to reveal the strength and weaknesses of the existing setup. It will also provide recommendations about how to save money by allowing all proper investments.

Firewalls

According to Al Duhaidahawi, et.al, (2020) Cyber security banking configuration is not having any type of application. This needs the right type of hardware for blocking attacks. Through an updated firewall former banks can block malicious activity before they are other parts of the network.

Antivirus and anti-malware applications

When there is an increment of firewall upgradation, it is not easy to stop attacks unless the anti-malware and antivirus applications are up to date (Nyakarimi, Kariuki, and Kariuki, 2020). The older software is not containing the latest rules and virus signatures. It can lack a potentially disastrous attack on the system.

Multi-factor authentication

This type of protection is extremely critical for protecting customers who are utilizing online or mobile apps for banking. Several users never change their passwords. By the usage of multi-factor authentication, attackers can be stopped from reaching the network. For example, for every transaction, a four or six-digit code is sent to the cell phone of the customer (Tiwari, et al., 2021).

Biometrics

It is known as another version of multi-factor authentication. It is a more secure version in which there is the enhancement of texted code. As given by Nobles (2021) In this type of authentication, there is relaying of retina scans, facial recognition, and thumbprints for confirming the identity of the user. Hackers can have access to this type of authentication in past but it is very difficult to accomplish.

Automatic logout

Many apps and websites allow users to stay logged in when they allow it. It provides access to information at any time without entering any login credentials (Pradesyah, Yuslem, and Batubara, 2021). This provides permission for attackers to easily obtain records. Automatic logout minimizes this through closing access of any user after a few minutes of inactivity.

Education

As stated by Al Duhaidahawi, et.al (2020), The measures explained above can increase cyber security in the banking sector. However, awareness and education about cyber security among customers will help in minimizing cyber attacks and fraud. Education is an essential factor that will provide enhancement and awareness of cyber security among clients

There is an increase in the chances of banking fraud because there is the usage of digital transactions nowadays. In the finance and banking sector, there is a direct relation between assets and funds people and because of this, frauds are the most common factor. When information technology and cyber security acts Are present for appropriate transactions in cyber phishing, it is easy to overcome cyber frauds and cybercrimes. Various cyber frauds are not known. This research, there will be focusing on solving the problem of cyber security in banks. Data breach is also a common problem in financial institutions and banks. It is known as the most common problem of banks and due to this, customers are facing various problems. As a result, there is declined customer satisfaction. This study will provide ways for overcoming this issue.

According to the views of Roy and Prabhakaran (2022), Digital misbehavior in which criminals are having access to any mobile device or computer, or website for performing different crimes like withdrawals and money transfers is known as a cybercrime (Al-Hashedi and Magalingam, 2021). This research depends upon the concept of cyber security in banking which is influencing fraud.

Technology has become one of the most essential factors which are involved in different daily activities of customers. There is a huge shift in digital space. Because of this, economies of various countries are rapidly shifting from cash to digital transactions. Because of this former, it is valuable for cyber-attacks (Bouveret, 2018). For making the transaction flexible for clients and customers, different bands have transitioned towards the incorporation of digital platforms that will enable digital modes of payments. It will play a significant role rising number of cyber deceptions in the context of the banking field.

The technological advancement in the banking sector has initiated a doorway that will increase the convenience of the banking experience. As there is an expansion of technology the scope of online banking has grown and due to this there is an exponential growth in banking frauds. Moreover, the increasing number of online transactions has increased the risk of committing malicious acts (Shukur, and Kurnaz, 2019).

The customers are mindful of various common risks that are linked with online finance stop, however; as there is the evolution of technology there is the development of sports stores that has come up with new and innovative ways for trapping claims. Several forward-moving digital platforms have failed to integrate the new technologies for claims to perform these financial transactions stop that is why banking institutions must provide dedication for more resources in the context of customer education and awareness programs (De Kimpe, et.al, 2022).

According to **Ileberi, and Wang, 2022,** the advance in e-payment and e-commerce systems has been sparked by increasing in financial fraud cases which as credit card fraud. There is a crucial case that has been implemented by the mechanism which can notice credit card fraud. Different features of credit card fraud have been playing a crucial role when machine learning has been used in credit card fraud detection which has been chosen appropriately. The machine learning algorithm (ML) is based on credit card fraud detection which has been used by the engine by the genetic algorithm (GA) for a different selection of the feature. After the features which have been chosen are optimized there is the proposed detection which used the following ML classifiers which are (RF), Random Forest, (DT) Decision Tree, (ANN) Artificial Neural Network, (LR) Logistic Regression, and (NB) Naïve Bayes. By validating the performance there is the proposed credit performance by the projected credit card fraud detection engine which is assessed by using the dataset to be generated by the European cardholders. The outcome has been proposed by the approach to outperforming the existing system. There has been significant improvement in the system of the internet. The proliferation is increasing the use of the services which are tap and pay systems, e-commerce and online bill payment schemes, and many more. The impostors have been increasing the number of machines that have been used in protecting credit card transactions such as credit card tokenization and encryption (Shukur, and Kurnaz, 2019). Different methods are effective in many cases as they are not majorly protected by the credit card transaction that is against fraud.

The dataset has been used in developing the ML models inky for credit card fraud detection which has been contained by the anonymized attributes. Credit card fraud detection is to be known as a challenging task that has been constantly changing by the different patterns and nature of fraudulent transactions. For credit card fraud detection the existing ML model has been majorly suffering from the detection accuracy which is not solving the major issue like the credit card fraud datasets.

GA is based features selection method majorly in conjunction with the AAN, RF, NB, DT, and LR has been proposed. The GA has been implemented by the RF which is the fitness function. GA has been applied to the European cardholder's credit card transaction dataset that has 5 optimal features to be generated. The result has been achieved by using the GA attributes to be selected by demonstrating by the GA-RF to achieve the accuracy of the optimal to 99.98%. Other classifiers are also achieved such as GA-DT with a remarkable accuracy that is 99.92%. The main result has been obtained by the superior mark which has been achieved by the existing methods. It has been implemented that the framework of the synthetic credit card fraud dataset has been validating the result to be obtained by the European credit card frauds.

According to **Dornadula, and Geetha, 2019,** using technology such as phishing techniques by internet banking fraud majorly means removing and transferring the money from the banker's account by not taking permission from the banker. Credit card fraud has been happening in a large amount where some banking companies have been giving services to the banks that have been facing the major issues. In this report, it has been implemented the model has been built by predicting the non-fraud and also the fraud transaction with the respect to the amount in the time of transaction which is used by the classification of the machine learning algorithm and statistics. There is a linear algebra that has been built by the complex machine learning algorithm model that has been understanding and predicting the data set. A credit card is the smallest fiber card that has been contained by the person named and signature which has been linked to the account. The card information has been read by the ATM by swiping up the machine, from the store's readers which are detected by the bank and online transaction. There is a unique number on the card which is one of the major important parts to keep the security which does rely on the physical security of the card that can prevent the privacy of the credit card number.

There is an increase in the transaction of credit cards that has led to considerable development in deceitful cases. There are many statistical methods and data mining that have been implemented by using artificial intelligence and also pattern matching. There is the detection of fraud has been used by secure and efficient methods that are very important (Zhang, et al., 2018). Credit card fraud has been increasing and the fraud of financial loss is also increasing. Due to the online internet transaction which has been growing as the new technology. There are different machine algorithm which has been used by detecting fraud through the hybrid algorithm, and also the artificial neural network which has been used by giving the best performance.

Credit card fraud has been resulting in the loss of money for many people and also the loss of banks and credit card companies. Helping the people who have lost money has been trying to develop the main model that is efficiently separating by fraud to have fewer transactions in developing the features in the data set has been given by Kegel. There is some machine learning algorithm that has been used which are decision tree, logistic regression, and supporting vector machine (Tiwari, et al., 2021). All of these have been supervised by the machine learning algorithm into the machine learning. Some features have been solved by the problem statement that has been used by another part of the artificial intelligence that has been seen in the time series analysis. The amount and time features have been predicted by the weather transaction which is a non-fraud or fraud transaction. In the time series analysis, this has been recommended by reducing the number of parameters where the features have been required by the model that can be easily achieved through the average method.

There are different types of frauds that are conducted in the banking sector due to online transactions or offline transactions. Some of these are mentioned below :

Credit card fraud

As per (Rodrigues, et.al (2022), Fraudsters have started scamming by the usage of specific targeting. There is a common scamming technique that includes credit cards which are vulnerable to unauthorized transactions. Users of credit cards must follow different preventive and security measures for protecting sensitive data from a type of deceitful card actions. Users need to use credit cards along with preventive measures two keeping car details, pin, and login details private (Minastireanu, and Mesnita, 2019). It is recommended to block the card if it is not in use. It is also suggested to not use the cards for online transactions to reduce the chances of fraud.

E-mail phishing

As studies expansion of technology, it has increased the opportunity for going paperless. There are lots of people who have stored information and data on the computer that includes failures or risks (Georgiadou, et.al, 2022). It requires customers to adopt preventive measures and an efficient way of protecting confidential information including clicking on the malicious link which is integrated into the e-mail received from a malicious and unknown source. It is recommended to evaluate the URL carefully before opening any attached link. In phishing sites or emails, there is a setup of a website that is the identical link to direct the users to an insecure web page.

Password use

It is the most important security tape which protects customers or clients from any banking fraud. When customers visit any website, that seeks for remembering or saving the password. Customers must know about the attended city of that website before saving the password for banking. This is a technique that will provide convenience to customers by not typing the online banking details each time. By saving passwords from the customers are providing organization to any third party for their passwords. This is an easy way to create fraud

Protection of confidential information

According to Georgiadou (2022), For having a good formatting banking experience, users must be having awareness of common fishing techniques which are used by cybercriminals. There is attention to gathering confidential information of the customer like debit cards or credit card details, usernames, banking details, and passwords by deceptive ways. For protecting the account from phishing attacks, users need to avoid strange messages and calls.

Cybercrime and fraud have become common problems in the context of the banking sector. The main objective of conducting this research is to evaluate the concept of cybercrime in the banking sector. This research will also focus on how transactions can be made more authentic and protected (Qabajeh, Thabtah, and Chiclana, 2018). It will also provide information about what are the different ways by which the banking sector can be made more effective.

Chapter 3: Methodology

This section explains the methodology that covers the research methodology and experiment methodology. This includes the approach used for the quantitative data collection, dataset collection, research, and theoretical approach of the practical implementation (Itoo, and Singh, 2021).

Research Methodology

This research study is based on the issue of fraud detection in the banking system. This research aims to design a cyber security fraud detection system. The secondary research method is used for the data collection for this work. This refers to the collection of quantitative and quantitative data using secondary sources of the data such as research papers, journals, books, and articles. The dataset for this research work is collected from the Kaggle website which is accessible for public use for free. The quantitative research approach is selected in this work. In this, only a theoretical approach is explained for the identification of fraud in the banking sector. The small size dataset is used for the representation of that how data encryption is performed for the protection of the data from cyber-attack in the application area of banking. Sensitive data is involved in the banking sector including credit card details, customer information, and account information. Therefore, it is required to protect such data from digital spying (Sadgali, et al., 2019). The machine learning approach is used for the preparation of the model for the detection of fraud in the baking sector. The main focus area of the work is to detect fraud activities using machine learning (MLA) algorithms. The quantitative research approach is used here as the machine learning algorithms that are used in this work which are KNN, decision tree and the neural network does not work with the qualitative data. The collected dataset is numerical data.

The data set for fraud detection that is collected from the Kaggle website is evaluated using the neural network technique that is part of machine learning algorithms. The supervised learning (SL) method is used for the effective examination of the data that is used in different problems. The supervised learning method refers to the collection of data and then dividing it into a small set of issues (Yousefi, Alaghband, and Garibay, 2019). The main objective is to predict the test label for further utilization. For defining the simulated network for fraud detection, the neural network approach is used here for resolving the cyber security challenges in the banking sector. The neural network can solve complex issues in a suitable form for ensuring cyber security in the sector of banking. The machine learning framework is used. This is helpful in the detection of details for resolving the issue of fraud detection.

Sample collection approach for dataset

Machine learning and other techniques are used for the analysis of data and behavior in the dataset. Suitable data is required to apply the machine learning algorithm and evaluation of results so that fraud detection can be identified in the early stage of the payment process. The dataset is obtained from the Kaggle website which is the online community for accessing a range of datasets. This was developed for providing data related to data science, artificial intelligence, and deep learning. It also offers a wide range of features to perform the desired set of activities on the dataset. The data is collected in the form of an excel sheet in an organized manner. For the effective and accurate evaluation of data, some cleaning techniques must be applied to the data to check the presence of Null values or outliers in the dataset (Carminati, et al., 2020). Data operation techniques are used to clear this. For enhancing the security of the data, two methods are continuous and discrete.

The dataset is the Paysim synthetic dataset of mobile money transactions. The used dataset is only 25% of the original dataset. There are a total of 11 columns in the dataset that is described in the following table.

Table 1 Description of the dataset

Name	**Description**
step	This attribute maps the unit of time. 1 step refers to 1 hour.
type	This refers to a type of transaction that is cash-in, cash-out, payment, debt, and transfer
amount	This refers to the number of transactions in the local currency.
nameOrig	This attribute shows the customer who started the communication.
oldbalanceOrg	This refers to the original balance before the transaction.
newbalanceOrig	This refers to the balance after the transaction.
NameDest	This refers to the ID of the recipient of the process of the transaction.
oldbalanceDest	This refers to the balance of the recipient before the process of the transaction.
newbalanceDest	This refers to the balance of the receiver after the transaction.
isFraud	This refers to the category of transaction. 1 represents the fraudulent transaction and 0 represents the non-fraudulent transaction.
isFlaggedFraud	This refers to the flag on the illegal attempts of transferring the amount of more than 200 in a single transaction.

Theoretical approach for the data analysis

It has been identified from the literature review that with the advancement in communication and information technology, the demand for cyber security in the banking sector has increased. Mostly, solutions are based on machine learning (ML) algorithms and Python language. In the proposed approach, the main focus area is to build an effective classifier model for the detection and labeling of credit card fraud in the banking sector. the complete approach is divided into mainly three phases. The first step includes the explanatory analysis of the dataset. The second step includes the implementation of a machine learning algorithm (MLA) on the dataset. In the final step of the analysis, evaluation and training of the model are done to determine the best model. Different tools are used in this process including sckitlearn, Numpy, and different python libraries such as imblearn and matplotlib for the development of the classification model for the detection of financial fraud (Thompson, Aborisade, and Odeniyi, 2019).

Performing exploratory data analysis

Different python modules are used for the evaluation of the data in the Visual Studio Code name IDE (Integrated Development Environment). The dataset is in the .csv format that is imported on the IDE. The python language code is used for the identification of the null values in the dataset. The occurrence of the class label in the dataset is identified using the python code. Then the matplotlib module is used for plotting the obtained results.

Application of the ML algorithms on the collected dataset

The second step is based on the implementation of the machine learning (ML) algorithms on the dataset. The dataset is divided into the testing dataset along with the training dataset and in the final step, the best performing ML model is selected. Some major varieties of algorithms are **Decision Tree, KNN, and Neural Networks** which have been applied to train the dataset. On the basis of the class imbalance, the problems have been removed in selecting the **Neural Network** which is concluded as the best performing model (Raiter, 2021).

Train and Evaluate Models on the Dataset

The Python technology-based model has been evaluated and tested by attempting in selecting the best performing model. Hence, the model has been trained by the dataset to use the fit () function and also then recoding the prediction which has been made by the model to use predict () function. The score of classifiers for each model has been evaluated by attempting in selecting the best model. The visualization in evaluating the model metrics and confusion matrix has been majorly carried out. This has been stated that the **neural network-based model** has the edge which is over the other machine learning algorithm.

The major problem is the class imbalance in the dataset has been removed so that there can be the adoption of the varieties of techniques that have been oversampling. Hence, the data augmentation is carried out by the minority of the class to the dataset (Varun Kumar, et al., 2020). There is one of the easiest solutions has been doubling up the entries to the minority class. Some new entries have been generated through the replication of the existing data entries. (SMOTE) Synthetic Minority Oversampling Techniques are used commonly carrying out the method for the data augmentation of the minority class.

Making the use of the neural network through the data classification model which is highly indicated as the non-linear method to be applied in the data set. There is a fact that has been implementing more complex than the other identifiers which is the usage of the classifiers for the machine learning-based financial fraud detection that has been gaining traction with each day passing.

SMOTE has been included by the imblearn package which can be resampled and imported of the data that have been carried out by using the train_test_split () through the spilled of 70-30.

Practical approach

The major requirement of different tools has been performing fraud detection in the banking sector. The major toll has been used by reducing cyber security which is machine learning. The machine learning tools have been developed in computers by a different system which has been learned and analyzed by real-life things to analyze the data. Therefore, machine learning has been using different statistical algorithms and models which can be easily analyzed by the data patterns and data. To perform the algorithm of the machine learning one major requirement is to scripting the language which is R python language that has been used by defining and training the functionality of the model (Gyamfi, and Abdulai, 2018).

To define the model of fraud detection, there is the supervised type algorithm which has been used to describe supervised machine learning by sub-categorizing artificial intelligence. This type of model has been distinguished between some parts which have been worked as the training algorithm so that it can easily predict the basic data and its outcome. There is a tool that has been used by a range of different libraries of python scripting languages. The main model of python is Pandas, matplotlib, and Seaborn which has been majorly performing the task. There are some models which have been required to write about the scripting of the model which are:

- Matplotlib
- Pandas
- Seaborn
- NumPy
- SkLearn
- NLTK
- Genism
- Imblearn

These are the model which is the most important to perform the given practical. The (NN) neural network has been performed by making use of python code for the machine learning (ML) model. The neural network (NN) is said to be useful and has been identified by the relation which is between the process of the machine learning (ML) model and different sets of data. This program of the model has been developed to be considered by the different sub-activities of the given model. Through the help of the neural network, there is a system that needs to be understood by the behavior through the normal transaction and also by the history of users. If any certainty has been occurring while sending during the time of money so the machine learning model has been warning about the user payment which shows the negative characteristics about the user which is according to the given complaint.

In the next step, the data have distinguished between the two parts that have been trained and also tested data. So, 80% of the data have been dedicated to the training whereas 20% has been used by the testing model. There is the major requirement of the data which has been visualized by the method by understanding the pattern of the transaction made by the users. The neural network has been used to display the output which is based on the labeled (categorized) data. The main model is giving the result about the accuracy of the model for the evaluation of performance.

Packages and libraries used

Numpy: This is the python library that stands for Numerical Python Library. This is used in linear algebraic operations and multidimensional arrays.

Pandas: This is also a python library that is used as a tool for data manipulation and analysis. This is mainly used for loading and reading the dataset.

Sckitlearn: This python package is mainly used in machine learning models and statistical models.

Keras: This refers to the advanced stage of the API (Application Programming Interface) of the neural network. this is mainly used for the implementation of deep learning algorithms such as RNN and CNN due to the modularity and user-friendliness. This runs on the tensor flow. This is used with the backend running tensor flow for the detection of fraud. This is an excellent choice for training the neural network architecture.

Classification techniques

Decision tree: A decision tree is a popular and powerful tool for predicting and classifying it. The decision tree is known for the flowchart which is like the structure of the tree where every internal node denotes the test to the attribute. Each of the branches represents the outcome of the test and the leaf node or the terminal node that have holding the class label. The Decision Tree has many analogies which have been turning around by influencing the wide implementation of machine learning. This includes both regression and classification. In the decision analysis, the decision tree has been used explicitly and visualized to represent the decision making and its decision. This has been used by the tree-like model for making the decision. There is a widely used tool in data mining that can derive the strategy which can reach a particular goal. This has been used widely by machine learning (Minastireanu, and Mesnita, 2019). Decision trees have been classified by sorting the tree from the root to have some leaves. This has been provided with the classification for some instances. Decision trees have been generated to understand the rules and also perform the classification without the requirement of the computation.

KNN: K- Nearest Neighbor is known for the simplest Machine Learning algorithm which is based on the supervised learning (SL) technique. This algorithm has been assumed to have the resemblance between the new data and new cases which have the available cases and then this can be put into new cases which have different categories. K-NN algorithm has the stores which have the available data to classifier the new data point which is grounded on the similarity. When the new data has appeared, there is an easy way which can classify it into the suit category based on the K-NN algorithm. K-NN algorithm has been used by the Regression that has been classified to be sued by the classification problems. K-NN is known as the non-parametric algorithm that doe not have any assumptions more on the underlying data. This is also known as the lazy leaner algorithm as this does not understand the training set (Soni, 2019). This majorly stored the dataset to the time of classification and then it acts as the dataset. This algorithm has the training part which stores the data to get the new data and after this, it can classify the data into a different class that is more like the new data.

ANN: Artificial Neural Network has been providing the basic as well as the advanced concept of the ANN. Artificial Neural Network (ANN) is the biological inspiration which is a sub-field to the article intelligence modeled which is after the brain. The artificial neural network (ANN) has a computational network that relies on the human neural network so that it can be easily constructed by of the human brain's structure. This is way similar to the human brain which has neurons that are linked to each other. This has neurons that are connected all together and form various layers of the network (Zhang, et al., 2018). These types of neurons can be understood as modes. A Neural network is the type of series of algorithms that have been trying in mimicking the human brain so that it can find the relationship between the different sets of data. This has been used in different use-case that as regression, Image recognition, regression, and many more. The neural network has various layers where each layer have been performing a specific function. There is the complexity of the model that is increased where the number of layers also increases which is described as the multi-layer perceptron.

Evaluation measures

The evaluation of the final result depends on the use of a confusion matrix, recall, precision, and accuracy measures. Different features affect the confusion matrix including the true positive (TP), true negative (TN), false positive (FP), and false negative (FN) rate. These measures are described below;

True positive: This denotes the values in which the actual value and predictive values are 1.

True negative: This mentions the values in which the actual value and predicted values are 0.

False positive: This presents the values in which the actual value is 0 and the predicted value is 1.

False negative: This refers to the values in which the actual value is 1 and the predicted value is 0.

Precision: This measure identifies the number of positive value predictions that belong to the actual class. This shows the quality of the positive prediction.

$$Precision = TP/\,(TP + FP)$$

Recall: This measure identifies the number of positive value predictions (true positives) found in the predicted results.

$$Recall \ = TP/\,(TP + FN)$$

Accuracy: This is an effective measure for the evaluation of classification models. This refers to the number of total correct or accurate predictions among the total quantity of predictions.

$$Accuracy \ = (TP + TN)/\,Total$$

Chapter 4: Results and discussion

Results of the explanatory data analysis

The data does not have any missing values or garbage values, data cleaning is not required. The data analysis is performed as there is high variation in different columns. The normalization helped in the improvisation of the overall accuracy of the machine learning model.

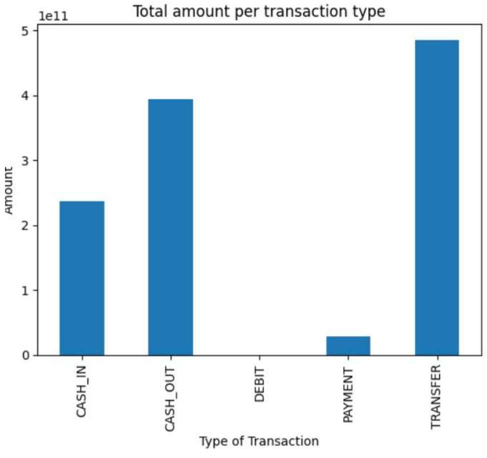


Figure 1 Grouping of the data

The results show that the maximum amount is shared in the TRANSFER and the minimum amount transferred in the DEBT. The CASH_OUT and TRANSFER are the two most commonly used methods of transaction by the customer that mainly act as the attack vector for fraud activities. Therefore, these modes are mainly focused on the analysis.

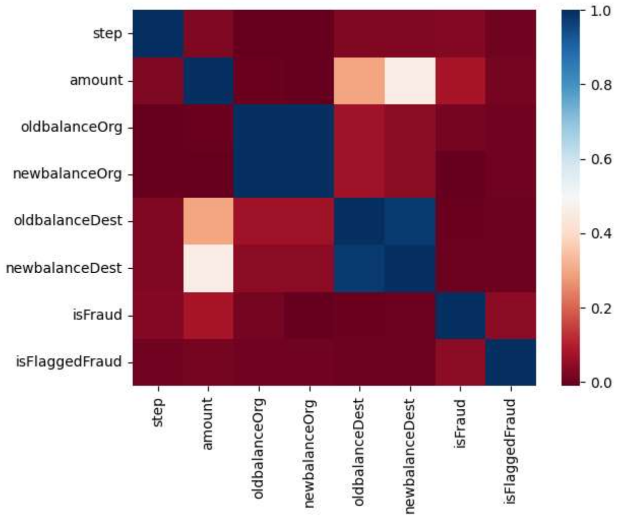


Figure 2 Heatmap between variables

The heatmap given above shows that there is a high correlation between the newbalanceOrg and OldbalanceOrg. Similarly, there is a high correlation between the NewbalanceDest and OldbalanceDest. The Fraud is related to the Amount. For better analysis, separate heatmaps are created of non-fraudulent and fraudulent activities.

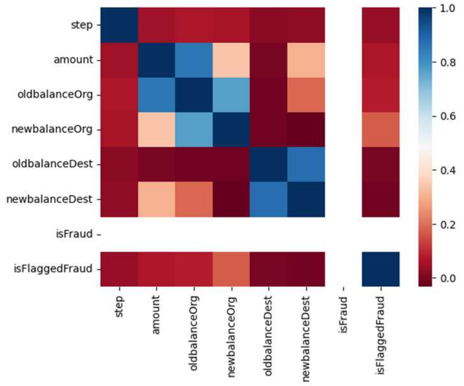


Figure 3 fraud and flagged fraud heatmap

From the above heatmap, it can be interpreted that there is some relation between the isFlaggedFraud and other columns. Therefore, it indicates some relation between isFraud.

From the explanatory analysis, it has been determined that the total number of a fraudulent transaction is 8213, the total number of labeled fraud transactions are 16 and the ratio between the fraud and non-fraud transaction is 1:773. The amount that is lost due to the fraud activity is $12056415427.

From the following graph, it can be determined that the number of transactions that are not flagged correctly is very high, therefore a system is required for fast and reliable detection of fraud.

Results of the machine learning model

In this study, the empirical comparison is conducted between machine learning algorithms and deep learning algorithms. The aim is to investigate the accuracy of different models for the dataset. The following bar graph is created before the machine learning model that the frequency of Not Fraud is very high than the Fraud transactions.

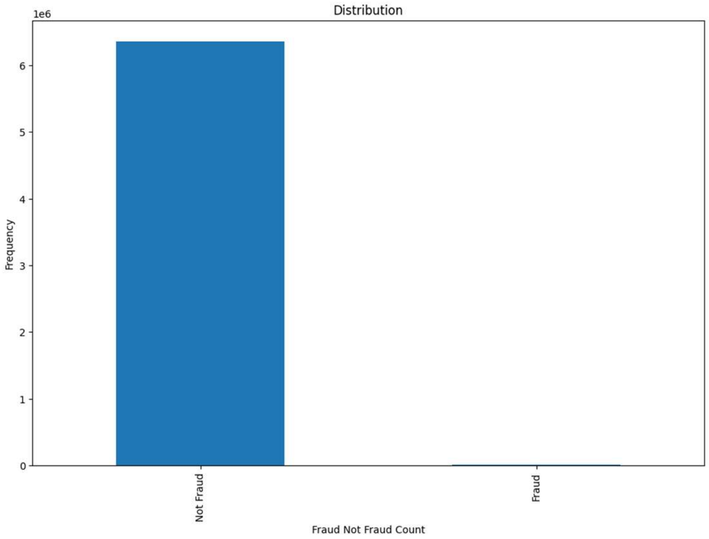
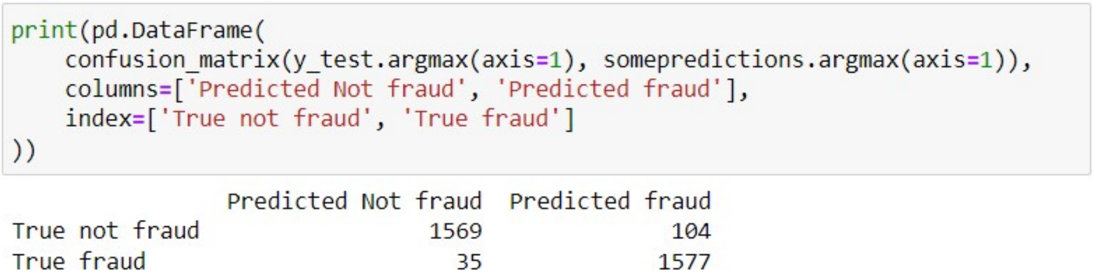


Figure 4Fraud Not Fraud Count bar graph

In the next step, under-sampling is done for Not Fraud. There are a lot of predictions of Not fraud and Fraud classes in the dataset. therefore, it is shuffled, and then scaling is done with the help of a standard scaler for the normalization of data and to avoid bias because the columns such as 'Amount' is having high values that can impact the prediction. Then oneHotEncoder is used on the target for dividing it into the NotFraud(0) and Fraud (1). Then, using Keras, a deep learning model is prepared. 50 percent of the model is used for training and 50 % of the model is used for testing. The evaluation score of the ANN model is 0.958.

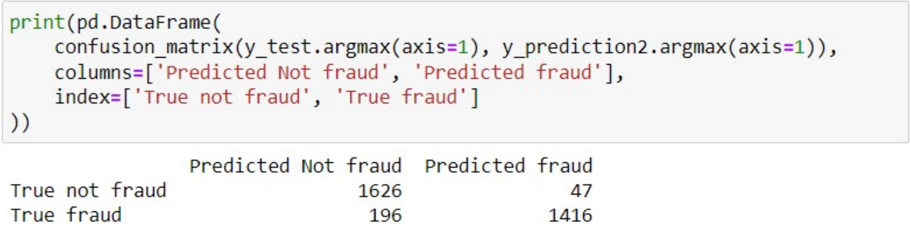


```
print(pd.DataFrame(
    confusion_matrix(y_test.argmax(axis=1), somepredictions.argmax(axis=1)),
    columns=['Predicted Not fraud', 'Predicted fraud'],
    index=['True not fraud', 'True fraud']
))
```

```
                Predicted Not fraud  Predicted fraud
True not fraud                 1569              104
True fraud                       35             1577
```

Figure 5 confusion matrix for CNN

Then in the next step, the K value is determined that will best work for the machine learning model using the KNN algorithm. The obtained value of K is 2. The following image shows the testing accuracy using KNN. Following is the confusion matrix or the KNN algorithm. For the KNN algorithm, a cross-validation approach is applied to the training dataset for the determination of the best K value. The best K value is used for the analysis of the dataset.



```
print(pd.DataFrame(
    confusion_matrix(y_test.argmax(axis=1), y_prediction2.argmax(axis=1)),
    columns=['Predicted Not fraud', 'Predicted fraud'],
    index=['True not fraud', 'True fraud']
))
```

```
                Predicted Not fraud  Predicted fraud
True not fraud                 1626               47
True fraud                      196             1416
```

Figure 6 confusion matrix for KNN

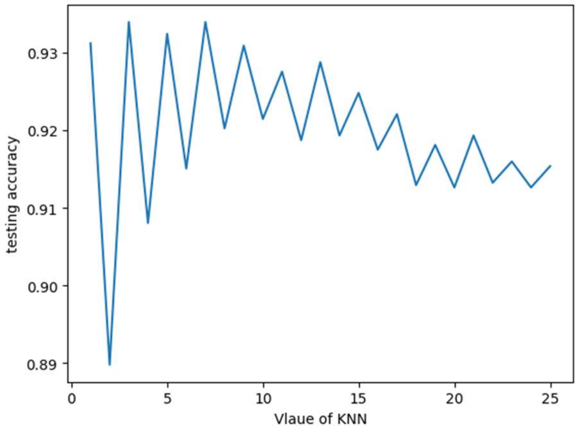
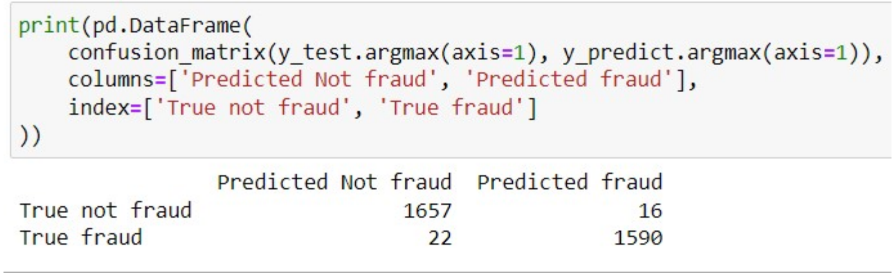


Figure 7 Testing accuracy using KNN

Then, the decision tree model is prepared using sklearn. The following image shows the confusion matrix for the decision tree.



```
print(pd.DataFrame(
    confusion_matrix(y_test.argmax(axis=1), y_predict.argmax(axis=1)),
    columns=['Predicted Not fraud', 'Predicted fraud'],
    index=['True not fraud', 'True fraud']
))
```

```
                Predicted Not fraud  Predicted fraud
True not fraud                 1657               16
True fraud                       22             1590
```

Figure 8 Confusion matrix for decision tree

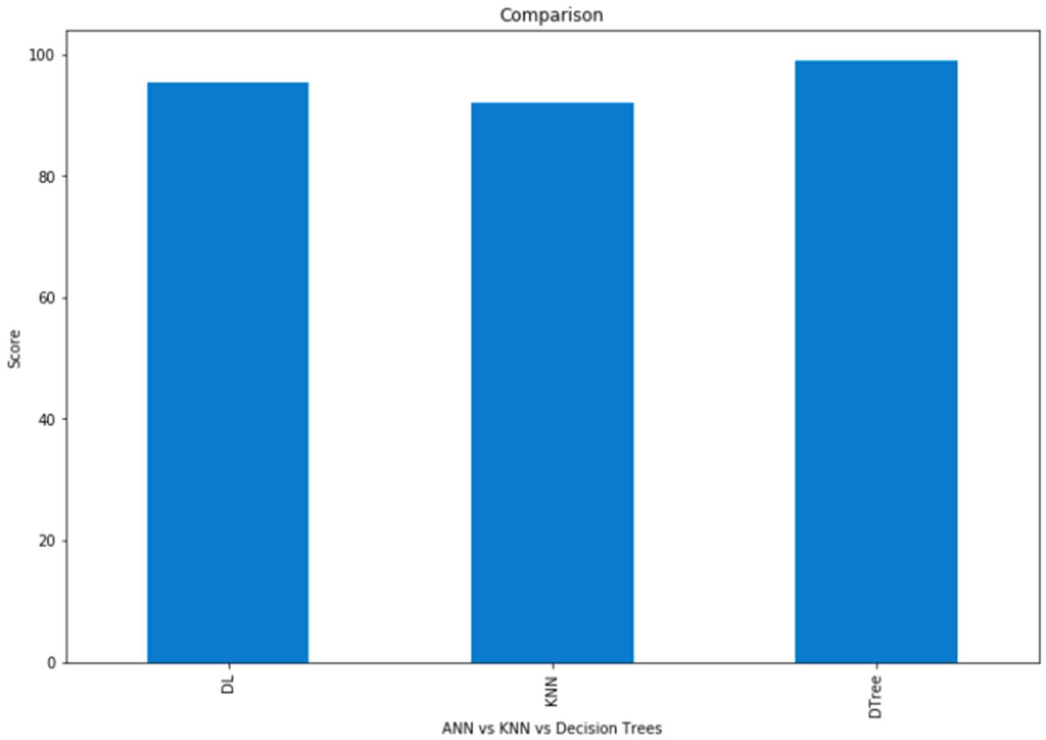


Figure 9 Comparison of ANN vs KNN vs decision tree

The following table presents the results of the different machine learning algorithms based on different performance metrics measures including precision, recall, and accuracy.

Table 2 performance analysis

Name of the ML algorithm	**Precision**	**Recall**	**Accuracy**
K-Nearest Neighbour	97.07	91.35	88.98
Artificial Neural network	94.29	96.24	95.34
Decision Tree	99.06	99.06	99.08

Chapter 5: Conclusion

In this research work, a method is presented for fraud detection in the banking sector that is based on the use of machine learning (ML) algorithms. This paper offers a detailed investigation of the fraud data using the different machine learning algorithms. Deep learning (DL) models are recently used in different application areas for getting higher computation power and low computing cost. In the first step, different machine learning algorithms are compared including KNN, ANN, and decision tree. Then in the final step, the neural network is used for preparing the model for the detection of fraud in the banking sector. In the proposed model, the artificial neural network offered almost 99% accuracy which referred that it is best suitable for fraud detection. This offered more accuracy in comparison to the unsupervised learning (UL) algorithms. This research study covered the pre-processing of the data, normalization, as well as under-sampling to avoid the issues caused by the imbalanced dataset (Soni, 2019). the limitation of this study is that only three machine learning algorithms are used which are decision tree, ANN, and CNN. Fraud patterns are hard to detect and this can be done efficiently using machine learning models.

In future work, other deep learning algorithms can also be used. Specifically, GANs (generative Adversarial Networks) can be used for anomaly detection in banking for the detection of cyber-attacks. The fuzzy systems can also be incorporated for the outlier detection that is used along with the neural network so that better results can be obtained. To increase the performance of the machine learning model, the ensemble learning method and stacking methods can also be used that combine multiple algorithms in a single model for enhancing the accuracy of the model.

REFERENCES

Boutaher, N., Elomri, A., Abghour, N., Moussaid, K. and Rida, M., 2020, November. A review of credit card fraud detection using machine learning techniques. In *2020 5th International Conference on Cloud Computing and Artificial Intelligence: Technologies and Applications (CloudTech)* (pp. 1-5). IEEE.

Carminati, M., Santini, L., Polino, M. and Zanero, S., 2020. Evasion attacks against banking fraud detection systems. In *23rd International Symposium on Research in Attacks, Intrusions, and Defenses (RAID 2020)* (pp. 285-300).

Daliri, S., 2020. Using harmony search algorithm in neural networks to improve fraud detection in banking system. *Computational Intelligence and Neuroscience*, *2020*.

Dornadula, V.N. and Geetha, S., 2019. Credit card fraud detection using machine learning algorithms. *Procedia computer science*, *165*, pp.631-641.

Gyamfi, N.K. and Abdulai, J.D., 2018, November. Bank fraud detection using support vector machine. In *2018 IEEE 9th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)* (pp. 37-41). IEEE.

Ileberi, E., Sun, Y. and Wang, Z., 2022. A machine learning based credit card fraud detection using the GA algorithm for feature selection. *Journal of Big Data*, *9*(1), pp.1-17.

Itoo, F. and Singh, S., 2021. Comparison and analysis of logistic regression, Naïve Bayes and KNN machine learning algorithms for credit card fraud detection. *International Journal of Information Technology*, *13*(4), pp.1503-1511.

Jansen, J. and Van Schaik, P., 2018. Testing a model of precautionary online behavior: The case of online banking. *Computers in Human Behavior*, *87*, pp.371-383.

Koibichuk, V., 2021. Innovation technology and cyber fraud risks of neobanks: gravity model analysis. *684080133*.

Krishna Rao, N.V., Harika Devi, Y., Shalini, N., Harika, A., Divyavani, V. and Mangathayaru, N., 2021. Credit Card Fraud Detection Using Spark and Machine Learning Techniques. In *Machine Learning Technologies and Applications* (pp. 163-172). Springer, Singapore.

Krishna, B., Krishnan, S. and Sebastian, M.P., 2022. Examining the Relationship between National Cybersecurity Commitment, Culture, and Digital Payment Usage: An Institutional Trust Theory Perspective. *Information Systems Frontiers*, pp.1-29.

Kulatilleke, G.K., 2022. Challenges and complexities in machine learning based credit card fraud detection. *arXiv preprint arXiv:2208.10943*.

Leo, M., Sharma, S. and Maddulety, K., 2019. Machine learning in banking risk management: A literature review. *Risks*, *7*(1), p.29.

Lim, K.S., Lee, L.H. and Sim, Y.W., 2021. A review of machine learning algorithms for fraud detection in credit card transaction. *International Journal of Computer Science & Network Security*, *21*(9), pp.31-40.

Minastireanu, E.A. and Mesnita, G., 2019. An Analysis of the Most Used Machine Learning Algorithms for Online Fraud Detection. *Informatica Economica*, *23*(1).

Mittal, S. and Tyagi, S., 2019, January. Performance evaluation of machine learning algorithms for credit card fraud detection. In *2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)* (pp. 320-324). IEEE.

Nasution, M.D.T.P., Rossanty, Y., Siahaan, A.P.U. and Aryza, S., 2018. The phenomenon of cybercrime and fraud victimization in the online shop. *Int. J. Civ. Eng. Technol*, *9*(6), pp.1583-1592.

Nobles, C., 2021. *Banking Cybersecurity Culture Influences on Phishing Susceptibility*. Temple University.

Nyakarimi, S.N., Kariuki, S.N. and Kariuki, P., 2020. Risk assessment and fraud prevention in the banking sector.

Patil, S., Nemade, V. and Soni, P.K., 2018. Predictive modelling for credit card fraud detection using data analytics. *Procedia computer science*, *132*, pp.385-395.

Popat, R.R. and Chaudhary, J., 2018, May. A survey on credit card fraud detection using machine learning. In *2018 2nd international conference on trends in electronics and informatics (ICOEI)* (pp. 1120-1125). IEEE.

Pradesyah, R., Yuslem, N. and Batubara, C., 2021, November. Fraud In Financial Institutions. In *Journal Of International Conference Proceedings (Jicp)* (Vol. 4, No. 2, pp. 341-348).

Priya, G.J. and Saradha, S., 2021, February. Fraud detection and prevention using machine learning algorithms: a review. In *2021 7th International Conference on Electrical Energy Systems (ICEES)* (pp. 564-568). IEEE.

Qabajeh, I., Thabtah, F. and Chiclana, F., 2018. A recent review of conventional vs. automated cybersecurity anti-phishing techniques. *Computer Science Review*, *29*, pp.44-55.

Raiter, O., 2021. Applying Supervised Machine Learning Algorithms for Fraud Detection in Anti-Money Laundering. *Journal of Modern Issues in Business Research*, *1*(1), pp.14-26.

Rodrigues, A.R.D., Ferreira, F.A., Teixeira, F.J. and Zopounidis, C., 2022. Artificial intelligence, digital transformation and cybersecurity in the banking sector: A multi-stakeholder cognition-driven framework. *Research in International Business and Finance*, *60*, p.101616.

Rout, M., 2021. Analysis and comparison of credit card fraud detection using machine learning. In *Artificial Intelligence and Machine Learning in Business Management* (pp. 81-93). CRC Press.

Roy, N.C. and Prabhakaran, S., 2022. Sustainable response system building against insider-led cyber frauds in banking sector: a machine learning approach. *Journal of Financial Crime*.

Ryman-Tubb, N.F., Krause, P. and Garn, W., 2018. How Artificial Intelligence and machine learning research impacts payment card fraud detection: A survey and industry benchmark. *Engineering Applications of Artificial Intelligence*, *76*, pp.130-157.

Sadgali, I., Nawal, S.A.E.L. and Benabbou, F., 2019, October. Fraud detection in credit card transaction using machine learning techniques. In *2019 1st International Conference on Smart Systems and Data Science (ICSSD)* (pp. 1-4). IEEE.

Shukur, H.A. and Kurnaz, S., 2019. Credit card fraud detection using machine learning methodology. *International Journal of Computer Science and Mobile Computing*, *8*(3), pp.257-260.

Singh, A. and Jain, A., 2020. An empirical study of aml approach for credit card fraud detection–financial transactions. *International Journal of Computers Communications & Control*, *14*(6), pp.670-690.

Soni, V.D., 2019. Role of Artificial Intelligence in Combating Cyber Threats in Banking. *International Engineering Journal For Research & Development*, *4*(1), pp.7-7.

Soomro, Z.A., Ahmed, J., Shah, M.H. and Khoumbati, K., 2019. Investigating identity fraud management practices in e-tail sector: a systematic review. *Journal of Enterprise Information Management*.

Susto, G.A., Terzi, M., Masiero, C., Pampuri, S. and Schirru, A., 2018, September. A fraud detection decision support system via human on-line behavior characterization and machine learning. In *2018 First International Conference on Artificial Intelligence for Industries (AI4I)* (pp. 9-14). IEEE.

Tariq, N., 2018. Impact of cyberattacks on financial institutions. *Journal of Internet Banking and Commerce*, *23*(2), pp.1-11.

Thennakoon, A., Bhagyani, C., Premadasa, S., Mihiranga, S. and Kuruwitaarachchi, N., 2019, January. Real-time credit card fraud detection using machine learning. In *2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)* (pp. 488-493). IEEE.

Thompson, A., Aborisade, L. and Odeniyi, E., 2019. A Fraud Detection Framework using Machine Learning Approach. In *The Fourth International Conference on Cyber Technology and Cyber Systems, 2019.*.

Tiwari, P., Mehta, S., Sakhuja, N., Kumar, J. and Singh, A.K., 2021. Credit Card Fraud Detection using Machine Learning: A Study. *arXiv preprint arXiv:2108.10005*.

Varun Kumar, K.S., Vijaya Kumar, V.G., Vijay Shankar, A. and Pratibha, K., 2020. Credit Card Fraud Detection using Machine Learning Algorithms. *INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume*, *9*.

Yang, W., Zhang, Y., Ye, K., Li, L. and Xu, C.Z., 2019, June. Ffd: A federated learning based method for credit card fraud detection. In *International conference on big data* (pp. 18-32). Springer, Cham.

Yousefi, N., Alaghband, M. and Garibay, I., 2019. A comprehensive survey on machine learning techniques and user authentication approaches for credit card fraud detection. *arXiv preprint arXiv:1912.02629*.

Zhang, Z., Zhou, X., Zhang, X., Wang, L., and Wang, P., 2018. A model based on a convolutional neural network for online transaction fraud detection. *Security and Communication Networks*, *2018*.